FOURTH EDITION

GLOBAL HEALTH

Diseases, Programs, Systems, and Policies

Michael H. Merson Robert E. Black Anne J. Mills

FOURTH EDITION

GLOBAL HEALTH

Diseases, Programs, Systems, and Policies

Michael H. Merson, MD
Wolfgang Joklik Professor of Global Health
Duke Global Health Institute
Duke University
Durham, North Carolina

Robert E. Black, MD, MPH
Professor and Director
Institute for International Programs
Department of International Health
Johns Hopkins Bloomberg School of Public Health
Baltimore, Maryland

Anne J. Mills, PhD
Professor and Deputy Director and Provost
London School of Hygiene & Tropical Medicine
London, UK

JONES & BARTLETT
LEARNING

World Headquarters
Jones & Bartlett Learning
5 Wall Street
Burlington, MA 01803
978-443-5000
info@jblearning.com
www.jblearning.com

Jones & Bartlett Learning books and products are available through most bookstores and online booksellers. To contact Jones & Bartlett Learning directly, call 800-832-0034, fax 978-443-8000, or visit our website, www.jblearning.com.

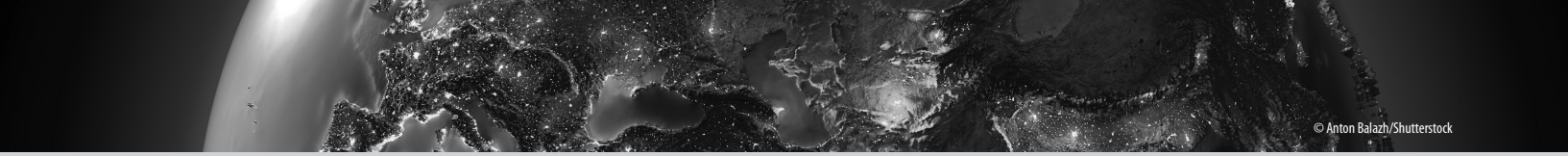
Substantial discounts on bulk quantities of Jones & Bartlett Learning publications are available to corporations, professional associations, and other qualified organizations. For details and specific discount information, contact the special sales department at Jones & Bartlett Learning via the above contact information or send an email to specialsales@jblearning.com.

Production Credits
VP, Product Management: David D. Cella
Director of Product Management: Michael Brown
Product Manager: Sophie Fleck Teague
Product Specialist: Danielle Bessette
Production Editor: Kelly Sylvester
Senior Marketing Manager: Susanne Walker
Manufacturing and Inventory Control Supervisor: Amy Bacus
Composition: codeMantra U.S. LLC
Cover Design: Scott Moden
Rights & Media Specialist: Merideth Tumasz
Media Development Editor: Shannon Sheehan
Cover Image (Title Page, Chapter Opener):
© Anton Balazh/Shutterstock
Printing and Binding: Sheridan Books
Cover Printing: Sheridan Books

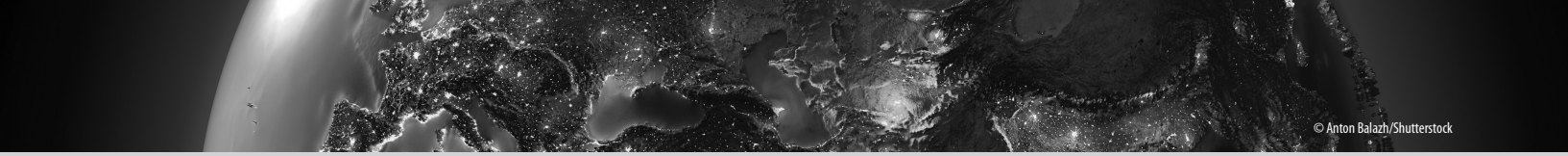
Library of Congress Cataloging-in-Publication Data
Names: Merson, Michael H., editor. | Black, Robert E., editor. | Mills, Anne, editor.
Title: Global health: diseases, programs, systems, and policies / [edited] by Michael H. Merson, Robert E. Black, and Anne J. Mills.
Other titles: Global health (Merson)
Description: Fourth edition. | Burlington, Massachusetts: Jones & Bartlett Learning, [2020] | Includes bibliographical references and index.
Identifiers: LCCN 2018023789 | ISBN 9781284122626 (casebound)
Subjects: | MESH: Global Health | Public Health | International Cooperation | Public Health Administration | Health Policy
Classification: LCC RA441 | NLM WA 530.1 | DDC 362.1—dc23
LC record available at https://lccn.loc.gov/2018023789

6048

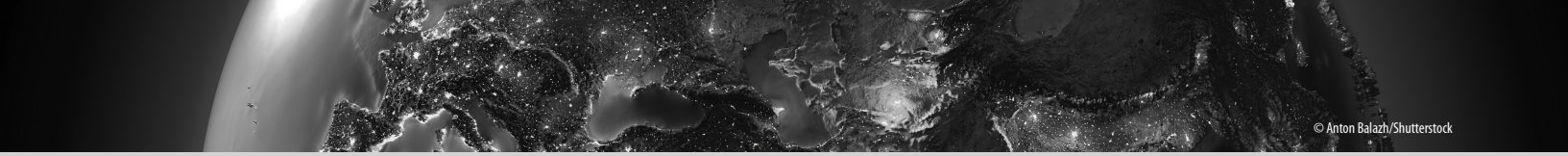
Printed in the United States of America
22 21 20 19 18 10 9 8 7 6 5 4 3 2 1



Brief Contents

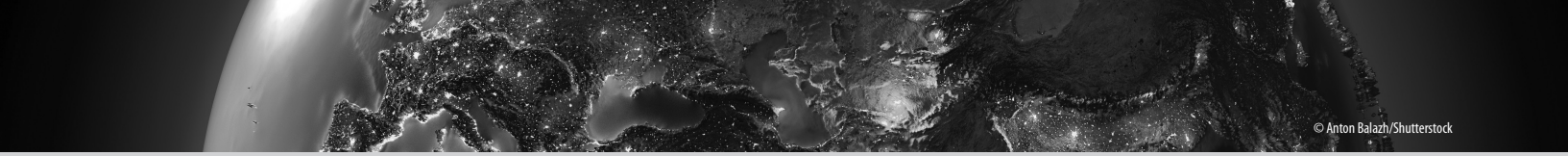


Contents



Acknowledgments

We would like to acknowledge the technical and editorial assistance of Michelle Pender and Rachel Genego.



Contributors

Stuart Anderson, MA, PhD, FRPharmS, FHEA
Emeritus Professor of History of Pharmacy, London School of Hygiene & Tropical Medicine
London, United Kingdom

Abdulgafoor M. Bachani, PhD, MHS
Assistant Professor, International Health
Director, Johns Hopkins International Injury Research Unit
Health Systems Program, Department of International Health
Johns Hopkins Bloomberg School of Public Health
Baltimore, Maryland

David E. Bloom, MA, PhD
Clarence James Gamble Professor of Economics and Demography
Chair, Department of Global Health and Population
Harvard School of Public Health
Boston, Massachusetts

Chris Beyrer, MD, MPH
Desmond M. Tutu Professor and Director
Center for Public Health and Human Rights
Johns Hopkins Bloomberg School of Public Health
Baltimore, Maryland

Alexander Butchart, MA, PhD
Coordinator, Violence Prevention
Department for Management of Noncommunicable Diseases, Disability, Violence and Injury Prevention
World Health Organization
Geneva, Switzerland

David Canning, PhD
Professor of Economics and International Health
Department of Global Health and Population
Harvard School of Public Health
Boston, Massachusetts

Rupa Chanda
RBI Chair Professor in Economics
Indian Institute of Management
Bangalore, India

Gillian P. Christie, MPhil, MA
Doctor of Public Health Candidate
Harvard T.H. Chan School of Public Health
Cambridge, Massachusetts

Alex Cohen, MA, PhD
Visiting Scientist
Department of Epidemiology
Harvard T.H. Chan School of Public Health
Boston, Massachusetts
&
Honorary Associate Professor
Department of Population Health
London School of Hygiene & Tropical Medicine
London, United Kingdom

Osman A. Dar, FRCP Edin, FFPH
Consultant in Global Health
Global Public Health Division
Public Health England
London, United Kingdom

Jessica Fanzo, PhD
Bloomberg Distinguished Associate Professor
Department of International Health
Bloomberg School of Public Health
Berman Institute of Bioethics
Johns Hopkins University
Baltimore, Maryland

Andrew Green, BA, MA, PhD
Professor Emeritus
Nuffield Centre for International Health and Development, Leeds Institute of Health Sciences
University of Leeds
Leeds, United Kingdom

Aubree Gordon, PhD
Assistant Professor
Department of Epidemiology
School of Public Health
University of Michigan
Ann Arbor, Michigan

Johanna Hanefeld, BA, MSc, PhD
Associate Professor
Health Policy and Systems Research

London School of Hygiene and Tropical Medicine
London, United Kingdom

Kara Hanson, MPhil, SD
Professor of Health System Economics
Faculty of Public Health and Policy
London School of Hygiene and
Tropical Medicine
London, United Kingdom

David J. Heller, MD, MPH
Assistant Professor
Arnhold Institute for Global Health
Icahn School of Medicine at Mount Sinai
New York, New York

David L. Heymann, MD
Professor of Infectious Disease Epidemiology
London School of Hygiene & Tropical Medicine
London, United Kingdom

Adnan A. Hyder, MD, MPH, PhD
Professor & Associate Chair, International Health
Director, Health Systems Program, International
Health
Director, Johns Hopkins International Injury
Research Unit
Associate Director, Johns Hopkins Berman Institute
of Bioethics
Johns Hopkins Bloomberg School of Public Health
Baltimore, Maryland

Brady Hunt
PhD Candidate
Richards-Kortum Laboratory
Department of Bioengineering
Rice University
Houston, Texas

Dean T. Jamison, MS, PhD
Professor, Global Health
Adjunct Professor, Health Services
Institute for Health Metrics and Evaluation
Department of Global Health
University of Washington
Seattle, Washington

Benjamin Johns, PhD
Senior Associate / Scientist
International Development Division
Abt Associates, Inc.
Rockville, Maryland

Sandeep P. Kishore, MD, PhD, MSc
Associate Director
Arnhold Institute for Global Health
Icahn School of Medicine at Mount Sinai
New York, New York

Nancy Kass, ScD
Vice Provost for Graduate and Professional
Education
Phoebe R. Berman Professor of Bioethics and Public
Health
Deputy Director of Public Health in the Berman
Institute of Bioethics
Johns Hopkins University
Baltimore, Maryland

Berit Kieselbach, MSc, MPH
Technical Officer, Violence Prevention
Department for Management of Noncommunicable
Diseases, Disability, Violence and
Injury Prevention
World Health Organization
Geneva, Switzerland

Tord Kjellstrom, MEng, PhD (Med)
Visiting Fellow
Australian National University
Canberra, Australia

Brandon A. Kohrt, MD, PhD
Charles and Sonia Akman Associate Professor in
Global Psychiatry Department of Psychiatry and
Behavioral Sciences
The George Washington University
Washington, DC

Sandra D. Lane, PhD, MPH
Laura J. and L. Douglas Meredith Professor
Professor of Public Health and Anthropology
Syracuse University
&
Research Professor
Obstetrics and Gynecology
Upstate Medical University
Syracuse, New York

Kelley Lee, BA MPA, MA, DPhil, FFPH, FCAHS
Professor
Faculty of Health Sciences
Simon Fraser University
Burnaby, British Columbia, Canada

Siobhan Leir, MSc
Research Assistant in Global Public Health
London School of Hygiene & Tropical Medicine
London, United Kingdom

Henry Lishi Li, MA, MSc
PhD Candidate
Department of Health Services Research and Policy
London School of Hygiene &
Tropical Medicine
London, United Kingdom

Tonya Littlejohn, MBBS, MA
Affiliated Scholar
Center for Public Health and Human Rights
Johns Hopkins Bloomberg School of Public Health
Baltimore, Maryland

Rene Loewenson, PhD, FRSPH
Director
Training and Research Support Centre
Harare, Zimbabwe

Crick Lund, PhD
Professor of Global Mental Health and Development
Centre for Global Mental Health
Institute of Psychiatry, Psychology, and Neuroscience
King's College London
London, United Kingdom
&
Professor of Public Mental Health
Alan J. Flisher Centre for Public Mental Health
Department of Psychiatry and Mental Health
University of Cape Town
Cape Town, South Africa

Alexandra Macmillan, MBChB, MPH, PhD, FNZCPHM
Senior Lecturer Environmental Health
University of Otago
Dunedin, New Zealand

Melisa Martinez-Alvarez, BSc, MSc, PhD
Senior Scientist
MRC Unit at The Gambia
London School of Hygiene & Tropical Medicine

Barbara McPake, PhD
Professor of Health Economics and
Director, Nossal Institute for Global Health
Melbourne School of Population and Global Health
University of Melbourne
Melbourne, Australia

Jane Menken, MS, PhD
Distinguished Professor of Sociology
Research Professor, Institute of Behavioral Science
University of Colorado, Boulder
Boulder, Colorado

Dipak Kumar Mitra, MBBS, MPH, PhD
Associate Professor of Epidemiology
School of Public Health
Independent University, Bangladesh (IUB)
Dhaka, Bangladesh

Tolib Mirzoev, MD, MA, PGCertLTHE, PhD
Associate Professor
Nuffield Centre for International Health and Development, Leeds Institute of Health Sciences
University of Leeds
Leeds, United Kingdom

Melinda K. Munos, PhD
Assistant Professor
Department of International Health
Bloomberg School of Public Health
Johns Hopkins University
Baltimore, Maryland

Mary Natoli
PhD Candidate
Richards-Kortum Laboratory
Department of Bioengineering
Rice University
Houston, Texas

Maria Oden
Teaching Professor
Department of Bioengineering
Rice University
Houston, Texas

Marsha Orgill, BAdmin (Hons), MPhil
Senior Researcher
Health Policy and Systems Division, School of Public Health and Family Medicine
University of Cape Town
Cape Town, South Africa

Benjamin Palafox, MSc
Research Fellow, Pharmaceutical Policy and Economics
Department of Global Health and Development
London School of Hygiene & Tropical Medicine
London, United Kingdom

Vikram Patel, MSc, MRCPsych, PhD, FMedSci
The Pershing Square Professor of Global Health and Wellcome Trust Principal Research Fellow
Harvard University
Boston, Massachusetts

Margaret Peden, BSc Nurs, BSc Med Hons, PhD
Senior Research Fellow & Head of Injury Programme
The George Institute for Global Health
Oxford University
Oxford, United Kingdom

Ajay Pillarisetti, MPH, PhD
Lecturer and Postdoctoral Scholar
Environmental Health Sciences
University of California, Berkeley
Berkeley, California

Jennifer Prah Ruger, BA, MA, MSc, PhD, MSL
Amartya Sen Professor of Health Equity, Economics, and Policy
School of Social Policy & Practice
University of Pennsylvania
Philadelphia, Pennsylvania

M. Omar Rahman, MD, MPH, DSc
Vice Chancellor and Dean
School of Public Health
Independent University, Bangladesh (IUB)
Dhaka, Bangladesh

Michael Kent Ranson, MD, MPH, PhD
Senior Economist (Health)
The World Bank
Washington, DC

Arthur Reingold, MD
Professor and Division Head, Epidemiology and Biostatistics
School of Public Health
University of California, Berkeley
Berkeley, California

Eric Richardson
Associate Professor of Practice
Department of Biomedical Engineering
Duke University
Durham, North Carolina

Rebecca Richards-Kortum, PhD
Malcom Gillis University Professor
Department of Bioengineering
Rice University
Houston, Texas

Sakib Rokadiya, MRCP, DTMH
Specialist Registrar in Infectious Diseases & General Medicine
Barts Health NHS Trust
London, United Kingdom

Giuliano Russo
Lecturer in Global Health
Centre for Global Public Health
Queen Mary University of London
London, United Kingdom

Susan C. Scrimshaw, PhD
Nevin Scrimshaw International Nutrition Foundation, Tufts University, & Former President of The Sage Colleges and of Simmons College
Thornton, New Hampshire

Rima Shretta, MSc, PhD
Associate Director
Malaria Elimination Initiative
Global Health Group
University of California, San Francisco
San Francisco, California

Sarah Simpson
Conjoint Lecturer
UNSW Sydney
Sydney, Australia

Director
EquiACT
Montluel, France

Kirk R. Smith, MPH, PhD
Professor Global Environmental Health
University of California, Berkeley
Berkeley, California

Richard Smith, BA, MSc, PhD, HonMFPH
Professor of Health System Economics
Dean of Faculty of Public Health and Policy
London School of Hygiene and Tropical Medicine
London, United Kingdom

Christine P. Stewart, MPH, PhD
Associate Professor
Department of Nutrition
Program in International and Community Nutrition
University of California, Davis
Davis, California

Michael J. Toole, BMedSc, MBBS, DTM&H
Principal Research Fellow
Burnet Institute
Professor, School of Public Health
Monash University
Melbourne, Australia

Cesar G. Victora, MD, PhD
Emeritus Professor of Epidemiology
Federal University of Pelotas in Brazil
Pelotas, Brazil

Ronald J. Waldman, MD, MPH
Professor of Global Health
Milken Institute School of Public Health
The George Washington University
Washington, DC

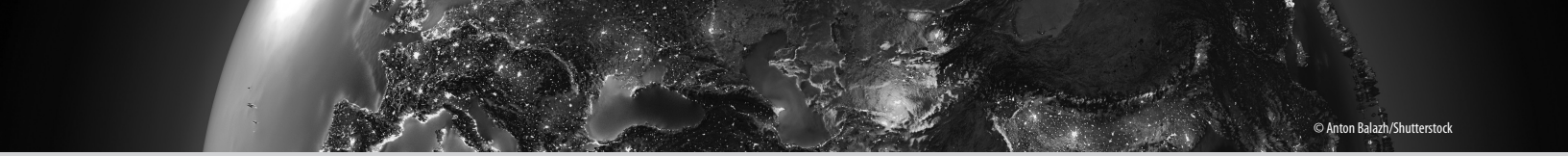
Keith P. West, Jr., DrPH, MPH
George G. Graham Professor, Infant and Child Nutrition
Department of International Health
Bloomberg School of Public Health
Johns Hopkins University
Baltimore, Maryland

Harvey Whiteford, MBBS, MPH, PhD
Professor of Population Mental Health
School of Public Health, The University of Queensland
Brisbane, Australia

James Whitworth, MD, FMedSci
Professor of International Public Health
London School of Hygiene & Tropical Medicine
London, United Kingdom

Alistair Woodward, MB BS, MMedSci, PhD, FNZCPHM
Professor Epidemiology and Biostatistics
University of Auckland
Auckland, New Zealand

Derek Yach, MBChB, MPH
Chief Health Officer
The Vitality Group
New York, New York



About the Authors

Michael H. Merson, MD, is the Wolfgang Joklik Professor of Global Health at Duke University. He joined the Duke faculty in November 2006 as the founding director of the Duke Global Health Institute and served in that role through 2017. In addition, Dr. Merson was Vice President and Vice Provost for Global Affairs from 2011 until 2018, and Vice Chancellor for Duke-National University of Singapore Affairs from 2010 until 2016.

Dr. Merson has held leadership positions at the Centers for Disease Control and Prevention and the International Center for Diarrheal Diseases Research, Bangladesh; served as Director of the World Health Organization's Diarrheal and Acute Respiratory Control Programs and Global Program on AIDS; and was the first Dean of Public Health at Yale University. Dr. Merson's research and writings have been primarily on the etiology of diarrheal diseases in low- and middle-income countries, AIDS prevention and policy, and academic global health. He has served as an advisor to a number of international organizations and advisory bodies; is an elected member of the National Academy of Medicine; and has received two honorary degrees and the U.S. Surgeon General's Exemplary Service Medal.

Robert E. Black, MD, MPH, is Professor and Director of the Institute for International Programs of the Johns Hopkins Bloomberg School of Public Health in Baltimore, Maryland. Dr. Black is trained in medicine, infectious diseases, and epidemiology. He has served as a medical epidemiologist at the Centers for Disease Control and Prevention and worked at institutions in Bangladesh and Peru on research related to childhood infectious diseases and nutritional problems. He was Chair of the Department of International Health of the Bloomberg School of Public Health from 1985 to 2013.

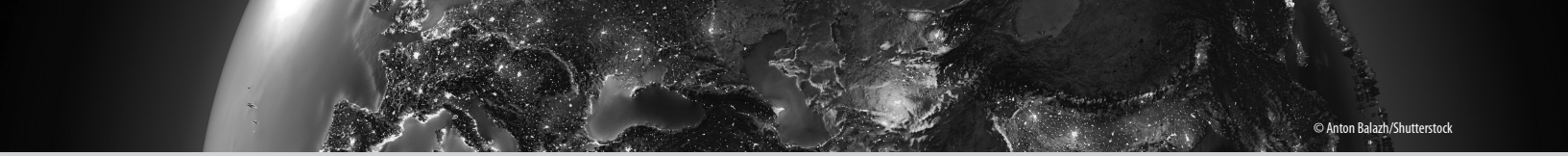
Dr. Black's current research includes field trials of vaccines, micronutrients, and other interventions; effectiveness studies of health programs; and evaluation of preventive and curative health service programs in low- and middle-income countries. In the last 20 years, he has led work that demonstrated the benefits of zinc supplements in prevention and treatment of childhood diarrhea and pneumonia. His other interests are related to the use of evidence in policy and programs, including estimates of the causes of child mortality, the development of research capacity, and the strengthening of public health training.

As a member of the U.S. National Academy of Medicine and advisory bodies of the World Health Organization, the International Center for Diarrheal Diseases Research, Bangladesh, and other international organizations, Dr. Black assists with the development of research and policies intended to improve child health. He chaired the Child Health and Nutrition Research Initiative and serves on the governing boards of Nutrition International and Vitamin Angels. He has more than 700 scientific journal publications. Dr. Black is the recipient of the Programme for Global Paediatric Research Award for Outstanding Contributions to Global Child Health in 2010, the Prince Mahidol Award for Public Health in 2010, the Canada Gairdner Global Health Award in 2011, the Nutrition Leadership Award from Sight and Life in 2013, and the Jimmy and Rosalynn Carter Humanitarian Award in 2016.

Anne J. Mills, PhD, is Deputy Director and Provost of the London School of Hygiene & Tropical Medicine (LSHTM), and Professor of Health Economics and Policy. She has degrees from Oxford University (MA), Leeds University (Diploma), and London University (PhD). After two years as Overseas Development Institute Fellow and Economist in the Ministry of Health in Malawi, she spent three years at the University of Leeds studying health planning in the National Health Service; she moved to the LSHTM in 1979.

Dr. Mills has researched and published widely in the fields of health economics and health systems in low- and middle-income countries and continues to be involved in research on financial protection in South Africa, Tanzania, India, and Thailand. She has had continuing involvement in supporting capacity development in health economics in universities, research institutes, and governments. She has been

involved in numerous policy initiatives, including the World Health Organization's Commission on Macroeconomics and Health. She is a Foreign Associate of the U.S. Institute of Medicine and a Fellow of the U.K. Academy of Medical Sciences. In 2009, Dr. Mills received the Prince Mahidol Award in the field of medicine. In 2013, she was elected a Fellow of the Royal Society; and in 2015, she was made a Dame in the Queen's New Year's Honours, for services to international health.



Foreword

The philosophy, skills, and knowledge essential for global health practice continue to improve with experience. The objective, of course, is the improvement of health for everyone in the world, or global health equity. The means to achieve that objective require harnessing every resource and skill possible in every country. When it comes to health, no country is more important than any other.

The United States has direct experience with the returns of global investment in health. When President Lyndon Johnson provided funds for global smallpox eradication in 1966, the United States had not had a case of smallpox for 17 years. It would have been easy to say this was not an American problem. Yet the investment to eliminate smallpox in the rest of the world had immediate financial benefits because treating Americans with adverse effects caused by the vaccine was no longer necessary. We did not have to tolerate 6 to 8 deaths each year due to the vaccine. And savings were realized in foreign quarantine procedures as well. This country recoups its investment in smallpox eradication every 3 months—more than 150-fold, to date, and it will continue forever, an infinite benefit–cost ratio.

Because of the increasing number of global health programs, health benefits continue to accrue for every country. More than 35,000 deaths are averted each day in children younger than the age of 5, as compared to 60 years ago, when my interest in global health was kindled. In addition to smallpox, we stand at the brink of eradicating Guinea worm disease and polio. The number of measles deaths has declined by more than 95% and, according to the World Bank, 250,000 people move out of poverty every day. It is a miraculous time for global health.

And science marches on, in ways that promise even more benefits in the next decade. Immunization rates are improving, new vaccines are becoming available, and the world is slowly moving to address chronic diseases, environmental health problems, and injury control in larger geographic areas. Even the health problems associated with global warming now engage many countries, including the largest polluters of the environment. Evolution is evident in humanity's responses to its global health problems.

No global health decision stands alone. Instead, such decisions are always dependent on political decisions and compromises. It is difficult enough to gain agreement on a local level, where groups are acquainted and share many experiences. It is many times more difficult to achieve agreement across political and cultural divides. Global health workers are forced to learn diplomatic skills and to understand the art of cross-cultural transactions. Despite such difficulties, some of my best professional memories involve meetings, discussions, and field activities involving people of various cultures and languages. Successful approaches thrive on the shared knowledge of people with different life experiences.

For all of the improvements, there are, nevertheless, areas of concern. The rapid expansion of academic programs in global health presents practical problems for students. We need more teachers who have actually worked in resource-poor situations. I attended a superb tropical public health program 50 years ago, but not a single faculty member had actually lived in tropical areas, although they were frequent consultants to field programs. The training of pediatricians, surgeons, or internists involves hands-on experiences. Likewise, we need to increase the ability of students to have problem-solving experiences in poor areas, while they are still in school.

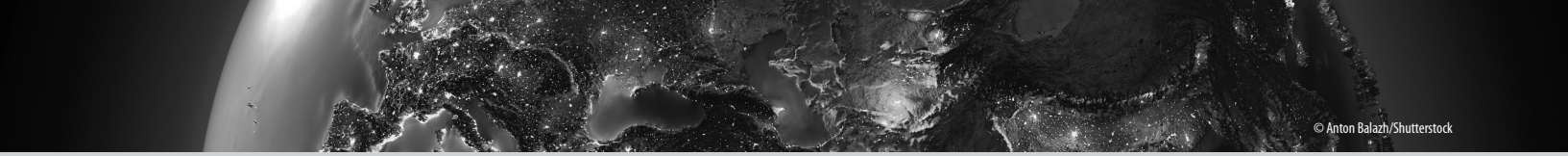
This text also addresses the very real shortcomings of how global health is organized. We are often critical of the World Health Organization (WHO) and other agencies without acknowledging that we organized them to be dysfunctional. WHO has regional offices that can undercut the WHO headquarters decisions because the United States insisted, 70 years ago, on strong regional offices in its attempt to protect the Pan American Health Organization. We imposed the burden of 195 health ministers as WHO's board of directors, and annually tell WHO to reduce its budget. Then we condemn WHO for an inadequate response to the Ebola virus. It is time to ask what has been learned in 70 years and how we can use that experience to

improve our global organizations to better protect the health of all.

Finally, WHO promoted primary health care at the Alma Ata conference 40 years ago, yet we are still struggling to implement village-level interventions and better understand the social determinants of health. This year the world will again come together at a conference in Kazakhstan to recommit to strengthening primary health care in order to achieve universal health coverage and the Sustainable Development Goals. No risk factor is as great as poverty. Nevertheless, while progress *is* being made and the World Bank has championed the cause, deliberate action to reduce poverty is still an underdeveloped tool in the daily lives of health departments in every country. Moreover, keeping people healthy helps to keep them productive. The world will not be a great place for any of us until it is a great place for all of us.

Drs. Merson, Black, and Mills have again brought us a comprehensive text that helps to make sense of an enlarging and challenging field. It is an enormously valuable guide for students, teachers, practitioners, and researchers alike, which helps to map out the journeys of those who will change the future and make our planet a healthier one.

William Foege
Professor Emeritus, Rollins School of Public Health,
Emory University



Introduction

Michael H. Merson, Robert E. Black, and Anne J. Mills

The three of us are privileged to serve as faculty at universities that provide education every year to hundreds of graduate and undergraduate students motivated to learn about global health issues and challenges. Many of these students plan to or have already begun careers in global health research, policy, practice, teaching, or administration. This text is written for these students around the world, as well as for those who teach and mentor them. In this Introduction, we define global health, provide a brief history of the field, and summarize its many accomplishments and challenges. We then explain how we put this *Fourth Edition* together and how we think it can best be used.

What Is Global Health?

Global health is a burgeoning field, which has seen a major surge of interest as an area of academic study in the last decade. Essentially, global health has replaced international public health in both concept and reality. International public health focuses on the application of the principles of public health to health problems and challenges that affect low- and middle-income countries (LMICs) and to the complex array of global and local determinants that influence them. Global health maintains this focus, but places much greater emphasis on health issues that concern many countries or that are affected greatly by transnational determinants, such as climate change or urbanization. This greater emphasis on the scope and location of health problems provides the opportunity to address cross-border issues as well as domestic health disparities in high-income countries.

While international public health primarily applies the principles of public health, there is now agreement that success and progress in improving health around the world requires a multidisciplinary and interdisciplinary approach that includes, yet extends beyond, public health. Professionals from many disciplines and academic fields possess the skills and knowledge needed to understand the various determinants of health and develop strategies that will address these determinants, thereby sharing goals to improve the health of populations. These disciplines and professional fields include social and behavioral sciences (including sociology, economics, psychology, anthropology, political science, and international relations), biomedical and environmental sciences, engineering, business and management, public policy, law, history, and divinity. Furthermore, while efforts to reduce health disparities should focus on prevention, treatment, care, and curative strategies must also be addressed when developing solutions to global health challenges. This call for multipronged action further emphasizes the need for a multidisciplinary approach.

In addition, while social justice must continue to be a central pillar of health, the approach to achieving health equity and finding solutions to reducing health disparities must now much more strongly emphasize global cooperation. Rather than following a model that transfers ideas and resources from high-income countries, organizations, or funding agencies to low- and middle-income settings, it is imperative to pursue "a real partnership, a pooling of experience and knowledge, and a two-way flow between developed and developing countries" when implementing health interventions or programs (Koplan et al., 2009, p. 1995).

Today we live in an increasingly connected world, but the challenges to reduce health disparities are considerable, and the tenets of global health provide a unique insight and strategic approach to addressing them. Given this evolution in our thinking, there has been an understandable interest in defining global health. In 1997, the U.S. Institute of Medicine (IOM) released a report that broadly defined global health as "health problems, issues, and concerns that transcend national boundaries, may be influenced by circumstances or experiences in other countries, and are best addressed by cooperative actions and solutions" (Board on International Health, 1997, p. 1). More than 10 years later, IOM amended its definition, describing

global health "not just as a state but also as the *goal of improving health for all people by reducing avoidable disease, disabilities, and deaths*" (Committee on the U.S. Commitment to Global Health, 2009, p. 5).

Although a number of other definitions for global health have been proposed, we prefer the definition of global health that was adopted by the Consortium of Universities for Global Health (CUGH). CUGH was formed to promote, facilitate, and enhance the growth of global health as an academic field of study. It has defined global health as follows:

> [A]n area for study, research, and practice that places a priority on improving health and achieving equity in health for all people worldwide. Global health emphasizes transnational health issues, determinants, and solutions; involves many disciplines within and beyond the health sciences and promotes interdisciplinary collaboration; and is a synthesis of population-based prevention with individual-level clinical care. (Koplan et al., 2009, p. 1995)

When providing this definition, an effort was made to explain the differences between public health, international health, and global health. While these terms certainly share areas of overlap, this comparison helps to draw out global health's distinctive qualities. (**EXHIBIT I-1**).

A Brief History of Global Health

Tracing the roots of global health brings us to the history of international public health. This history encompasses the origins of public health and can be viewed as the story of how populations experience health and illness; how social, economic, and political systems create the possibilities for healthy or unhealthy lives; how societies create the preconditions for the production and transmission of disease; and how people, both as individuals and as social groups, attempt to promote their own health or avoid illness (Rosen & Morman, 1993). A number of authors have documented this history (Arnold, 1988; Basch, 1999; Leff & Leff, 1958; Rosen & Morman, 1993; Winslow & Hallock, 1933). A brief history is presented here primarily to provide a perspective for the challenges that face us today (**EXHIBIT I-2**).

The Origins of Public Health

It is difficult to select a date for the origins of the field of public health. Some would begin with Hippocrates, whose book *Airs, Waters, and Places*, published around

EXHIBIT I-1 Global Health, International Health, and Public Health

Global Health	International Health	Public Health
Focuses on issues that directly or indirectly impact health but can transcend national boundaries.	Focuses on health issues of countries *other* than one's own, especially those of LMICs.	Focuses on issues that impact the health of the *population* of a particular community or nation.
Development and implementation of solutions often require global cooperation.	Development and implementation of solutions usually involve binational cooperation.	Development and implementation of solutions usually do not involve global cooperation.
Embraces both prevention in populations and clinical care of individuals.	Embraces both prevention in populations and clinical care of individuals.	Mainly focused on prevention programs for populations.
Health equity among nations and for all people is a major objective.	Seeks to help people of other nations.	Health equity within a nation or community is a major objective.
Highly interdisciplinary and multidisciplinary within and beyond health sciences.	Embraces but has not emphasized multidisciplinarity.	Encourages multidisciplinary approaches, particularly within health sciences and with social sciences.

EXHIBIT I-2 History of Global Health: A Summary

400 BC: Hippocrates presents the causal relationship between environment and disease.

First century AD: Romans introduce public sanitation and organize a water supply system.

14th century: The "Black Death" (bubonic plague) leads to quarantine and *cordon sanitaire*.

Middle Ages: Colonial expansion spreads infectious diseases around the world.

1750–1850: The Industrial Revolution results in extensive health and social improvements in cities in Europe and the United States.

1850–1910: Knowledge about the causes and transmission of communicable diseases is greatly expanded.

1910–1945: Significant reductions in child mortality occur. Schools of public health and international foundations and intergovernmental agencies interested in public health are established.

1945–1990: The World Bank, World Health Organization (WHO), and other United Nations agencies are created. WHO eradicates smallpox. The Alma Ata conference gives emphasis to primary health care. The HIV/AIDS pandemic begins. The United Nations Children's Fund (UNICEF) leads efforts to ensure universal childhood immunization. Greater attention is given to noncommunicable diseases (NCDs).

1990–2000: Priority is given to health-sector reform, the impact of and responses to globalization, cost-effectiveness, and public–private partnerships in health.

2000–2010: Priority is given to equity, social determinants of health, health and development, use of innovative information and communications technologies, declaration of the Millennium Development Goals (MDGs), and response to influenza.

2010–2020: Priority is given to climate change's impact on health, growing burden of NCDs, increasing threat of emerging and re-emerging diseases, global health security, universal health coverage, proliferation of innovative technologies for delivery of prevention and care services, and the declaration of the Sustainable Development Goals (SDGs).

400 BC, was the first systematic effort to present the causal relationships between environmental factors and disease and offer a theoretical basis for an understanding of endemic and epidemic diseases. Others would cite the introduction of public sanitation and an organized water supply system by the Romans in the first century AD. Many would select the bubonic plague ("Black Death") pandemic of the 14th century, which began in Central Asia; was carried on ships to Constantinople, Genoa, and other European ports; and then spread inland, killing 25 million persons in Europe alone. In responding to this devastating infectious disease, the Great Council of the city of Ragusa (now Dubrovnik, Croatia) followed a contagion theory, which recommended the separation of healthy and sick populations; it issued a document stating that outsiders entering the city must spend 30 days in the restricted location of nearby islands (Stuard & NetLibrary, 1992). The length of time for this isolation period, dubbed *trentino*, was eventually increased from 30 to 40 days, introducing the concept of the modern quarantine (Gensini, Yacoub, & Conti, 2004).

The Middle Ages was also the period when many cities in Europe, particularly through the formation of guilds, took an active part in establishing hospitals and other institutions to provide medical care and social assistance. At the same time, many European countries began to expand their horizons abroad, by exploring and colonizing new lands. The travelers brought some diseases with them (e.g., influenza, measles, smallpox), and those who settled in these colonial outposts were forced to confront diseases that had never been seen in Europe (such as syphilis, dysentery, malaria, and sleeping sickness). European explorers also carried pathogens from one part of Africa to another, and from one area of the globe to another (e.g., from Africa to North America through the slave trade). On long voyages, the greatest enemy of the sailor was often scurvy—at least until 1875, when the British government issued its famous order that all men-of-war should carry a supply of lemon juice as a preventive measure.

The Age of Enlightenment (1750–1830) was a pivotal period in the evolution of public health. It was a time of social action in relation to health, as reflected by the new interest taken in the health problems of specific population groups. During this period, rapid advances in technology led to the development of factories. In England and elsewhere, this industrialization was paralleled by expansion of the coal mines. The Industrial Revolution had arrived. During this period,

sanitaire the populations of the cities of England and other industrialized nations grew enormously, with overcrowded, unsanitary conditions in these urban areas leading to outbreaks of cholera and other epidemic diseases, which ultimately resulted in high rates of child mortality. Near the end of this period, significant efforts were made to address these problems. Improvements were made in urban water supplies and sewerage systems, municipal hospitals arose throughout cities in Europe and the east coast of the United States, laws were enacted limiting children's ability to work, and data on deaths and births began to be systematically collected in many places.

As industrialization continued, it became obvious that more efforts to protect the health of the public were needed. These changes occurred first in England, regarded as the first modern industrial country, through the efforts of the noted social reformer Edwin Chadwick. Beginning in 1832, he headed up the royal Poor Law Commission, which undertook an extensive survey of health and sanitation conditions throughout the country. The work of this commission led in 1848 to the Public Health Act, which created a General Board of Health that was empowered to appoint local boards of health and medical officers of health to deal effectively with public health problems. The impact of these developments was felt throughout Europe and especially in the United States, where it stimulated creation of health departments in many cities and states.

Cholera, which in the first half of the 19th century spread in waves from South Asia to the Middle East and then to Europe and the United States, did the most to stimulate the formal internationalization of public health. The policy of establishing a *cordon sanitaire*—an action applied by many European nations in an effort to control the disease—had become a major restraint on trade, necessitating an international agreement. In 1851, the First International Sanitary Conference was convened in Paris to discuss the role of quarantine in the control of cholera, plague and yellow fever, which were causing epidemics throughout Europe. Although no real agreement was reached, the conference laid the foundations for international cooperation in health.

The latter part of the 19th century was distinguished by the enormous growth of knowledge in the area of microbiology, as exemplified by Louis Pasteur's proof of the germ theory of disease, Robert Koch's discovery of the tubercle bacillus, and Walter Reed's demonstration of the role of the mosquito in transmitting yellow fever. Between 1880 and 1910, the etiologic causes and means of transmission of many communicable diseases were discovered in laboratories in North America and Europe. The development of this knowledge base was paralleled by related discoveries in the sciences of physiology, metabolism, endocrinology, and nutrition. Dramatic decreases were soon seen in child and adult mortality thanks to improvements in social and economic conditions, discovery of vaccines, and implementation of programs in health education. The way was now clear for the development of public health administration based on a scientific understanding of the principles involved in the transmission of communicable diseases.

The first two decades of the 20th century witnessed the establishment of three formal intergovernmental public health bodies: the International Sanitary Bureau to serve nations in the western hemisphere (in 1904); l'Office Internationale d'Hygiene Publique in Paris, which was concerned with prevention and control of the main quarantinable diseases (in 1909); and the League of Nations Health Office (LNHO) in Geneva, Switzerland, which provided assistance to countries on technical matters related to health (in 1920). In 1926, LNHO commenced publication of *Weekly Epidemiological Record*, which evolved into a weekly publication of the World Health Organization (WHO) and still is published today. LNHO also established many scientific and technical commissions, issued reports on the status of many infectious and chronic diseases, and sent its staff around the world to assist national governments in dealing with their health problems.

In North America and countries in Europe, the explosion of scientific knowledge in the latter part of the 19th century and the belief that social problems could be solved stimulated universities, such as Johns Hopkins, to establish schools of public health. In France, public subscriptions helped to fund the Institut Pasteur (named in honor of Louis Pasteur) in Paris, which subsequently developed a network of institutes throughout the francophone world that produced sera and vaccines and conducted research on a wide variety of tropical diseases. Another significant development during this period was the founding of the Rockefeller Foundation (in 1909) and its International Health Commission (in 1913). During its 38 years of operation, the commission cooperated with many governments in campaigns against endemic diseases such as hookworm, malaria, and yellow fever. The Rockefeller Foundation also provided essential financial support to help establish medical and public health schools around the world; and later international health programs in a number of American and European

schools of medicine and public health. All of these developments were paralleled by the development and strengthening of competencies in public health among the militaries of the United States and the countries of Europe, stimulated in great part by the buildup to and realities of World War I. Following the war, there was increasing recognition that much ill health in the colonial world was not easily solvable with medical interventions alone, but rather was intractably linked to malnutrition and poverty.

Some historians would date the beginning of international public health to the end of World War II. The ending of European colonialism, the need to reconstruct the economies of the United States and the countries of Western Europe, and the rapid emergence of newly independent countries in Africa and Asia were all forces that led to the creation of many new intergovernmental organizations. The United Nations Monetary and Financial Conference, held in Bretton Woods, New Hampshire, in 1944 and attended by representatives from 43 countries, resulted in the establishment of the International Bank for Reconstruction and Development (more commonly known as the World Bank) and the International Monetary Fund. The former initially lent money to countries only at prevailing market interest rates, but in 1960 it began to provide loans to poorer countries at much lower interest rates and with far better terms through its International Development Association. It was not until the early 1980s, however, that the World Bank began to accelerate greatly its provision of loans to countries for programs in health and education. By the end of that decade, these loans had become the greatest source of foreign assistance to LMICs (Ruger, 2005).

In the decade after World War II, many other United Nations organizations (e.g., UNICEF) and specialized agencies (such as WHO) were formed to assist countries in strengthening their health, social, and economic sectors. In addition, most of the wealthier industrialized countries established agencies or bureaus that funded bilateral projects in specific LMICs. Among the historical colonial powers, such assistance was most often provided to their former colonies.

Many of the international health efforts in the 1960s and 1970s were dedicated to the control of specific diseases. A global effort to control malaria was hampered by a number of operational and technical difficulties, including the vector's increasing resistance to insecticides and the parasite's resistance to available antimalarial drugs. In contrast, the campaign to eradicate smallpox, led by WHO, successfully eliminated the disease in 1981 and stimulated the establishment of the Expanded Program on Immunization, which focused on the delivery of effective vaccines to infants. Also, during the 1970s, two large international research programs were initiated under the co-sponsorship of various United Nations agencies: The Special Program for Research on Human Reproduction (focusing on development and testing of new contraceptive technologies) and the Tropical Disease Research Program (providing support for the development of better means of diagnosis, treatment, and prevention of six tropical diseases, including malaria). Greater attention also was gradually given to chronic diseases, commonly known as noncommunicable diseases (NCDs), such as cardiovascular and cerebrovascular diseases and cancer.

In 1978, WHO organized a conference in Alma Ata in the former Soviet Union that prioritized the delivery of primary healthcare services and set the goal of "health for all by the year 2000." Rather than focusing solely on control of specific diseases, this conference called for international efforts to strengthen the capacities of LMICs to extend their health services to populations with poor access to prevention and care. The concerns of tropical medicine, which were concentrated on the infectious diseases of warm climates, were replaced by an emphasis on the provision of health services to reduce morbidity and premature mortality in resource-poor settings (De Cock, Lucas, Mabey, & Parry, 1995). Given the limited financial and managerial capacities of many governments, increased attention was paid to the role of nongovernmental organizations (NGOs) in providing these services. As a result, many mission hospitals, particularly in sub-Saharan Africa, expanded their activities in their local communities, the number of local NGOs began to increase, and a number of international NGOs (e.g., Save the Children, Oxfam, Médecins Sans Frontières) greatly expanded their services, often with support from bilateral agencies. Disease-specific efforts—most notably UNICEF's Child Survival Program, with its acronym GOBI (growth charts, oral rehydration, breastfeeding, immunization) and its goal of universal childhood immunization by 1990—were seen by many as programs that both focused on specific health problems and provided a means of strengthening health systems.

The emergence of what is sometimes called "the new public health" was heralded by the Ottawa Charter of 1986, which was meant to provide a plan of action to achieve the "health for all" targets set forth at Alma Ata. The Ottawa Charter pioneered the

definition of health as a resource for development, rather than merely a desirable outcome of development. The prerequisites for health that were outlined in the charter were diverse and included peace, shelter, education, food, income, a stable ecosystem, sustainable resources, social justice, and equity. Moreover, the charter emphasized the importance of structural factors that affect health on a societal level, rather than focusing only on the risk behaviors of individuals. It called on the worldwide health community to address health disparities by engaging and enabling people to take charge of their health at community and policy-making levels. This shift from a "risk behavior" focus to an emphasis on "risk environment" continues to resonate in contemporary public health practice and research.

One hugely influential development in the 1980s was the onset of the human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) pandemic. By the time a simple laboratory test to detect HIV was discovered in 1985, more than 2 million persons in sub-Saharan Africa had been infected. In 1987, WHO formed the Global Programme on AIDS, which within 2 years became the largest international public health effort ever established, with an annual budget of $90 million and 500 staff working in Geneva, Switzerland, and in more than 80 LMICs. In 1995, with some 20 million persons infected with HIV (mostly living in LMICs), and with the understanding that the pandemic could be brought under control only through a multisectoral effort, the program was transformed into a joint effort of UN agencies known as the Joint United Nations Programme on HIV/AIDS (UNAIDS). The global response to HIV/AIDS helped to shape the field of global health from its emphasis on intersectoral collaboration, health and human rights, global advocacy for health, and focus on prevention and treatment.

The Origin and Growth of Global Health

The end of the Cold War ushered in dramatic changes that stimulated the development of the new concept of global health. Major shifts in political and economic ideologies led to a reconsideration of the role of governments, including how they should finance and deliver public services. Greater attention was given to increasing the role of civil society and the private sector in achieving universal health coverage. Indeed, global health as it relates to health systems in the last decade of the 20th century and the first two decades of the 21st century can be characterized by its emphasis on health-sector reform, cost-effectiveness as an important principle in the choice of interventions, and public–private partnerships in health, paralleled by a rapid expansion of innovative technologies.

Although rising incomes have long been known to improve health status, increased attention has been paid to the relationship between health and poverty, and the importance of a healthy population for achieving economic development. Participation of sectors other than the health sector is now viewed as essential for achieving a healthy population. More and more countries are experiencing the demographic transition to societies with rapidly increasing numbers of middle-aged and older adults, and in turn are being challenged with providing preventive and care services that address health problems of both the poor and the wealthy simultaneously. Increasing life expectancy, urbanization, and resultant changing lifestyles have contributed to an ever-increasing burden of NCDs. India and China, for example, now have high rates of cardiovascular disease, stroke, and diabetes. Not surprisingly, issues regarding equity in the availability of drugs and vaccines and in access to other technological advances have drawn greater attention. Healthy populations are also now viewed as essential for domestic security.

The first decade of the 21st century witnessed the addition of new multifaceted and complex issues to the list of global health challenges—among them, human migration and displacement, bioterrorism, emerging pathogens, climate change, and disaster preparedness. It was within this context that the United Nations General Assembly adopted the Millennium Declaration in September 2000 as a set of guiding principles and key objectives for international cooperation. The declaration underscored the need to address inequities that have been created or worsened by globalization, and to form new international linkages to achieve and protect peace, disarmament, poverty eradication, gender equality, a healthy environment, human rights, and good governance. The goals dealing specifically with development and poverty eradication become known as the Millennium Development Goals (MDGs); three of them pertained primarily to health (shown in bold in **EXHIBIT I-3**). In addition, Goal 1 included reduction in childhood undernutrition. All 191 member states of the UN pledged to meet the MDGs by 2015.

Building on the achievements of the MDGs (covered later), as well as learning from their limitations, the United Nations adopted the 2030 Agenda for Sustainable Development in 2015. Central to this agenda are the SDGs which comprise 17 goals, 169 associated

EXHIBIT I-3 Millennium Development Goals

1. Reduce extreme poverty and hunger by one-half
2. Achieve universal primary education
3. Promote gender equality and empower women
4. **Reduce under-5 mortality by two-thirds**
5. **Reduce maternal mortality by three-fourths**
6. **Reverse the spread of HIV/AIDS, malaria, tuberculosis, and other major diseases**
7. Ensure environmental sustainability
8. Develop a global partnership for development, with targets for aid, trade, and debt relief

EXHIBIT I-4 Sustainable Development Goals

Goal 1: End poverty in all its forms everywhere
Goal 2: End hunger, achieve food security and improved nutrition, and promote sustainable agriculture
Goal 3: Ensure healthy lives and promote well-being for all at all ages
Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all
Goal 5: Achieve gender equality and empower all women and girls
Goal 6: Ensure availability and sustainable management of water and sanitation for all
Goal 7: Ensure access to affordable, reliable, sustainable, and modern energy for all
Goal 8: Promote sustained, inclusive, and sustainable economic growth, full and productive employment, and decent work for all
Goal 9: Build resilient infrastructure, promote inclusive and sustainable industrialization, and foster innovation
Goal 10: Reduce inequality within and among countries
Goal 11: Make cities and human settlements inclusive, safe, resilient, and sustainable
Goal 12: Ensure sustainable consumption and production patterns
Goal 13: Take urgent action to combat climate change and its impacts
Goal 14: Conserve and sustainably use the oceans, seas, and marine resources for sustainable development
Goal 15: Protect, restore, and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, halt and reverse land degradation, and halt biodiversity loss
Goal 16: Promote peaceful and inclusive societies for sustainable development, provide access to justice for all, and build effective, accountable, and inclusive institutions at all levels
Goal 17: Strengthen the means of implementation and revitalize the global partnership for sustainable development

targets, and 230 indicators to guide global sustainable development through to 2030; the goals are shown in **EXHIBIT I-4**. The SDGs seek a more integrated approach to address the challenges of an increasingly global and integrated world. While there is only one health-specific goal—Goal 3: "Good health and well-being"—it has 13 associated targets and there are many linkages between health and the other SDGs and associated targets. For example, Goal 2 includes elimination of nutritional deficiencies. Achieving the SDGs will require consistent monitoring and evaluation and both global and national commitment.

In the years since the previous edition of this text was published, there have been a number of noteworthy successes in global health, partly due to the MDGs. Notably, substantial progress was made in achieving the three health-related MDG goals cited earlier. First, mortality among children younger than age 5 dropped 56%, from 12.6 million deaths in 1990 to 5.6 million deaths in 2016 (WHO, 2017a). Second, maternal mortality worldwide decreased by 44% between 1990 and 2015 (WHO, 2016b). Third, there was an almost 26% reduction in malaria deaths between 2006 and 2016 (Global Burden of Disease [GBD] Causes of Death Collaborators, 2017). In addition, in 2016, the number of people who were infected annually with HIV declined to 1.8 million—a more than 50% reduction since 1996, when there were 3.5 million new infections. The number of AIDS-related deaths fell by 48% (to 1 million deaths) since the peak in 2005 (UNAIDS, 2017a. In addition, for the first time since the eradication of smallpox, we are

on the verge of eliminating another major global disease from the world: 2017 saw the lowest case count of polio in recorded history (Polio Global Eradication Initiative, 2017).

The recent successes in fighting malaria and HIV/AIDS are attributable in great part to the expansion of access to treatment, financed primarily by the Global Fund to Fight AIDS, Tuberculosis and Malaria (the Global Fund), and the President's Emergency Plan for AIDS Relief (PEPFAR). Moreover, the Global Fund's performance-based funding and decision-making processes have made important contributions to the practice of aid, particularly in encouraging management for results, participation of civil society, mutual accountability, and broad-based country and local ownership.

Current Challenges in Global Health

We have witnessed major improvements in the health of populations over the past century, with the pace of change increasing rapidly in LMICs since the Bretton Woods Conference. Global health—and, more broadly, an improved understanding of how social, behavioral, economic, and environmental factors influence the health of populations—has contributed to these improvements to a greater extent than expanded access to medical care. Nevertheless, these improvements have not been universal, disparities between rich and poor both between and within countries remain, and the challenges of global health have never been greater.

Despite recent progress, we still have far to go in terms of maternal and child health. Millions of children still die before reaching the age of 5 due to diseases that could be simply and affordably prevented and treated (UNICEF, 2016). Hundreds of thousands of women continue to die annually from preventable complications of pregnancy and childbirth, and most of these deaths occur in LMICs (WHO, 2016b).

> Infectious diseases—once thought to have been vanquished as major killers—have emerged or reemerged around the world as top threats to health and well-being. Despite recent progress, the AIDS pandemic is far from over. In particular, prevention efforts need to be targeted toward vulnerable populations who are still at high risk for HIV/AIDS, such as girls and young women in sub-Saharan Africa—a population that accounts for more than 70% of new HIV infections among adolescents. (USAID, 2017)

We have seen infectious diseases travel from endemic regions to previously unaffected areas (e.g., Zika virus), and others newly emerge. The 2013-2016 Ebola outbreak in West Africa was an alarming wake-up call regarding the danger of epidemics in urban areas and the frail state of our global health security. The underlying causes of many emerging infectious diseases can be traced to human-initiated social and environmental changes, including climatic and ecosystem disturbances, trends in food consumption and production, close proximity of humans and animals in household settings, and unsafe medical practices (Kuiken, Fouchier, Rimmelzwaan, & Osterhaus, 2003). This relationship has given rise to the disciplines of One Health, an approach that seeks to address and mitigate the effects on health arising from the interfaces between humans, animals, and environments (Gibbs, 2014) and Planetary Health, which studies the health of human civilizations and the systems—political, economic, and social—on which they depend (The Lancet Planetary Health, 2017). Furthermore, globalization forces, including increased trade and movement of people, have led to far greater opportunities for infectious disease to spread around the world quickly, as evidenced by the severe acute respiratory syndrome (SARS) epidemic in 2002–2003. In 2012, an estimated 12.6 million deaths were attributable to environmental factors, many of which were related to effects of climate change (Pruss-Ustun, Wolf, Corvalan, Bos, & Neira, 2016). Other health consequences of climate change and environmental biodegradation will be experienced through increased water and food insecurity, extreme climactic events, displaced populations, and vulnerable human settlements (Jamison et al., 2013). Resistance to antibiotics is also rising among populations around the globe, making infections harder and more expensive to treat and threatening the gains we have made in combatting infectious diseases.

NCDs were once considered a problem afflicting only high-income nations whose populations had achieved long life expectancies. Today, NCDs are the leading causes of death worldwide; they accounted for more than 70% of global deaths in 2016, an increase of 16% since 2006 (GBD Causes of Death Collaborators, 2017). Ischemic heart disease and stroke are the leading causes of death globally, accounting for 15 million deaths in 2015 (WHO, 2017b). The rise of

NCDs is especially apparent in LMICs, where 78% of global NCD deaths occurred in 2015 (WHO, 2017b). Globalizing forces that have imported Western lifestyle habits, such as increased trade and trade liberalization, tobacco use, and increased consumption of processed foods, have fueled these disease trends. In addition, despite the high burden imposed by mental illness, and especially depressive disorders, and the high percentage of global disability and mortality attributable to them, only recently have these disorders been given the attention they deserve (Summergrad, 2016). As a result of the overall increase in all types of NCDs and the lingering problem of infectious diseases, many LMICs face a double burden of disease, putting more strain on their already frail healthcare systems.

The importance of improving the performance of health systems to achieve reductions in mortality and morbidity has become widely accepted, including the need to address the global health workforce crisis (WHO, 2006). Health systems need adequate resources to meet the changing and growing needs of the populations they serve. An estimated 40 to 50 million new health and social care workers will be needed by 2030 to reach the SDGs, 18 million of whom are needed in LMICs (WHO, 2016a). WHO has identified the critical need to strengthen health systems so as to fight poverty, foster development, and maintain and improve the health of people around the world. It has set the goal of achieving universal health coverage, meaning that all persons can access needed health services of sufficient quality to be effective and not cause financial hardship. As the world faces increasing threats from emerging and re-emerging infectious diseases, strong basic healthcare systems will also be essential to support global health security and to avoid crippling national and global pandemics (Horton, 2018).

One means for expanding access to health services has been the use of mobile phone technology. By 2021, 5.5 billion people are predicted to have access to mobile phones, and more people in Africa will have access to mobile phones than to running water (Thornton, 2017). Mobile phone initiatives are now aimed at improving healthcare services in many countries, as they are increasingly being used for disaster management, reminders for people to get vaccinations, health screening tests, and social marketing. Emerging drone technology also has been harnessed to improve access to health care by transporting blood, contraceptive products, and medical supplies to remote areas, or during natural disasters. These technologies, and others like them, will surely play a pivotal role in the future of global health.

In recent years, the growth in the number of refugees and displaced persons around the world has been startling. By the middle of June 2017, more than 65 million people had been forced to leave their homes (United Nations High Commissioner for Refugees, 2017)— the highest number recorded in 70 years (Mohammadi, 2016). In addition to humanitarian and economic repercussions, this population upheaval puts displaced populations at increased risk for both infectious and NCDs due to overcrowding, poor sanitation, and lack of access to health care, and overburdens the health systems in the refugees' host countries.

There is a broad consensus that poverty is the most important underlying cause of preventable death, disease, and disability on a global level. While the number of people living in poverty remains unacceptably high, progress has been made in reducing the global poverty rate: 10.7% of the world population lived below the poverty line in 2013, compared to 42% in 1981 (*The Economist*, 2017). This progress has been partly due to advances made in literacy, access to housing, safe water, sanitation, food supplies, and urbanization. Even so, 50% of the world's extremely poor live in sub-Saharan Africa (Hollenhorst, 2016), reflecting the particular challenges of social and economic development in that region.

Meeting the global health challenges we face today and in the future will require new forms of financing and cooperation. Changing global economies have altered traditional global health funding mechanisms. Contributions to health assistance from donor countries have more or less flattened in recent years, while LMICs are rightfully taking on more responsibility for financing their domestic health sector due in part to the increasing economic growth in these countries. In 2015, for example, domestic resources accounted for 57% of total HIV/AIDS funding (UNAIDS, 2017b). However, many low-income countries are far from being able to raise enough domestic revenue to replace development assistance for health. International and intersectoral cooperation between UN agencies with an established health role, other international bodies such as the World Trade Organization, regional bodies such as the European Union, bilateral agencies, NGOs, foundations, and the private sector, including pharmaceutical companies, will need to be enhanced. Global organizations such as the Bill & Melinda Gates Foundation, the Global Fund, and GAVI, the Vaccine Alliance, have taken on significant leadership roles and in many cases have surpassed the older,

more traditional actors in global health governance. These organizations have not only injected significant amounts of funds into the global system, but also brought a new, more informal and personal style of operation. Ensuring the ideal structure, effective functioning, and financing of this global health system will itself be an enormous challenge for the next decade of global health (National Academies of Sciences, Engineering, and Medicine, 2017).

Use and Content of This Text

This text has been prepared with future global health challenges foremost in mind. Its focus is on diseases, programs, health systems, and health policies in LMICs, making reference to and using examples from the United States, Western Europe, and other high-income countries as appropriate.[1] Individual chapters present information on health issues that transcend national boundaries and are of concern to many countries.

Our intent has been, first and foremost, to provide a text for graduate students from various disciplines and professions who are studying global health. Given its broad range of content, the text as a whole may serve as the main source for an introductory graduate course on global health. Experience with the previous editions has shown that it also can be used as a reference text for undergraduate courses in global health. Alternatively, some chapters (or parts of chapters) can be used in graduate or undergraduate courses dedicated to more specific subjects and topics. Ideally, students who use the text in this way will be stimulated to explore other chapters once they have read the assigned material. Moreover, the text can serve as a useful reference for those already working in the field of global health in government agencies, health and development agencies, NGOs, or the private sector.

Because of the many dynamic areas and subjects we wanted to cover, we chose to prepare an edited text. We selected content experts for each chapter rather than presuming to have the expertise to write the entire text ourselves. We recognize that an edited text has its shortcomings, such as some inconsistencies in style and presentation and occasional overlap in chapter content. We have done our best to limit these disadvantages, and hope the reader will agree that those that remain are a small price to pay for fulfilling our goal of providing the reader with the highest-quality content.

Another consequence of the dynamic nature of global health is the occasional difficulty in providing the most up-to-date epidemiologic information on all causes of mortality and morbidity. To assist the reader in obtaining this information, we have provided salient references in various chapters, including internet resources.

This is the fourth edition of the text. In planning its preparation, we sought advice on how to improve it from those who prepared chapters in the first three editions, as well as from faculty in various countries who were using the text in their courses; we also examined important current and emerging trends in global health. The text has 21 chapters, including four new chapters that have been added in response to feedback from these reviewers.

The first four chapters set the background. The *Measures of Health and Disease in Populations* chapter reviews the importance of using quantitative indicators for decision making in health. It presents the latest developments in the measurement of health status and the global burden of disease, including the increasing use of composite measures of health that combine the effects of disease-specific morbidity and mortality on populations. It then reviews current estimates and future trends in selected countries and regions, as well as the global burden of disease.

The *Culture, Behavior, and Health* chapter examines the social, cultural, and behavioral parameters that are essential to understanding public health efforts. This chapter describes key concepts in the field of anthropology, particularly as they relate to health belief systems. It presents theories of health behavior that are relevant to behavior change with examples of specific national and community programs in various areas of health. The importance of combining qualitative and quantitative methodologies in measuring and assessing health status and programs is emphasized.

The 1946 WHO constitution established that "the highest attainable standard for health as a fundamental right of every human being" (WHO, 2014, p. 1). The *Global Health, Human Rights and Ethics* chapter is new to this edition of the text in recognition of the inextricable link between human rights and health.

[1]A classification of countries can be found on the World Bank's website: https://datahelpdesk.worldbank.org/knowledgebase/topics/19280-country-classification.

It begins by defining global health in the context of human rights and ethics and by describing the importance of these concepts in global health approaches. Ebola, Zika, HIV/AIDS, and sugar-sweetened beverages are used as case studies to explore the role of human rights and ethics in a global health response. Finally, the chapter discusses the many challenges faced in implementing these approaches.

Addressing the underlying causes of ill health that stem from complex social, political, historical, economic, and environmental factors is crucial to improving health and addressing health inequity.

The *Understanding and Acting on Social Determinants of Health and Health Equity* chapter presents and explains these determinants. It focuses on the development and application of conceptual frameworks for understanding the social determinants of health, the implications of social determinants on health systems and services, the role of social determinants of health in an increasingly globalized world, and the importance of evaluating actions that address social determinants and health equity.

The next three chapters are devoted to the three greatest public health challenges traditionally faced by LMICs: reproductive health, infectious diseases, and nutrition. Reproductive health has long been addressed primarily through family planning programs directly intended to reduce fertility, and through programs that prevent sexually transmitted infections. The *Reproductive Health* chapter presents more current views of this area that broaden the concept to include empowerment of women in making decisions about their health and fertility. It provides information on population growth and demographic changes around the world, reviews how women control their fertility, indexes the effects of various social and biological determinants of fertility, and examines the impact of family-planning services and programs.

Collectively, infectious diseases have historically been the most important causes of premature mortality and morbidity in LMICs. The *Infectious Diseases* chapter presents the descriptive epidemiologic features and available prevention and control strategies for the infectious diseases that are of greatest public health significance in these countries today—namely, vaccine-preventable diseases; diarrhea and acute respiratory infections in children; tuberculosis, influenza, malaria, and other parasitic diseases; HIV/AIDS and other sexually transmitted diseases; and emerging infectious diseases and new disease threats, including Zika and Ebola. The chapter also addresses concerns about antimicrobial resistance. Examples of successful control programs are described, as are the challenges and obstacles that confront LMICs in implementing them.

Nutritional concerns in LMICs are diverse, ranging from deprivation and hunger to consequent deficiencies in health, survival, and quality of life in some regions. The *Nutrition* chapter focuses on several spheres of nutrition that are of utmost concern in these countries, including undernutrition and its components of protein- and energy-related malnutrition and micronutrient deficiencies at various stages of life; food insecurity; the interaction of nutrition and infections; the role of breastfeeding and complementary feeding in ensuring healthy children; and currently implemented nutrition-specific interventions and programs. It also considers the nutrition transition and the growing problem of overweight and related metabolic diseases in adults observed in more affluent segments of populations in rapidly developing countries. Finally, the chapter discusses the cost-effectiveness of nutritional interventions.

The next four chapters address the shifting and emerging global health priorities in NCDs, injury, mental health, environmental and occupational health, and complex emergencies.

NCDs are generally characterized by a long latency period, prolonged course of illness, noncontagious origin, functional impairment or disability, and incurability. The *Chronic Diseases and Risks* chapter covers the global epidemiology of these diseases, particularly their prevalence in LMICs, as well as the impact they have on health and economies. The chapter also discusses the four key risk factors—poor diet, tobacco, alcohol and substance abuse, and low physical activity—that serve as determinants and proposes approaches, programs, and policies required to adequately prevent and manage these diseases at national and global levels.

The *Unintentional Injuries and Violence* chapter discusses both unintentional injuries (ones for which there is no evidence of predetermined intent, such as road accidents and occupational injuries) and intentional injuries or violence that is planned or intended (including injuries related to self-directed, interpersonal, and collective violence). The chapter provides an overview of the global burden of injuries, outlines the causes of and risk factors for them, describes evidence-based interventions that can successfully reduce their impact, and considers the opportunities and challenges that can move forward an injury prevention agenda at the global level.

Mental health is finally starting to receive the attention commensurate with its great importance to the disability and disease burden in LMICs. The *Global Mental Health* chapter charts the historical development of public mental health; considers various concepts and classifications of mental disorders, taking into account the influence of cultural factors in the development of psychiatric classifications; and reviews what is known about the epidemiology and etiology of the more common disorders, including anxiety and mood disorders, psychotic disorders, substance abuse disorders, epilepsy, developmental disabilities, and dementia. The chapter reviews mental health policies, human resources for mental health care, and the evidence for the prevention and treatment of major mental disorders, and provides guidance on implementing mental health services into health systems. Lastly, the chapter addresses the importance of research and emerging research priorities.

The *Environmental and Occupational Health* chapter provides a comprehensive review of environmental and occupational health issues in LMICs, and its new title reflects an increased focus on occupational health and its link to environmental health. The chapter examines the global burden of environmental health hazards in five categories: household (e.g., water and sanitation), workplace (e.g., on farms, in mines, and in factories), community (e.g., outdoor air pollution), and regional and global levels (e.g., climate change, ozone depletion, and biodiversity). It covers important methodology issues in the field and provides several case studies on the most pressing issues. The chapter concludes with a discussion on future priorities for environmental health research and policy.

The *Complex Emergencies* chapter focuses on the global health challenges that characterize conflicts that occur within and across state boundaries, have political antecedents, and are protracted in duration. It considers the causes of complex emergencies (particularly the political causes) and their impact on populations and health systems, and reviews the technical interventions that can limit their adverse effects on the health of populations. Attention is drawn to the importance of an effective and efficient early response in influencing the long-term survival of populations and health systems and the nature of any post-conflict society that is established. The chapter also reviews the impact of natural disasters.

The next three chapters are concerned with the development and implementation of effective health systems, which have a crucial influence on the ability of countries to address their disease burden and improve the health of their populations. The *Design of Health Systems* chapter focuses mainly on the healthcare system. It provides a conceptual map of the health system along with its key elements; addresses the fundamental and often controversial question as to the role of the state; and considers the key functions of any health system—namely, regulation, financing, resource allocation, and services provision. It includes extensive references to country experiences in strengthening the various dimensions of health systems, and reviews current issues. Five country examples are used throughout the chapter to illustrate key differences in health systems across the world.

Public Health Infrastructure is a new chapter that defines public health functions and the tangible and intangible aspects that make up the public health infrastructure. These include the institutions, knowledge, and commodities involved in supporting public health at local, national, and global levels. The chapter discusses ways to strengthen the public health system, including universal health coverage, human and material resources, and emergency response systems. It also examines issues surrounding global health security, the International Health Regulations, and the need for research and development to improve public health practice.

As multipurpose and multidisciplinary endeavors, health systems require coordination among numerous individuals and units. Thus they require effective and efficient management and planning. The *Management and Planning for Global Health* chapter is dedicated to this topic. It details the important aspects of the political, social, and economic context in which a management process must operate; discusses the organizational structures under which healthcare systems may be organized, including the role of the private sector; examines the critical processes of planning and priority setting; looks at issues in the management of resources, focusing on finance, resource allocation, staffing, transport, and information; and concludes by discussing some cross-cutting themes, such as management style, accountability, and sustainability.

The *Pharmaceuticals* chapter is dedicated to this key element of a health system. It focuses on access, availability (both upstream issues and country-level distribution and management systems), and affordability of pharmaceuticals and their safe and effective use. It also discusses the pharmaceutical system architecture and reflects on coordination and priority setting in a complex global environment.

Health technology and innovation resources have grown exponentially in recent years and have the ability

to transform healthcare delivery. The newly added *Innovation, Technology, and Design* chapter examines the process of technical innovation and design, and explores the impact it can have on health. The chapter focuses on three critical areas of innovation in global health: mobile health technologies; point-of-care diagnostics; and improved access technologies. As innovative technologies achieve their full potential only when they are successfully scaled up, the chapter outlines methods on how to ensure technologies can be translated into the field. It concludes by discussing some of the gaps in health care that might potentially be addressed by new innovations.

The *Evaluations of Large-Scale Health Programs* chapter covers the important area of evaluation science and addresses the rationale for and the design of summative impact evaluation of programs being scaled up and delivered to large populations and aimed at delivering several biological and behavioral interventions together. To describe the planning, design, and execution of program evaluations and data analyses, the authors use three evaluations as examples: An Integrated Community Case Management of Childhood Illness program, an accelerated child survival development initiative, and a voucher scheme for insecticide-treated bed nets.

Health and health systems interrelate with a nation's economy in two main ways. The first, as noted earlier, comprises the bidirectional relationships between health status and national income and development. For example, health affects income through its impact on labor productivity, saving rates, and age structure; in turn, a higher income improves health by increasing the capacity to produce food and have adequate housing and education, and through incentives for fertility limitation. The second type of interaction concerns linkages among healthcare delivery institutions, health financing (including insurance) policies, and economic outcomes. The *Health and "The Economy"* chapter reviews both of these challenging and closely related topics, which are critical to government policy makers seeking the best ways to improve the quality of life of their populations, and are particularly important for those countries transitioning between low-, middle-, and high-income status.

The global trading system has both direct and indirect impacts on the health of populations, but is a factor that is often overlooked by health professionals and was addressed only tangentially in the previous edition of this text. Given its importance, we have added a new *International Trade and Health* chapter that examines the relationships among international trade, determinants of health, and health care. It outlines international trade agreements, describes the surrounding political context, and reviews the global health community's involvement in this field.

Finally, the *Global Health Governance and Diplomacy* chapter addresses the role of governance and globalization in global health, bringing together material included in the previous edition's two chapters on cooperation in global health and globalization and health. It covers the various institutional arrangements that respond to global health needs through collective action, considers how these have shifted over time, and assesses the governance challenges faced by the global health community. It concludes with thoughts on how to strengthen global health governance and its institutional architecture in the future.

Many case studies can be found in exhibits scattered throughout the text. They provide concrete examples and illustrations of key points and concepts covered in each chapter. At the conclusion of each chapter is a list of questions that can help course instructors stimulate classroom discussions about important issues covered in the chapter.

The editors recognize that this text could include separate chapters on many other topics—maternal and child health and implementation or delivery science, for example. We have opted instead to provide in-depth information on the core subjects that were selected, although we did our best to cover some aspects of these and other topics in one or more chapters.

In many ways, global health stands today at an important crossroads. Its greatest challenge is to confront global forces, while at the same time promoting local, evidence-based, cost-effective, and responsive programs that deal with both disease-specific problems and more general issues, such as poverty and gender inequality. Global health–related research is essential to gain a better understanding of the determinants of illness and of innovative approaches to prevention and care, as well as to find means of improving the efficiency and coverage of health systems. Whether as practitioners, policy makers, or researchers, global health professionals can make an enormous difference by being well trained and sensitive to the beliefs, culture, and value systems of the populations with whom they collaborate or serve. We hope this text will aid in this process.

References

Arnold, D. (1988). *Introduction: Disease, medicine, and empire. Imperial medicine and indigenous societies* (pp. 1–26). Manchester, UK/New York, NY: Manchester University Press/ St. Martin's Press.

Basch, P. F. (1999). *Textbook of international health*. New York, NY: Oxford University Press.

Board on International Health, Institute of Medicine. (1997). *America's vital interest in global health: Protecting our people, enhancing our economy, and advancing our international interests.* Washington, DC: Institute of Medicine.

Committee on the U.S. Commitment to Global Health. (2009). *The U.S. commitment to global health: Recommendations for the new administration*. Washington, DC: Institute of Medicine.

De Cock, K. M., Lucas, S. B., Mabey, D., & Parry, E. (1995). Tropical medicine for the 21st century. *British Medical Journal*, *311*(7009), 860–862.

Gensini, G. F., Yacoub, M. H. & Conti, A. A. (2004). The concept of quarantine in history: From plague to SARS. *Journal of Infection*, *49*(4), 257–261.

Gibbs, E. P. J. (2014). The evolution of One Health: A decade of progress and challenges for the future. *Veterinary Record*, *174*(4), 85–91.

Global Burden of Disease (GBD) 2016 Causes of Death Collaborators. (2017). Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet, 390*(10100), 1151–1210.

Hollenhorst, M. (2016). Global poverty declines even amid economic slowdown, World Bank says. North Carolina Public Radio. Retrieved from https://www.npr.org/sections/thetwo-way/2016/10/02/496099777/global-poverty-declines-even-amid-economic-slowdown-world-bank-says

Horton, R. (2018). Offline: From 1918 to 2018—the lessons of influenza. *The Lancet*, *391*(10115), 14.

Jamison, D. T., Summers, L. H., Alleyne, G., Arrow, K. J., Berkley, S., Binagwaho, A., & Ghosh, G. (2013). Global health 2035: A world converging within a generation. *The Lancet*, *382*(9908), 1898–1955.

Koplan, J. P., Bond, T. C., Merson, M. H., Reddy, K. S., Rodriguez, M. H., Sewankambo, N. K., . . . Wasserheit, J. N. (2009). Towards a common definition of global health. *The Lancet*, *373*(9679), 1993–1995.

Kuiken, T., Fouchier, R., Rimmelzwaan, G., & Osterhaus, A. (2003). Emerging viral infections in a rapidly changing world. *Current Opinion in Biotechnology*, *14*(6)4, 641–646.

Leff, S., & Leff, V. (1958). *From witchcraft to world health*. New York, NY: Macmillan.

Mohammadi, D. (2016). 2016: Year in review. *The Lancet*, *388*(10063), 2971–2972.

National Academies of Sciences, Engineering, and Medicine. (2017). *Promoting cardiovascular health and preventing cancer: National Academies of Sciences, Engineering, and Medicine report on global health and the future role of the United States.* Washington, DC: National Academies Press.

Polio Global Eradication Initiative. (2017). Working toward a polio-free future: 2017 in review. Retrieved from http://polioeradication.org/news-post/working-toward-a-polio-free-future-2017-in-review/

Pruss-Ustun, A., Wolf, J., Corvalan, C., Bos, R., & Neira, M. (2016). *Preventing disease through healthy environments: A global assessment of the burden of disease from environmental risks.* Geneva, Switzerland: World Health Organization.

Rosen, G., & Morman, E. T. (1993). *A history of public health*. Baltimore, MD: Johns Hopkins University Press.

Ruger, J. P. (2005). The changing role of the World Bank in global health. *American Journal of Public Health*, *95*(1), 60–70.

Stuard, S. M., & NetLibrary. (1992). *A state of deference: Ragusa/ Dubrovnik in the medieval centuries*. Philadelphia, PA: University of Pennsylvania Press.

Summergrad, P. (2016). Investing in global mental health: The time for action is now. *The Lancet Psychiatry*, *3*(5), 390–391.

The Economist. (2017). Fewer but still with us: The world has made great progress in eradicating extreme poverty. Retrieved from https://www.economist.com/news/international/21719790-going-will-be-much-harder-now-world-has-made-great-progress

The Lancet Planetary Health. (2017). Welcome to the Lancet Planetary Health. *The Lancet Planetary Health, 1*(1), e1.

Thornton, J. (2017). Drones and phones: How mobile tech is fighting global diseases. London School of Hygiene & Tropical Medicine. Retrieved from https://www.lshtm.ac.uk/research/research-action/features/drones-and-phones-how-mobile-tech-fighting-global-diseases

UNAIDS. (2017a). Fact sheet: Latest statistics on the status of the AIDS epidemic. Retrieved from http://www.unaids.org/en/resources/fact-sheet

UNAIDS. (2017b). Resources and financing. Retrieved from http://www.unaids.org/en/topic/resources

UNICEF. (2016). The state of the world's children 2016: A fair chance for every child. Retrieved from https://www.unicef.org/publications/files/UNICEF_SOWC_2016.pdf

United Nations High Commissioner for Refugees. (2017). Figures at a glance. Retrieved from http://www.unhcr.org/en-us/figures-at-a-glance.html

USAID. (2017). Dreams: Partnership to reduce HIV/AIDS in adolescent girls and young women. Retrieved from https://www.usaid.gov/what-we-do/global-health/hiv-and-aids/technical-areas/dreams

Winslow, C.-E. A., & Hallock, G. T. (1933). *Health through the ages*. New York, NY: Metropolitan Life Insurance Company.

World Health Organization (WHO). (2006). *The world health report 2006: Working together for health*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2014). *Basic Documents: Forty-eighth Edition*. Geneva, Switzerland: Author. Retrieved from http://apps.who.int/gb/bd/PDF/bd48/basic-documents-48th-edition-en.pdf

World Health Organization (WHO). (2016a). Global strategy on human resources for health: Workforce 2030. Draft 1.0. Retrieved from http://who.int/hrh/resources/global_strategyHRH.pdf?ua=1

World Health Organization (WHO). (2016b). Maternal mortality. Retrieved from http://www.who.int/mediacentre/factsheets/fs348/en/

World Health Organization (WHO). (2017a). Children: Reducing mortality. Retrieved from http://www.who.int/mediacentre/factsheets/fs178/en/

World Health Organization (WHO). (2017b). The top 10 causes of death. Retrieved from http://www.who.int/mediacentre/factsheets/fs310/en/

CHAPTER 1

Measures of Health and Disease in Populations

Abdulgafoor M. Bachani and Adnan A. Hyder

In its 1948 charter, the World Health Organization (WHO) defined health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity." Although this is an important ideological conceptualization, for most practical purposes, objectives of health programs are more readily defined in terms of prevention or treatment of disease. Disease has been defined in many ways and for a variety of reasons; distinctions may be made between disease, sickness, and illness. For purposes of defining and measuring disease burden, a general definition will be used in this text: *Disease* is anything that a person experiences that causes, literally, "dis-ease"—that is, anything that leads to discomfort, pain, distress, disability of any kind, or death constitutes disease. It may be due to any cause, including injuries or psychiatric conditions.

It is also important to be able to diagnose and classify specific diseases to the extent that such classification aids in determining which health intervention programs would be most useful. Thus, defining disease, understanding the pathogenesis of the disease process, and knowing which underlying risk factors lead to this process are critical for understanding and classifying causes so as to determine the most effective prevention and treatment strategies for reducing the effects of a disease or risk factor. Just as the purpose of diagnosis of a disease in an individual patient is to provide the right treatment, so the major purpose of working through a burden of disease analysis in a population is to provide the basis for the most effective mix of health and social program interventions.

Developments in the measurement of population health status and disease burden over the past two decades include the increasing use of summary, composite measures of health that combine the mortality and morbidity effects of diseases into a single indicator; the availability of results of Global Burden of Disease (GBD) studies, which make use of such summary indicators; and developments in the measurement of disability and risk factors. The more traditional approaches to measuring health are widely available in other public health textbooks and will be used for illustrative and comparative purposes here.

This chapter is divided into five sections. The first section explains the reasons for and approaches to measuring disease burden in populations, describes the need for using quantitative indicators, highlights the importance of using data for decision making in health, and lists a variety of major health indicators currently in widespread use. The second section critically reviews methods for developing and using composite measures that combine the mortality and morbidity from diseases in populations at national and regional levels. It explores the potential utility of these measures and discusses their limitations

and implications. The third section demonstrates the application of these methods for measurement of health status and assessment of global health trends. It reviews current estimates and forecasts trends in selected countries and regions, as well as examines the global burden of disease. The fourth section reviews important underlying risk factors of disease and discusses recent efforts to measure the prevalence of major risk factors and to determine their contributions to regional and global disease burdens. The final section provides conclusions for the chapter.

Reasons for and Approaches to Measuring Health and Disease

Rationale

The many reasons for obtaining health-related information all hinge on the need for data to guide efforts toward reducing the consequences of disease and enhancing the benefits of good health. These include the need to identify which interventions will have the greatest beneficial effect, to identify emerging trends and anticipate future needs, to assist in determining priorities for expenditures, to provide information for education to the public, and to help in setting health research agendas. The primary information requirement is for understanding and assessing the health status of a population and its changes over time. In recent years, practitioners have emphasized the importance of making evidence-based decisions in health care. There is little reason to doubt that evidence is better than intuition, but realizing its full benefits depends upon recognizing and acting upon the evidence. This chapter examines evidence—the facts of health and disease—and demonstrates how to assemble this evidence so that it can assist in better decision making concerning health and welfare.

A well-documented example of the relationship between decision making and data can be seen in a classic health systems project in Tanzania (**EXHIBIT 1-1**). This

EXHIBIT 1-1 Using Evidence to Improve a Health System: An Example from Africa

The **Tanzania Essential Health Interventions Project** (TEHIP), a joint venture of the Tanzanian Health Ministry, the International Development Research Centre (IDRC), and the Canadian International Development Agency, starting in 1996 was conducted in two rural districts—Morogoro and Rufiji—with a combined population of approximately 700,000. The annual health spending in Tanzania was about $8 per capita. In Morogoro and Rufiji, TEHIP added resources on the condition that they be spent rationally; in other words, the amount of money spent on interventions should reflect the burden of disease. TEHIP conducted burden of disease analysis for the two districts and established a demographic surveillance system (DSS). The routine data from the DSS provided information for the district teams to support resource allocation based on disease burden. The organization found that the amount that the local health authorities spent on addressing each disease bore little relation to the actual burden of disease. Although childhood problems (e.g., pneumonia, diarrhea, malnutrition, measles) constituted 28% of the disease burden, only 13% of the budget was devoted to addressing them. Other conditions, meanwhile, attracted more than their fair share of resources. For example, 22% of the budget was targeted to tuberculosis, even though it accounted for less than 4% of years of life lost.

TEHIP promoted the use of burden of disease analysis, district accounts, and other mapping tools for more rational decision making in the districts. It also brought management tools and community voice techniques to the district teams. The result was better ability of district-level healthcare workers and managers to control and allocate resources and processes related to healthcare provision. The district teams decided to spend more on neglected diseases for which cost-effective treatments or preventive measures were available. The extra $1 per capita was enough to allow the district health authorities to align their spending to reflect the real disease burden. For example, sexually transmitted diseases received 3% of the budget prior to TEHIP's intervention; that percentage changed to 9.5% after the realignment. Malaria accounted for 30% of the years of life lost because of death and debilitating illness; the budget for malaria prevention and treatment programs increased from 5% of total spending in 1996 to 25% in 1998.

The results of TEHIP were documented as changes in health outcomes. In Rufiji, for example, infant mortality fell by 40% in 5 years. In fact, just between 1999 and 2000, infant mortality fell from 100 deaths per 1,000 live births to 72 deaths per 1,000 live births, while the proportion of children dying before their fifth birthdays dropped by 14%, from 140 per 1,000 to 120 per 1,000. The success of TEHIP and its approach led to replication and further innovation in not only Tanzania but also many other low- and middle-income countries (LMICs), including Burkina Faso, Ghana, and Nigeria.

For additional information on the TEHIP success story, visit the websites identified in the following sources.

Data from https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/7814.pdf; https://www.idrc.ca/en/article/tanzanias-healthcare-breakthrough; https://www.researchgate.net/publication/268341357_Tanzania_Essential_Health_Interventions_Project_An_Overview_of_TEHIP_Interventions_Tanzania_Ministry_Of_Health_TANZANIA_ESSENTIAL_HEALTH_INTERVENTIONS_PROJECT_TEHIP_Interventions_-An_Overview

case illustrates how able people with good intentions had been making decisions routinely, only to find that using established methods to collect evidence on the burden of disease changed the nature and effectiveness of their own decisions. A major reason for the effective use of the evidence was that it was collected locally and put forward in a form helpful to decision makers.

Measuring Health and Disease

The relative importance (burden) of different diseases in a population depends on their frequency (incidence or prevalence), severity (the mortality and extent of serious morbidity), consequences (health, social, economic), and the specific people affected (gender, age, social and economic position).

Counting Disease

The first task in measuring disease in a population is to count its occurrence. Counting disease frequency can be done in several ways, and it is important to understand what these different methods of counting actually mean. The most useful way depends on the nature of the disease and the purpose for which it is being counted. There are three commonly used measures of disease occurrence: cumulative incidence, incidence density, and prevalence.

Cumulative incidence, or *incidence proportion*, is the number or proportion of new cases of disease that occur in a population at risk for developing the disease during a specified period of time. For this measure to have meaning, three components are necessary: a definition of the onset of the event, a defined population, and a particular period of time. The critical point is *new cases* of disease—the disease must develop in a person who did not have the disease previously. The numerator is the number of new cases of disease (the event), and the denominator is the number of people at risk for developing the disease. Everyone included in the denominator must have the potential to become part of the group that is counted in the numerator. For example, to calculate the incidence of prostate cancer, the denominator must include only men, because women are not at risk for prostate cancer. The third component is the period of time. Any time unit can be used as long as all those counted in the denominator are followed for a period comparable with those who are counted as new cases in the numerator. The most commonly used time denominator is one year.

Incidence density, which is often simply called *incidence rate*, is the occurrence of new cases of disease per unit of person-time. This metric directly incorporates time into the denominator and is generally the most useful measure of disease frequency; it is often expressed as new events per person-year or per 1,000 person-years. Incidence is a measure of events (in this case, the transition from a nondiseased state to a diseased state) and can be considered a measure of risk. This risk can be looked at in any population group, defined by age, sex, place, time, sociodemographic characteristics, occupation, or exposure to a toxin or any other suspected causal factor.

Prevalence is a measure of present status rather than of newly occurring disease. It measures the proportion of people who have the defined disease at a specific point of time. Thus, it is a composite measure made up of two factors—the incidence of the disease that has occurred in the past and its continuation to the present or to some specified point in time. That is, prevalence equals the incidence rate of the disease multiplied by the average duration of the disease. For most chronic diseases, prevalence rates are more commonly available than are incidence rates.

Severity of Disease

To understand the burden of disease in a population, it is important to consider not only the frequency of the disease but also its severity, as indicated by the morbidity and premature mortality that it causes. *Premature mortality* is defined as death before the expected age of death had the disease not occurred. *Morbidity* is a statement of the extent of disability that a person suffers as a consequence of the disease over time and can be measured by a number of indicators, as discussed later in this chapter.

Mortality

Traditionally, mortality has been the most important indicator of the health status of a population. John Grant developed the first known systematic collection of data on mortality with the *Bills of Mortality* in the early 1600s in London. He described the age pattern of deaths, categorized them by cause as understood at the time, and demonstrated variations from place to place and from year to year. Mortality rates according to age, sex, place, and cause continue to be central information about a population's health status and a crucial input for understanding and measuring the burden of disease. Considerable literature exists on the use of mortality to indicate health status and its application to national and subnational levels (Murray & Chen, 1992).

The fact of death by age, sex, and place is required by law in most countries through death registration, and in many countries the cause of death through death

certification is required as well. Both provide essential information about the health status of a population. Nevertheless, in many low-income countries, the fact of death, let alone its cause, is still not reliably available.

In high-income countries, vital statistics (i.e., the registration of births and deaths by age, sex, and place) are routinely collected and highly reliable. In most middle-income countries, the reliability and completeness of these data have been steadily improving and often are fairly satisfactory. In contrast, the collection of vital statistics remains grossly incomplete in many low-income countries. An analysis of death registration in the course of the Global Burden of Disease study showed that vital registration data together with sample registration systems still do not cover 100% of global mortality. Survey data and indirect demographic techniques are needed to provide information on levels of child and adult mortality to paint a complete picture of global mortality (GBD 2015 Mortality and Causes of Death Collaborators, 2016). Nevertheless, even in low-income countries, increasing use of survey methods is delivering useful estimates of the mortality rates for the population younger than age 5 years and other populations.

Obtaining information about cause of death remains difficult even in many middle-income countries; a lot of information depends on special surveys or studies of select populations. Verbal autopsies (VAs) have been used increasingly for judging the likely cause of death, especially for children younger than age 5. This method comprises structured questions administered by trained interviewers with family members after a death; the information is then reviewed by physicians (or computers) to assign a cause of death using algorithms. VAs are useful for assessing some causes of death such as neonatal tetanus and severe diarrhea, but their sensitivity and specificity may be limited for diseases whose symptoms are variable and nonspecific, such as malaria (Anker et al., 1999; Thatte, Kalter, Baqui, Williams, & Darmstadt, 2009). Recently, automated systems for analyzing VA data have been developed and are being tested.

Age-specific mortality profiles are a prerequisite for a burden of disease analysis. Although extensive work has been done to document and analyze child mortality in low- and middle-income countries (LMICs), less has been done for adult mortality (Hill, 2003). LMICs have higher rates of age-specific adult mortality than do high-income nations (GBD 2015 Mortality and Causes of Death Collaborators, 2016; Lopez et al., 2002; Murray & Chen, 1992). Indeed, mortality rates are higher for both women and men in LMICs at every age when compared with the high-income world. In Africa, the enormous increase in deaths of young and middle-aged women and men from acquired immunoeficiency syndrome (AIDS) has had a profound impact on mortality and survival (**EXHIBIT 1-2**).

Traditional indicators of mortality have been the standard for assessing population health status. Neonatal mortality rates (NMR; deaths of live-born infants before 28 days of age per 1,000 live births), infant mortality rates (IMR; deaths of live-born infants before 12 months of age per 1,000 live births), and child mortality (deaths of children younger than 5 years of age) are considered sensitive indicators of the overall health of nations. The United Nations Children's Fund (UNICEF) publishes an annual global report that includes a ranking of nations based on these indicators (UNICEF, 2015). These indicators have the added advantage of having been studied for their relationships with other indicators of the social and economic development of nations. For example, a clear relation exists between the gross national product (GNP) per capita, which is an indicator of national wealth, and child mortality. In general, the higher the level of economic development, the lower the rate of child mortality. However, there are exceptions, and they need to be examined carefully. For example, Sri Lanka and the Indian state of Kerala are both low-income regions that have traditionally had low child mortality rates. These examples demonstrate that the relationship between mortality and poverty is complex and needs in-depth investigation.

There are major deficiencies in cause-specific mortality data in low- and most middle- income countries. In keeping with demographic and epidemiologic transitions, the pattern of cause-specific mortality changes at different levels of total mortality, with a general trend of decreasing infectious and parasitic disease cause-specific mortality with declining total mortality. Indeed, mortality from these communicable causes has been a major reason for the historical difference between high- and low-mortality populations (Murray & Chen, 1992).

The cause of death certification system based on WHO's *International Classification of Diseases* (ICD) has been used widely in many countries for many years (WHO, 2016). Despite the existence of this standardized process for categorizing deaths, variations in the reliability of these data occur because of variations in the training and expertise of the people who are coding causes of death, as well as the supervision and feedback provided. Nevertheless, there have been steady improvements in many countries, including automation of data collection and analysis, and these kinds of data provide some of the best information available on major causes of mortality.

EXHIBIT 1-2 Trends of the HIV/AIDS Epidemic

Globally, AIDS is the leading infectious cause of death among 15- to 49-year-olds (Global Burden of Disease Risk Factors Collaborators et al., 2015). Untreated disease caused by the human immunodeficiency virus (HIV) has a case fatality rate that approaches 100% (WHO, 2003). Unknown more than 30 years ago, the HIV/AIDS epidemic has killed more than 35 million people.

At the end of 2016, an estimated 36.7 million people were living with HIV/AIDS, with 69.5% of those individuals living in sub-Saharan Africa (WHO, 2017) . The prevalence of new HIV infections has increased by 60% in eastern Europe and Central Asia (**TABLE 1-1**) (UNAIDS, 2017; WHO, 2017).

HIV/AIDS is the tenth leading cause of disability-adjusted life years, accounting for 2.7% of this global burden. In terms of mortality, it is the eleventh leading cause of death among people of all ages, accounting for 2.1% of all deaths (Global Burden of Disease Risk Factors Collaborators et al., 2015). Nearly 42% of the 1 million global deaths from HIV/AIDS have occurred in East and Southern Africa (**FIGURE 1-1**) (UNAIDS, 2017).

TABLE 1-1 Global Summary of HIV and AIDS Epidemic

Number of people living with HIV	**Total**	**36.7 million**	**(30.8–42.9 million)**
	Adults	34.5 million	(28.8.4–40.2 million)
	Women	17.8 million	(15.4–20.3 million)
	Children	2.1 million	(1.7–2.6 million)
Number newly infected with HIV	**Total**	**1.8 million**	**(1.6–2.1 million)**
	Adults	1.7 million	(1.4–1.9 million)
	Children	160,000	(100,000–220,000)
AIDS deaths	**Total**	**1.0 million**	**(830,000–1.2 million)**
	Adults	890,000	(740,000– 1.1 million)
	Children	120,000	(79,000–160,000)

Reproduced from Joint United Nations programme on HIV/AIDS (UNAIDS). (2017). *UNAIDS Data 2017*. Geneva, Switzerland: Author. Retrieved from http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf

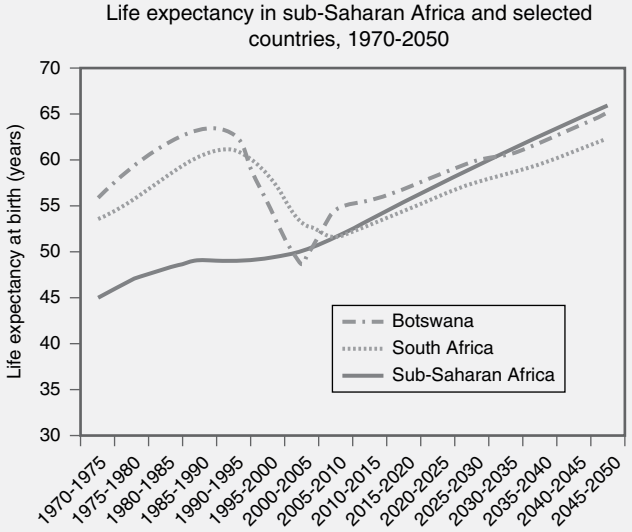


FIGURE 1-1 Trends in life expectancies and the HIV/AIDS epidemic.

Courtesy of Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat. (2008). *World population prospects: The 2008 revision*. Retrieved from www.un.org/esa/population/publications/wpp2008/wpp2008_highlights.pdf

Reproduced from Joint United Nations programme on HIV/AIDS (UNAIDS). (2017). *UNAIDS Data 2017*. Geneva, Switzerland: Author. Retrieved from http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf; World Health Organization (WHO). (2017). *Global health observatory (GHO) data – HIV/AIDS*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/gho/hiv/en/

Mortality can be expressed in two important quantitative measures: (1) *mortality rate* (MR) and (2) *case fatality ratio* (CFR). The MR, a form of incidence rate, is expressed as the number of deaths in a defined population in a defined time period. The numerator can be total deaths, age- or sex-specific deaths, or cause-specific deaths; the denominator is the number of persons at risk of dying in the stated category as defined earlier for incidence. Demographers use the notation XqY for the probability of dying in the Y years following age X at the then prevailing age-specific mortality rates for the population. Thus, 5q0 is the probability of death of newborns by age 5 years (see Table 1-2 later in this chapter), and 30q15 is the probability of death in young adults from age 15 to 45. The CFR is the proportion of those persons with a given disease who die of that disease (at any time, unless specified). The MR is equal to the CFR multiplied by the incidence rate of the disease in the population.

The distinction between the proportion of deaths attributable to a cause (number of deaths due to the cause divided by total number of deaths in a given population in a given time period) as compared to the probability of death from the cause (disease-specific MR) is important to understand. For example, the probability of death (and disability) from noncommunicable causes (indeed, from virtually all causes) is higher in low- and middle-income regions than in the high-income world. However, the proportion of deaths and disability attributable to these chronic causes is smaller in LMICs than in wealthier countries because of the much larger toll taken by infectious and nutritional causes. With increasing economic development, the risk of death and disability from chronic disease does not increase; rather, the proportion of deaths attributable to chronic disease increases as the proportion of deaths attributable to communicable and nutritional disease declines.

Demographic and Epidemiologic Transitions

The demographic transition describes the changes in birth and death rates that historically have accompanied the shift from a traditional society to a modern society; it is detailed in other chapters. With modernization, sharp declines in mortality have been followed by a reduction in fertility, albeit one that commonly lags behind the change in the death rate by years or decades. The term *transition* refers to the shift away from a stable population in which very high birth rates are balanced by very high death rates to a stable population in which low birth rates are balanced with low death rates. In between these extremes, as a society undergoes modernization, there is a lag between falling mortality, especially in the under-5 age group, and the drop in birth rates that leads to explosive population growth. Thereafter birth rates fall and a new stage is reached in which birth and death rates are low and balance resumes. The result is a striking change in the age structure of the population, with a decreased proportion of children and an aging population. These changes in the population age distributions are reflected in the shift from a wide-based pyramid, reflecting larger numbers in the younger age groups, to a structure with a narrow base, nearly rectangular configuration, and nearly equal percentages in each age group.

In 1971, Omran described the underlying reasons for the demographic transition and used the term *epidemiologic transition* to explain the changing causal factors of disease that accounted for the dramatic drop in under-5 mortality, which was largely due to reduction in malnutrition and communicable diseases. Although high rates of maternal mortality are characteristic of the low- and middle-income world, reductions in maternal mortality occur in a different time frame from those in under-5 mortality. Reductions in maternal mortality require a better-developed infrastructure, including ready availability of surgical and blood transfusion capacity plus improved communication and transportation systems. Thus, drops in maternal mortality occur much further along the road toward economic development, and changes occur only after shifts in the child mortality have been seen.

Major changes in the patterns and causes of injury are also likely to occur with modernization. For example, road traffic injuries tend to increase as countries go through the stage of development in which there is a great increase in vehicles and in the speeds at which they are operated before improved roads, appropriate laws and regulations, and law enforcement are in place (Crooper & Kopits, 2003; WHO, 2015). There may also be important shifts in the nature of violence and the people toward whom it is directed, related to crime patterns, civil unrest, ethnic conflicts, and intrafamily tensions (WHO, 2002b). The profound impact of the HIV/AIDS epidemic was discussed earlier in Exhibit 1-2.

Other Health-Related Metrics

In addition to basic measures of mortality, morbidity, and life expectation that are central for population health status assessment, a variety of important health-related indicators are useful for specific purposes. Many are discussed more fully in other chapters of this text; they are summarized in **TABLE 1-2**. Those related to the Sustainable Development Goals (SDGs) are discussed in **EXHIBIT 1-3**.

TABLE 1-2 Health-Related Metrics

Type	Indicator	Definition/Interpretation
Demographic indicators: reproductive health	Maternal death	Death of a woman while pregnant or up to 42 days post-delivery from any cause except accident
	Maternal mortality ratio	Maternal deaths per number of pregnancies (maternal deaths per 100,000 live births)
	Maternal mortality rate	Maternal deaths per number of women of reproductive age (maternal deaths per 100,000 women aged 15–49)
	Lifetime risk of maternal mortality	Cumulative loss of human life due to maternal death over the female life course
	Total fertility rate	Average number of children a woman would bear if she lived to the end of her reproductive period
	Life expectation at birth	Average number of years a newborn would live if his or her life were lived under the mortality conditions for the place and year in question
Anthropometric indicators: nutrition	Weight for age	Underweight
	Height for age	Stunting
	Weight for height	Wasting
	Mid-upper arm circumference	Wasting
Mortality (death) indicators	Mortality rate	Number of deaths in a specified time period/number of persons at risk of dying during that period
	Infant mortality rate	Number of deaths of live born infants before 12 months of age per 1,000 live births
	Under-5 mortality rate	Number of deaths of children younger than age 5 per 1,000 live births averaged over the last 5 years
	5q0	Probability of death of a newborn by age 5
	Neonatal mortality rate	Number of deaths of live-born infants before 28 days of age per 1,000 live births
	Stillbirth rate	Number of babies born with no signs of life at or after 28 weeks' gestation per 1,000 births
	Perinatal mortality rate	Number of fetal deaths (28 or more weeks of gestation) + postnatal deaths (first week) per 1,000 live births
Disease frequency	Endemic	Usual occurrence of a given disease in a defined population
	Epidemic	Occurrence of a given disease in a defined population clearly in excess relative to its usual occurrence
	Pandemic	A worldwide epidemic involving large numbers

EXHIBIT 1-3 Sustainable Development Goals

On September 25, 2015, UN member states adopted 17 *Sustainable Development Goals* (SDGs) as part of the post–Millennium Development Goals (MDGs) development agenda. The SDGs build on the goals identified under MDGs and broaden the scope to include new areas of focus—for example, climate change, economic inequality, innovation, sustainable consumption, and peace and justice—to improve overall well-being and life of current and future generations through sustainable means.

The 17 SDGs were divided into 169 quantifiable targets that are measured by 230 indicators. Of these, 21 targets and 39 indicators are directly related to health. The health-related indicators include a variety of indicator types: incidence rates, prevalence "rates," mortality rates, mortality ratios, birth rates, and proportion of target populations receiving an intervention.

Appendix 1 provides examples of some of the health-related SDG indicators.

Morbidity and Disability

Measures of mortality have been the principal indicators of population health status for generations. Their relative ease of observation, availability of data, and history of use make mortality information useful for assessing and monitoring the health status of populations. However, the key limitation with mortality-based indicators is that they "note the dead and ignore the living" (Kaplan & Anderson, 1996). Measurements of morbidity, by comparison, are more problematic because there is not a clearly defined endpoint such as death provides. In addition, several components of disability need to be assessed, and there may be a substantial subjective aspect to grading the extent or severity of a condition.

The *International Classification of Impairments, Disabilities, and Handicaps* (ICIDH) was developed in the 1970s to classify nonfatal health outcomes as an extension of WHO's ICD system (WHO, 1980). It was developed to more fully describe the impact of a given disease on an individual and on society, and to account for that disease's heterogeneity of clinical expression and evolution in different individuals and societies. ICIDH categories included *impairment* (loss or abnormality of psychological, physiological, or anatomic structure or function), *disability* (restriction or lack of ability to perform an activity considered normal), and *handicap* (disadvantage from a disability or impairment for a given individual based on the inability to fulfill a normal role as defined by age, sex, or sociocultural factors). These distinctions clarified more than just processes—they helped define the contribution of medical services, rehabilitation facilities, and social welfare to the reduction of disability.

In 2002, WHO built on the ICIDH to develop the *International Classification of Functioning, Disability, and Health*, commonly known as ICF (WHO, 2002c). In this system, health-related domains are classified from the perspectives of the body, of the individual, and of society by means of two lists: a list of body functions and structures, and a list of domains of activity and participation. Because an individual's functioning and disability occur within a context, the ICF also includes a list of environmental factors that provide a description of that context. The ICF has become WHO's framework for measuring health and disability at both individual and population levels. It was officially endorsed by all 191 WHO member states in the Fifty-Fourth World Health Assembly on May 22, 2001 (resolution WHA 54.21). Unlike its predecessor, which was endorsed for field trial purposes only, the ICF was endorsed for use in member states as the international standard to describe and measure health and disability.

Using such classifications, indicators of disability—such as *impairment-free*, *disability-free*, and *handicap-free* life expectancies—have been developed. These, in turn, have been used to estimate health-adjusted life expectancies using severity and preference weights for time spent in states of less than perfect health.

Hospital inpatient discharge records—when they are based on good clinical evidence and coded by staff well trained in coding procedures—can provide high-quality data on the major causes of morbidity serious enough to require hospitalization. They also can provide good cause-of-death data for hospitalized persons, and some sense of the outcome status of those with serious conditions. Hospital data are generally improving in quality, especially in middle-income countries and in selected sentinel (usually tertiary care) teaching hospitals in low-income countries. Such information is inevitably biased because of the highly skewed distribution of those using such hospitals, but in many situations it is possible to have a good understanding of those biases and make appropriate adjustments to draw useful conclusions.

Generally, outpatient records in most of the world are highly deficient in terms of diagnosis; indeed,

they often identify only the patient's chief complaint and the treatment dispensed. The main value of most such records is limited to establishing the fact of using a facility. There are usually strong biases in terms of those patients who use outpatient facilities because of access factors (distance and cost of use), nature and severity of the disease problem, and opportunity for using alternative services.

Visits to healthcare facilities, functional disability (a measure of activity that is less than the norm), and time spent away from work (absenteeism, work days lost) have all been used to assess the magnitude of morbidity from various conditions. A commonly used approach to evaluating morbidity in a population has been the assessment of the impact on social roles or functional performance, such as days missed from work or spent in bed (Kaplan & Anderson, 1996). A considerable body of literature focuses on the wide variety of instruments used to measure such functional capacity, especially in the clinical medical literature, that are not directly useful for population-based morbidity assessment.

Data about morbidity are often based on self-perceived assessments, and are frequently gleaned from survey-based interview information. The perception of morbidity and its reporting, the observation of morbidity and its impact, and other factors are responsible for the wide variations between reported and measured prevalence of conditions (Murray & Chen, 1992). This has resulted in an underestimation of the presence and impact of morbidity in both LMICs as compared with high-income nations. This situation also underscores the variations in morbidity data, which are often interpreted as indicating that wealthy individuals and low-mortality populations report higher rates of morbidity (*Global Burden of Disease and Risk Factors*, 2006; Woolf et al., 2015).

Measurement of health-related quality of life has also been discussed in the medical literature for decades. *Health-related quality of life* refers to how well an individual functions in daily life and his or her perception of well-being. Various domains of quality have been defined, such as health perception, functional status, and opportunity, and several instruments have been developed to evaluate them. Both disease-specific and general instruments exist, with such tools abounding in fields dealing with patients having chronic disabled states, such as psychiatry, neurology, and counseling. These scales are often dependent on self-reported information, although some incorporate observational data as well. However, concerns have been raised about their reliability and validity. These measures are not discussed further in this text, because they have been primarily used in clinical assessments of individuals, rather than larger populations.

Measuring Disability

If all the various forms of disability—physical, functional, mental, and social—are to be compared with mortality, they must be measured in an equivalent manner for use in health assessments. To do so, measurement of disability must quantify the duration and severity (extent) of this complex phenomenon. A defined process is needed that rates the severity of disability as compared with mortality, measures the duration of time spent in a disabled state, and converts various forms of disability into a common scale. General measures of disability without regard to cause (often carried out by special household surveys) are useful to determine the proportion of the population that is "disabled" and unable to carry out normal activities, but are not much help for quantifying the extent of disability.

In general, three components of disability need to be assessed. The first component is the *case disability ratio* (CDR)—the proportion of those diagnosed with the disease who have disability. For most diseases that are diagnosed clinically, the CDR will be 1.00 because, by the definition of disease given earlier, patients will have signs or symptoms. In contrast, when the diagnosis is based on, for example, infection rather than disease (such as tuberculosis) or on a genetic marker rather than the physical manifestation (such as sickle cell trait), the CDR is likely to be less than 1.00.

The second component of disability is its *extent or severity*—how incapacitated the person is as a result of the disease. The extent of disability is expressed on a scale, such as from 0 (indicating no disability) to 1.00 (equivalent to death). The assessment of severity can be quite subjective, particularly because so many different types and dimensions of disability exist. A number of methods have been introduced in an effort to achieve comparability and obtain consistency (Murray, Salomon, Mathers, & Lopez, 2002).

Measurement of individual preferences for different health states to determine relative severity of disability has been done by a variety of methods (Kaplan & Anderson, 1996; Murray et al., 2002; Torrance, 1986). Factors that influence the assessment of such preferences include the type of respondent, the type of instrument used to measure the response, and the time from entry into the disabled state. Individuals who are in a particular state, healthy individuals, healthcare providers, caretakers, and family members have all

been interviewed in studies. Adaptation, conditioning, development of special skills, and vocational training can all change the response of individuals over time within a particular health state, thereby affecting the value of that state to the individual. As a consequence, the valuation is time dependent—for example, the value placed on a year of life by a paraplegic soon after entering that health state would be different from that obtained after several years of adjustment to that state (Murray & Lopez, 1994).

Instruments used to extract such preferences involve visual and interview techniques (*Global Burden of Disease and Risk Factors*, 2006; Torrance, 1986). Two alternative scenarios are often presented to the subject and the point of indifference sought (as in standard gamble techniques). Despite much work in this area, there is no consensus or accepted standard method for such elicitation. Severity of disability scales have been developed by group consensus using community surveys (Kaplan & Anderson, 1996), a mixture of community and expert groups (Ghana Health Assessment Team, 1981), experts only (World Bank, 1993), and population surveys (Global Burden of Disease Risk Factors Collaborators et al., 2015; Murray et al., 2002; Salomon et al., 2012). These scales usually compare perfect health states to death on a scale of 0 to 1 (**TABLE 1-3**).

In the first Global Burden of Disease 1990 study, the disability severity estimates were based on expert opinion. Twenty-two indicator conditions were selected and used to construct seven disability classes (Table 1-3). Outcomes from all other health conditions were categorized within these seven classes (with special categories

TABLE 1-3 Examples of Disability Classification Systems

Ghana Health Assessment Team, 1981		
Class	**Severity**	**Equivalent to (Maximum)**
1	0	Normal health
2	0.01–0.25	Loss of one limb's function
3	0.26–0.50	Loss of two limbs' function
4	0.51–0.75	Loss of three limbs' function
5	0.76–0.99	Loss of four limbs' function
6	1	Equivalent to death

Global Burden of Disease Study, 2013*	
Category	**Disability Weight for Unique Health States**
Infectious disease	Acute episode, mild: 0.006 (0.002–0.012) Acute episode, moderate: 0.051 (0.032–0.074) Diarrhea, mild: 0.074 (0.049–0.104) Ear pain: 0.013 (0.007–0.024)
Cancer	Diagnosis and primary treatment: 0.288 (0.193–0.399) Metastatic: 0.451 (0.307–0.600)
Cardiovascular and circulatory disease	Acute myocardial infarction (MI), days 1–2: 0.432 (0.288–0.579) Acute MI, days 3–28: 0.074 (0.049–0.105) Heart failure, mild: 0.041 (0.026–0.062)
Diabetes and digestive and genitourinary disease	Diabetic foot: 0.020 (0.010–0.034) Gastric bleeding: 0.325 (0.209–0.462) Infertility, primary: 0.008 (0.003–0.015)

Chronic respiratory disease	Asthma, controlled: 0.015 (0.007–0.026) Chronic obstructive pulmonary disease (COPD) and other chronic respiratory diseases, mild: 0.019 (0.011–0.033) COPD and other chronic respiratory diseases, severe: 0.408 (0.273–0.556)
Neurologic disorders	Dementia, mild: 0.069 (0.046–0.099) Multiple sclerosis, moderate: 0.463 (0.313–0.613) Parkinson's disease, severe: 0.575 (0.396–0.730)
Mental, behavioral, and substance use disorders	Alcohol use disorder, very mild: 0.123 (0.082–0.177) Anxiety disorders, moderate: 0.133 (0.091–0.186) Anorexia nervosa: 0.224 (0.150–0.312)
Hearing and vision loss	Hearing loss, mild: 0.010 (0.004–0.019) Hearing loss, profound, with ringing: 0.277 (0.182–0.387) Distance vision, severe impairment: 0.184 (0.125–0.258)
Musculoskeletal disorders	Low back pain, moderate: 0.054 (0.035–0.079) Neck pain, severe: 0.229 (0.153–0.317) Gout, acute: 0.295 (0.196–0.409)
Injury	Burns, lower airway, with or without treatment: 0.376 (0.240–0.524) Crush injury, short or long term, with or without treatment: 0.132 (0.089–0.189) Concussion: 0.110 (0.074–0.158)
Other	Abdominopelvic problem, mild: 0.011 (0.005–0.021) Anemia, moderate: 0.052 (0.034–0.076) Hypothyroidism: 0.019 (0.010–0.032)

* Health states included in this table are only examples. The full list of 235 unique health states and their disability weights are available in Salomon et al., 2015.
Data from Salomon, J. A., Haagsma, J. A., Davis, A., de Noordhout, C. M., Polinder, S., Havelaar, A. H., et al. (2015). Disability weights for the Global Burden of Disease 2013 study. *Lancet Global Health, 3*(11), e712-723. doi: 10.1016/S2214-109X(15)00069-8

for treated and untreated groups). This approach was revised for the 2010 iteration of the study, for which the process included empirical studies comprising simple paired questions, and survey of the general public through household surveys in countries such as Bangladesh, Indonesia, Peru, Tanzania, and the United States, as well as a web-based survey in English, Spanish, and Mandarin. The result was a categorization that included 220 health states (Salomon et al., 2012). Generally, for most conditions a reasonable degree of consensus can be reached within broad categories (e.g., 25% disabled as compared with 50%), but efforts to reach much finer distinctions have proved equivocal. The need to seek out more refined scales for purposes of health program decision making ought to be a national or local decision.

The third component of disability is its *duration*. The duration is generally counted from onset of the disability until cure, recovery, or death. Sometimes there is continuing permanent disability after the acute phase is completed; in such a scenario, the duration would be the remaining life expectation from the time of onset of disease.

Data for Decisions

In the collection and assessment of information, the level of precision required should be guided by the purpose for collecting the information and depend on the decisions to be taken. Even rough estimates may be helpful; though disconcerting to some, the time and cost of efforts to realize further precision need to be justified by the increased precision's potential impact on decision making. LMICs, with their scarce resources, need timely and appropriate information to plan and implement health interventions that maximize the health of their populations. Methods, indicators, and assessments of disease must support and contribute to this primary purpose of health systems.

Decisions concerning deployment of interventions against diseases and underlying risk factors ideally should be taken such that maximum healthy life per resource expenditure is obtained in an equitable, fair, and just fashion. The ultimate reason for obtaining health data is to have the information to guide such decision making.

Summary Measures of Population Health

This section focuses on the major approaches used for developing composite measures of population health status that summarize mortality and morbidity occurring in a population through the use of a single number. It discusses the rationale for composite measures, reviews the origins of each major approach, examines methodological differences among these approaches, and outlines the advantages and limitations of each.

Rationale for Composite Measures

Rationing of healthcare resources is a fact of life everywhere; choices about the best use of funds for health must inevitably be made (Hyder, Rotllant, & Morrow, 1998; WHO, 2000). The global scarcity of resources for health care is a challenge for every country, rich and poor (Evans, Hall, & Warford, 1981; Figgis & Walters, 2015–2017; World Bank, 1993), but the realities in LMICs paint the choices in much starker terms. It is even more important for LMICs to choose carefully how to optimize health expenditures so as to obtain the most health in the most equitable fashion from these expenditures. Important tools under development to assist in making better choices for health spending are based on measures of the effectiveness of health interventions in improving health status in relation to their cost.

In most sectors, decisions on resource allocation are based on perceived value for money. The health sector, however, has had no coherent basis for determining the comparative value of different health outcomes (from different health programs). To make decisions about whether to put money into programs that reduce mortality in children, as compared with those programs that reduce disabling conditions in adults, a common denominator is needed. In recent decades, work has been carried out to develop composite indicators combining morbidity and mortality into a single measure that may serve as a common denominator for comparing different health outcomes. A common unit of measure for these different health outcomes is *time lost from healthy life*.

The most important reason for attempting to capture the complex mix of incommensurable consequences resulting from disease within a single number is the need to weigh the benefits of health interventions against their costs. Costs of health programs are expressed in a unidimensional measure, such as U.S. dollars; therefore, the benefits to be achieved from their expenditure should be expressed in the same manner. *Healthy lifetime* is a unidimensional measure that can be used to compress health benefits and losses into a single time dimension. An explicit, objective, quantitative approach should enable better budgetary decisions and permit resource allocation in the health sector to be undertaken in a more effective and equitable fashion.

Note that a composite indicator is simply a tool to be used to assist decision makers in resource allocation. Like any tool, it can be misused. Conclusions that are reached on the basis of these indicators must be carefully examined. Not only do problems arise in trying to put so many dimensions together, which inevitably may lead to distortions, but serious issues also emerge concerning the reliability and validity of the information on which these indicators are based. Thus, all the problems associated with determining causes of death, counting the number of cases of disease, and assessing the extent of disability from a condition will lead to uncertainties when these factors are added and multiplied together. The development of a single indicator consisting of a specific number implies deceptive stability about something that may actually be composed of fragile data. Continuing vigilance in how these data are obtained, compiled, and used is critical, and those responsible for using the tool must have a clear technical understanding of what lies behind the numbers and which underlying assumptions and limitations are associated with these approaches. Despite all of these caveats, alternative approaches to improved decision making leave even more to be desired.

Uses of Composite Indicators

Measures of health status that combine mortality and morbidity facilitate comparisons both within and across populations. They can be used to estimate the quantitative health benefits from interventions and serve as tools to assist in the allocation of resources. The development of such measures entails two major processes: the measurement of healthy life, including losses of time from premature mortality and disability; and the valuing of life, which incorporates issues of duration, age, extent of future life, productivity, dependency, and equity (Morrow & Bryant, 1995). The purpose of developing such measures and the need for refining them become clear if the following objectives are to be achieved:

- The use of such methods at the country level for evaluating the impact of diseases
- Their use in the allocation of resources within the health sector
- The generation of more relevant and useful data for policy makers

Understanding Summary Measures

Precursors of composite indicators have been discussed in the literature for decades and generally were developed to assist with prioritization of health issues. Usually these metrics were based on the measurement of losses of time, losses of productive time, income forgone, or other costs incurred as a result of diseases. The earlier indicators generally focused on economic losses and estimated time loss due to disease and converted these losses into a dollar value. Thus, these measures are more economic measures than disease burden measures.

Two types of composite summary measures have been developed: *health gap measures* (healthy life lost), such as healthy life years (HeaLYs) or disability-adjusted life years (DALYs), and *health expectancies*, such as disability-free life expectancy (DFLE) or health-adjusted life expectancy (HALE). Both types use healthy lifetime lost through disability and death as a common measure of the impact of mortality and nonfatal health outcomes. These two types of measures are complementary and can be studied using survivorship curves, as discussed by Murray and Lopez (1994) (**FIGURE 1-2**).

In Figure 1-2 (Murray et al., 2002), the darker line is the survivorship curve based on a standard hypothetical life table population that demonstrates the proportion (*y*-axis) of an initial birth cohort that remains alive at any age (*x*-axis). The area A + B is the total life expectancy at birth of this cohort. A part of this life is spent in full health (area A); the lighter line is the survivor curve of those persons in full health. Thus, area A represents time lived in full health, whereas area B is time lived in suboptimal health (with disability). Area C represents time lost due to mortality. The area of the complete rectangle (A + B + C) represents the ideal survivorship curve—the theoretical maximum of healthy life for a cohort who lived in full health until a maximum age when all died.

Health expectancies are summary measures that estimate expectancy of life in a defined state of health. Examples include DFLE, active life expectancy, and HALE. These indicators extend the concept of life expectancy to expectations of various states of health, not just of life per se. Health expectancies assign lower weights to life lived in less than full health on a scale of 0 to 1, in which full health is rated 1. In Figure 1-2, health expectancy is given by the following equation:

$$\text{Health expectancy} = A + f(B)$$

where f is some function that assigns weights to years lived in suboptimal health.

Health gaps are summary measures that estimate the difference between actual population health and some specified norm or goal. In Figure 1-2, that difference is indicated by area C (loss due to mortality) plus some function of area B—that is, survivorship with disability:

$$\text{Health gap (healthy life lost)} = C + g(B)$$

where g is some function that assigns weights to health states lived during time B. Weights range between 0, meaning no disability (full health), and 1, meaning complete disability (equivalent to death). Note that this measure is equivalent to healthy life lost based on the natural history of disease in a population as discussed in the section "Healthy Life Year" later in this chapter.

Although some believe that health expectancies such as the HALE indicator are more readily understood (because they are conceptual extensions of the

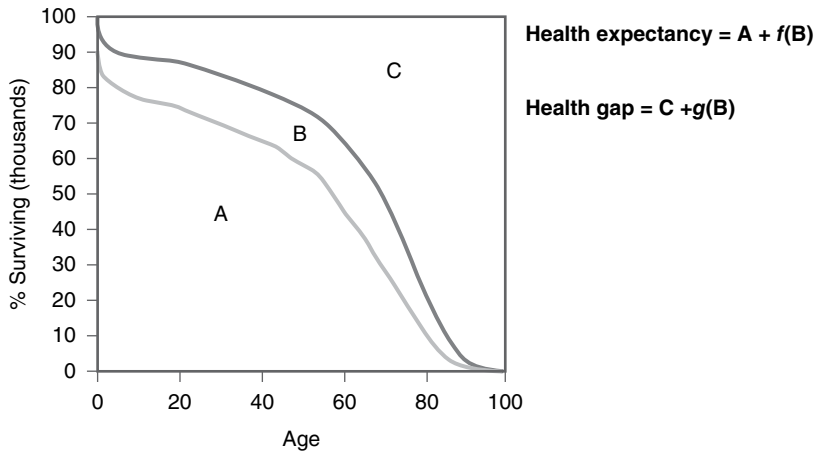


FIGURE 1-2 Survivorship curve of a hypothetical population showing health gaps and health expectancies.

Modified from Murray, C. J. L., et al. (2002). *Summary measures of population health*. Geneva, Switzerland: World Health Organization.

widely used life expectancy measure), health gap measures have important advantages for the purposes of health policy, planning, and resource allocation decisions. Both HeaLYs and DALYs are developed on the basis of disability and death attributable to a specific disease in an individual person. In their construction, great care is taken to ensure that there is *categorical attribution* using the ICD, so that each event (death or disability) is mutually exclusive and collectively exhaustive. With these measures, therefore, summing deaths and disabilities from each disease provides the total amount of death and disability for the population (a property termed *additive decomposition*). Health gap measures have this property, whereas health expectancies do not (Murray et al., 2002).

Composite Indicators

A number of composite summary indicators for burden of disease assessment have been developed. We will focus on four of these indicators: three of the health gap type (the healthy life year, the disability-adjusted life year, and the quality-adjusted life year) and one of the health expectancy type (HALE). In addition to measures of morbidity and mortality per se, these composite indicators may incorporate certain social value choices either explicitly or implicitly: the choice of life expectancy tables, valuing future life as compared with present life, valuing life lived at different ages, valuing social or economic productivity, and valuing equity in relation to cost-effectiveness. These social value choices are discussed later in this chapter (see the section "Valuing Life: Social Value Issues"), but because some social value choices are integral to the calculations of some composite indicators, they are briefly mentioned in this section.

Healthy Life Year

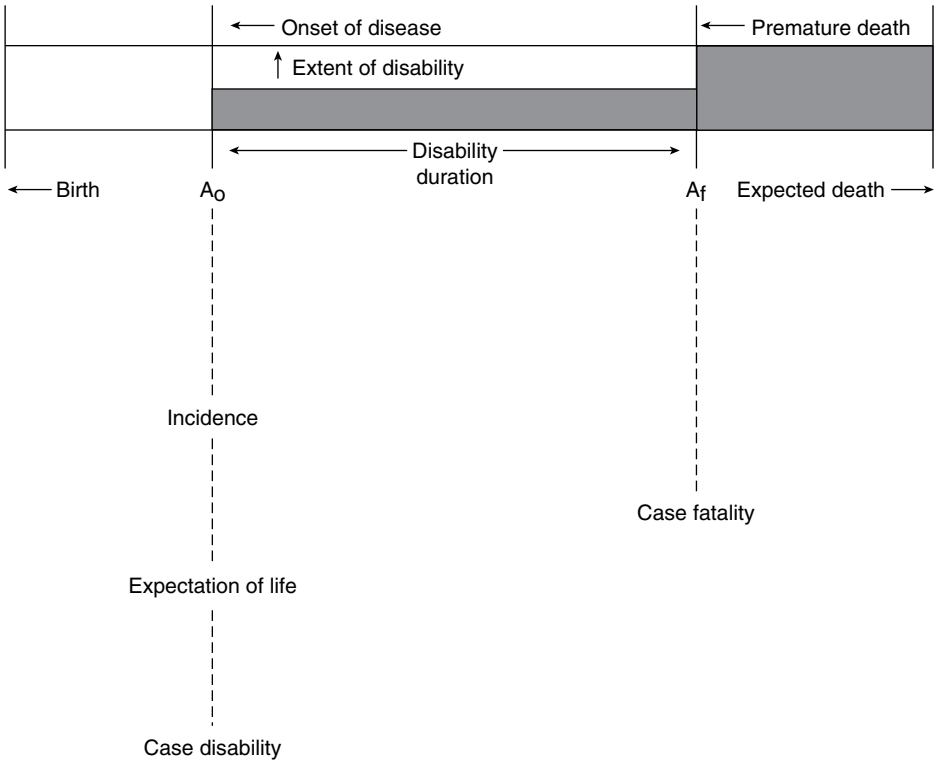
The healthy life year (HeaLY) is a composite measure that combines the amount of healthy life lost due to morbidity with that lost due to death—that is, loss of life expected had the disease not occurred (Hyder et al., 1998). We discuss the HeaLY first because it is conceptually straightforward, serves as a prototype for other health gap indicators, and was the first of the composite measures to be used as a tool in national health planning (Ghana Health Assessment Team, 1981).

Measuring the loss of healthy life from disability is more challenging than measuring the comparable loss from death, and many approaches have been used (Murray & Lopez, 1994). To incorporate loss from disability in a composite measure, such a loss must have comparable dimensions to that for life lost due to death. The HeaLY includes three components for disability: case disability ratio (comparable to the case fatality ratio), extent of disability, and duration of disability. The CDR and duration of disability can be determined objectively, but assessment of the extent of disability, which usually ranges from 0 (no disability) to 1 (equivalent to death), has a substantial subjective element (Morrow & Bryant, 1995).

The healthy life approach focuses on knowledge of the pathogenesis and natural history of disease (Last, Spasoff, & Harris, 2000) as the conceptual framework for assessing morbidity and mortality and for interpreting the effects of various interventions (**FIGURE 1-3**).

The onset of disease usually will be dated from the start of symptoms or signs, as determined by the individual afflicted, a family member, or a medical practitioner, or as the result of a lab test. Several different patterns of disease evolution are possible, of course. **FIGURE 1-4** illustrates healthy life lost from disability and premature death due to typical cases of cirrhosis, polio, and multiple sclerosis, respectively, in terms of onset, extent and duration of disability, and termination. The conclusion of the disease process depends on the natural history of the disease as modified by possible interventions. The possible outcomes include clinical recovery (the complete disappearance of clinical signs and symptoms), progression to another disease state (such as chronic hepatitis progressing to cirrhosis), and death. The last outcome includes death directly caused by the disease as well as death indirectly brought on by the disease as a result of disability.

The definitions of variables and formulas to calculate HeaLYs are provided later in this section and summarized in **TABLE 1-4**. Each disease will have a distribution of ages at which onset or death may occur, but for most diseases the average age will provide a satisfactory approximation for a population. In view of the limitations of data, this is the starting assumption for the application of the HeaLY method in LMICs. Nevertheless, as with other choices in this method, if sensitivity testing indicates that the average age is not satisfactory, then estimates may be based on age distributions. Similarly, if the natural history of a disease or response to interventions is different in different age groups, then the disease can be specifically classified by age (e.g., neonatal tetanus as compared with adult tetanus, and childhood pneumonia as compared with adult pneumonia). In recurrent diseases or diseases with multiple episodes (e.g., diarrhea), age at onset denotes the average age at first episode.



Note: A_o = average age at onset; A_f = average age at death; ■ = healthy life lost.

FIGURE 1-3 The HeaLY model: Loss of healthy life from disability and death.

The expectation of life in HeaLYs were based on normative expectations of what should occur under usual circumstances at the time of that work. Women in Japan, who had the highest global expectation of life, approximated this norm with an expectation of life at birth of 82.5 years for females (Model Life Table West, level 26) (Coale, Demeny, & Vaughan, 1983; Coale & Guo, 1989).

The definition of disease ("dis-ease") makes the value of the case disability ratio 1 by default for most disease states, because all cases are disabled (to varying degrees and duration) if those persons have been labeled as diseased. For some conditions (e.g., sickle cell trait or HIV positivity) and risk factors, however, cases may not be considered diseased by definition, but the condition nonetheless needs to be assessed.

The duration of disability can be either temporary or permanent (lifelong). If the disability is temporary, then Dt is the duration of that disability until recovery (see Table 1-4). If the disability is permanent and the disease does not affect life expectation, then Dt is the expectation of life at age of onset of disease [Dt = E(Ao)]. If the disability is permanent and the disease reduces life expectation, then Dt is the expectation of life at age of onset reduced by the difference between ages of fatality and onset [Dt = E(Ao) − (Af − Ao)]. A disability severity scale needs to be used to estimate extent (severity) of the disability (see Table 1-4).

The HeaLYs lost from death and from disability are added and expressed as the total years of life lost per 1,000 population per year. The loss is attributed to the year in which disease onset occurs and includes the stream of life lost from disability and death at any time after onset, even if these events happen many years later. This method offers a prospective view of the event (disease onset) and its natural history (or as modified by interventions) over time.

An important benefit of the HeaLY formulation is that the effects of different kinds of interventions can be readily explored to determine their expected gains in terms of healthy life. The HeaLY spreadsheet (available upon request from the authors) incorporates these concerns; it also includes options for considering the proportion of the population that will be covered by an intervention and allows for different levels of coverage for different segments of the population for each intervention.

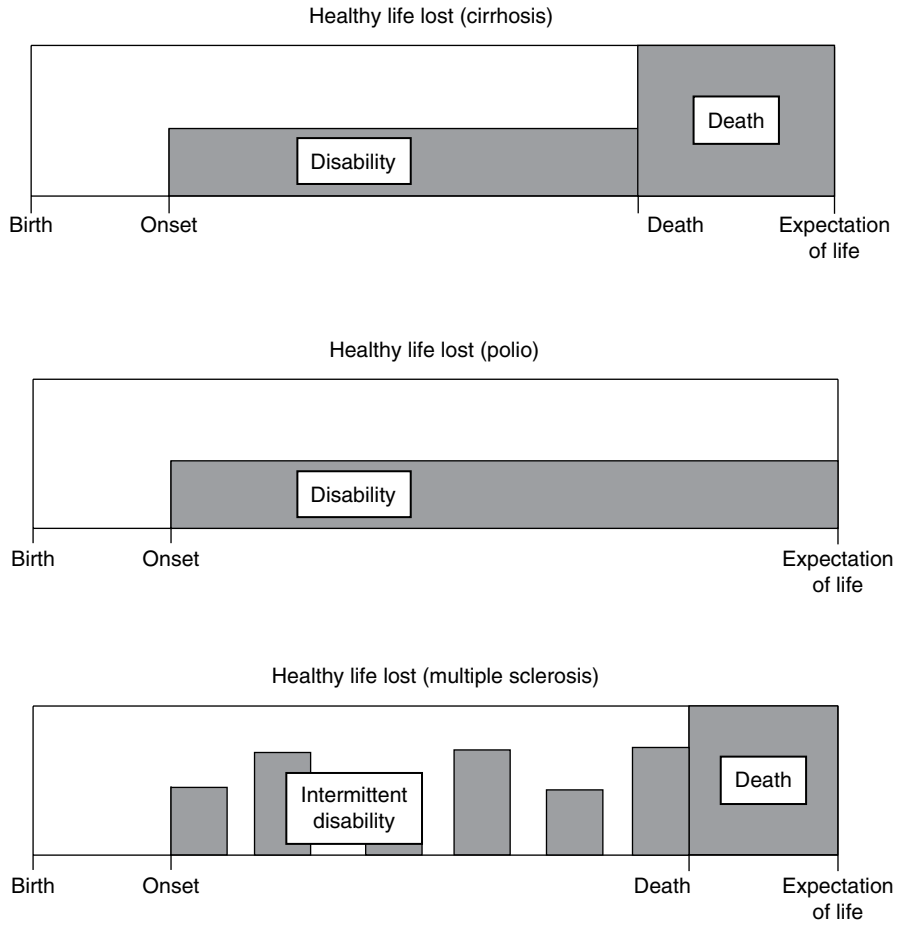


FIGURE 1-4 Different patterns of healthy life lost.

Reproduced from Hyder A., Rollant G., and Morrow R. H. (1998). Measuring the burden of disease: Healthy life-years. *American Journal of Public Health, 88*(2), 196–202. Figure 1, p. 197.

Disability-Adjusted Life Year

The disability-adjusted life year (DALY) is a health gap population summary measure that combines time lost due to disability with that lost due to death (life that would have been expected had the disease not occurred), in a manner similar to the HeaLY measure. It first appeared in the World Development Report of 1993 and has become the most widely used composite measure of population health (Global Burden of Disease Risk Factors Collaborators, 2016; Global Burden of Disease Risk Factors Collaborators et al., 2015; Jamison et al., 2006; Lim et al., 2012; Lopez et al., 2002; Murray & Lopez, 1994; Murray et al., 2002).

DALYs are calculated as two separate components for the measurement of life lost due to disease, and they may also directly include three social value choices. The two components are (1) years of life lost (YLL), referring to the loss of healthy life from death, and (2) years of life lived with disability (YLD), referring to the loss of healthy life from disability. Thus

$$\text{DALY} = \text{YLL} + \text{YLD}$$

The social value choices that may be included in DALYs are (1) life expectation values, (2) discount rates for future life, and (3) weighting for life lived at different ages, as discussed later.

Since the GBD 2010 study (Lozano et al., 2012), age weighting is no longer a default value choice for the DALY. Instead, users have the option to calculate these values with or without age weighting, as well as with or without discounting. The following three options could be applied: (1) use both age weights and discounting, (2) use either age weights or discounting, and (3) use neither age weights nor discounting.

The calculation for YLL in a population uses the age distribution of all deaths by cause in one year multiplied by life expectation at each age to estimate the loss of life for each disease that would have been expected if not for that disease. The expectation of life can be obtained either from a model life table (Coale & Guo, 1989) or based on the best achievable low levels of mortality such as those found in Japan; thus the DALY, as does the HeaLY, directly incorporates this social value choice.

TABLE 1-4 Variables for Estimating Healthy Life Years (HeaLYs)

Symbol	Explanation	Expression
I	Incidence rate per 1,000 population per year	/1,000/year
Ao	Average age at onset	years
Af	Average age at death	years
E(Ao)	Expectation of life at age of onset	years
E(Af)	Expectation of life at age of death	years
CFR	Case fatality ratio: proportion of those developing the disease who die from the disease	0.00–1.00
CDR	Case disability ratio: proportion of those developing the disease who have disability from the disease	0.00–1.00
De	Extent of disability (from none to complete disability equivalent to death)	0.00–1.00
Dt	Duration of disability in years	years
	Disability can be either permanent or temporary:	
	■ If temporary, then Dt = duration of that disability (i.e., until recovery or death)	
	■ If permanent and disease does not affect life expectation, then Dt = E(Ao)	
	■ If permanent and the disease does reduce life expectation, then Dt = Af − Ao	
HeaLY	Healthy life years lost per 1,000 population per year: $I \times \{[CFR \times \{E(Ao) - [Af - Ao]\}] + [CDR \times De \times Dt]\}$	HeaLYs per 1,000 per year

For disability, the DALY uses estimates of incidence, duration, and severity to calculate the time lived with disability (YLD) for each disease. The YLD component equals the number of incident cases in the period multiplied by the average duration of disease multiplied by a weight factor for the degree of severity (extent) of the disease. A description of the severity scale used in one version of DALY was given earlier in this chapter, in the section on measurement of disability (see Table 1-4).

The second social value choice, which was directly incorporated in the original version of DALY, is the discount rate of 3% per annum. This social time preference has been used for most estimates; recently, DALY results discounted at 0% have also become available.

The third social value choice concerns weighting life lived at different ages. Earlier DALYs were age weighted according to an arbitrary exponential curve designed to give the most value to life lived as a young adult (Hyder et al., 1998; World Bank, 1993). Weighting by age was the most controversial component of the DALYs when they appeared and caused great dissent from other health professionals (see the section "Valuing Life Lived at Different Ages" later in this chapter). Recent DALY listings from GBD studies also include results with no age weighting (all years equally valued). It has been argued that age weighting

of DALYs does not affect final results, but this depends on the purpose for making the estimates and has been challenged (Anand & Hanson, 1997; Barendregt, Bonneux, & Van der Maas, 1996; Barker & Green 1996; Hyder et al., 1998).

An important difference between the HeaLY and the DALY is the fact that the starting point for the HeaLY is the onset of disease; that is, the loss of healthy life is based on the natural history of the disease (as modified by interventions), illustrated in Figures 1-2 and 1-3. This is true for the YLD component of the DALY, but the YLL is based on mortality in the current year. In a steady state, there is no difference in these perspectives. However, when incidence is changing—such as with HIV in many parts of the globe—the DALY approach can potentially understate the true situation (Hyder & Morrow, 1999).

The calculation for DALYs can be expressed in the form of an integral that was first published in the World Bank literature (Murray & Lopez, 1994). This single equation incorporating all technical and value choices had the advantage of standardization to ensure comparability of the multiple calculations undertaken in the GBD studies, and greatly facilitated the actual computations. Nevertheless, for national and local priority setting, it may be preferable to use an indicator constructed such that the social value choices can be adjusted to suit national and local preferences (Bobadilla, 1998; Hyder et al., 1998; Morrow & Bryant, 1995). Recent DALY formulations allow for this possibility; indeed, it is useful to think of DALYs as a family of related measures using terminology specifying the following formulation: DALY (*r, K*) uses a discount rate of *r* and age weighting indexed to *K*. Other parameters can be added in a similar fashion (Jamison et al., 2006).

HeaLYs and DALYs are both "health gap" measures and can be considered part of the same family of measures. In fact, DALYS exactly equal HeaLYs when the following conditions are met: (1) the condition in question is in steady state or equilibrium (that is, the incidence, CFR, and disability variables remain constant during the time intervals under consideration); (2) age weighting is not applied ($K = 0$); and (3) the same measures of disability (weights) are used.

Quality-Adjusted Life Year

The quality-adjusted life year (QALY) was introduced in 1976 to provide a guiding principle for selecting among alternative tertiary healthcare interventions (Zeckhauser & Shepard, 1976). The idea was to develop a single measure of quality of life that would enable investigators to compare expected outcomes from different interventions—a measure that valued possible health states both for their quality of life and for their duration.

The central notion behind the QALY is that a year of life spent in one health state may be preferred to a year spent in another health state. This generic measure sums time spent in different health states using weights on a scale of 0.00 (dead) to 1.00 (perfectly healthy) for each health state; it is the arithmetic product of duration of life and a measure of quality of life (health state weight). For example, 5 years of perfect health = 5 QALYs; 2 years in a state measured as 0.5 of perfect health followed by 5 years of perfect health = 6 QALYs.

The QALY was originally developed as a differentiating indicator for individual choices among tertiary healthcare procedures, not as a measure of disease burden in a population. It was used to assess individual preferences for different health outcomes from alternative interventions (Morrow & Bryant, 1995). The QALY, too, comprises a large family of measures. Since its introduction, a wide variety of QALY measures have been developed, along with a voluminous literature on alternative methods incorporating a range of disability domains and a diversity of methods to assign weights to generate QALYs (Kaplan & Anderson, 1996; Nord, 1993). The most widely used measure is the EQ-5D (European Quality of Life with Five Domains and three levels of quality for each domain; www.euroqol.org).

Perhaps the most important use of QALYs has been as a common denominator to measure utility in cost-utility analysis (and effectiveness in cost-effectiveness analysis) to assist in resource allocation among alternative health interventions by ranking interventions in terms of cost per QALY (Kaplan & Anderson, 1996; Nord, 1992; Torrance, 1986). An early and widely publicized attempt to make the best use of healthcare resources by maximizing QALYs per dollar expended was the well-intentioned but rather unfortunate effort undertaken in Oregon in the early 1990s (**EXHIBIT 1-4**).

In the United Kingdom, as part of its 1997 National Health Service (NHS) reforms, the National Institute for Clinical Excellence (NICE; www.nice.org.uk) was created to advise public health officials about the effectiveness and cost-effectiveness of various health interventions. In an explicit attempt to introduce economic considerations in addition to medical judgments for the allocation of resources, NICE has produced a large collection of studies on the cost per QALY produced by the interventions it appraises. Some of these appraisals have been the source of considerable controversy. If a treatment is considered cost-effective for a group of patients,

EXHIBIT 1-4 Oregon: Historical Application of the QALY for Allocation of Resources

An early and well-known attempt to apply the QALY approach for allocation of health resources occurred in the state of Oregon (Blumstein, 1997). In 1988, Oregon faced a budgetary shortfall for its Medicaid program, and coverage for organ transplants was denied. In an effort to prioritize its health services, Oregon undertook a bold attempt to explicitly ration health services. A coalition including consumers, healthcare providers, insurers, business, and labor representatives launched a broad and courageous healthcare reform. It began with a series of "experiments" in which the decision-making process was based on a cost-effectiveness approach using quality of well-being (QWB—essentially a QALY) for comparing the outcomes of treatment options among people.

The initial list, published in 1990, consisted of 1600 condition/treatment pairs drawn up as follows:

Cost-effectiveness ratio = cost of services / (health gain × duration)
Cost of services = charges for treatment including all services and drugs
Quality of well-being (QWB) = sum of QWB weight (W) × each QWB state × probability that symptoms of that QWB state would occur
Health gain = QWB with treatment − QWB without treatment

From the beginning, there was great opposition to the very notion of rationing; consequent denial of services to those who had conditions that did not make the list contributed to the rancor. There were also unfortunate technical blunders in the generation of the first list. For example, treatment for thumb sucking was ranked higher than hospitalization for starvation, and treatment for crooked teeth higher than early treatment for Hodgkin's disease. Such inconsistencies, together with objections raised by groups advocating for the disabled, gave rise to alternative approaches for establishing rankings.

Although enormous public effort went into the reform and much was accomplished, the explicit cost-effectiveness approach with QALYs as the outcome measure was eventually dropped (Blumstein, 1997; Eddy, 1991; Morrow & Bryant, 1995; Nord, 1993).

NICE will recommend its use throughout the NHS; if not, it will recommend against its use in the NHS. The use of these cost-effectiveness studies as an aid to decision making is intended to increase the total healthcare benefits gained from the money spent by the NHS.

The QALY as originally used is essentially equivalent to the YLD of the DALY; in fact, it would be exactly the same as the YLD when the following conditions are met: (1) there is no discounting ($r = 0$); (2) there is no age weighting ($K = 0$); and (3) the same disability weights are used. More recently (as used in some cost-effectiveness studies), QALYs have incorporated life expectation as well.

Health-Adjusted Life Expectancies

Several types of health expectancies exist in the literature. During the 1990s, *disability-free life expectancy* (DFLE) and related measures were calculated for many countries (Robine, 1994; Mathers et al., 2001). However, these measures incorporate a dichotomous weighting scheme in which time spent in any health state categorized as disabled is assigned, arbitrarily, a weight of zero (equivalent to death). Thus, DFLE is not sensitive to differences in the severity distribution of disability in populations. In contrast, *disability-adjusted life expectancy* (DALE) adds up expectation of life for different health states with adjustment for severity weights. In 2001, WHO replaced the DALE terminology with *health-adjusted life expectancy* (HALE); the latter term will be used throughout the remainder of this text.

The HALE is a composite summary measure of population health status that belongs to the family of health expectancies; it summarizes the expected number of years to be lived in what might be termed the equivalent of "full health." WHO has used it as the measure of the average level of health of the populations of member states for annual reporting on population health for a few years (WHO, 2000).

Health expectancy indices combine the mortality experience of a population with the disability experience. The HALE is calculated using the prevalence of disability at each age so as to divide the years of life expected at each age (according to a life table cohort) into years with and without disability. Mortality is captured by using a life table method, while the disability component is expressed by additions of prevalence of various disabilities within the life table. This indicator allows an assessment of the proportion of life spent in disabled states. When compared with the total expectation of life, it translates into a measure of the total disability burden in a population.

Comparison of the various methods and specific indicators is available in the literature (Robine, 1994). Alternative methods are given in WHO's *National Burden of Disease Studies* manual (Mathers et al., 2001a). As originally designed, the HALE does not relate to

TABLE 1-5 Comparisons of Original Versions of Composite Summary Measures of Population Health

	Quality-Adjusted Life Years	Healthy Life Years	Disability-Adjusted Life Years	Health-Adjusted Life Expectancy
Origin	North America, 1976	Ghana Ministry of Health, 1981	World Bank development report, 1993	World Health Organization report, 2000
Purpose	Assess individual preferences for various outcomes from complex interventions	Assist in resource allocation decisions	Compare disease burdens in many different populations	Compare national disease burdens
Level of use	Personal decisions	National- and district-level decisions	Broad policy decisions	Global comparisons
Data	Tertiary hospital data and personal interviews	National and local data from multiple sources; expert review	Global data and expert opinion	Global data and expert opinion
Original discipline base	Economists, clinicians	Epidemiologists, clinicians, national planners	Economists, statisticians	Demographers, economists, statisticians
Social values that may be incorporated	Generally not included	Future life discounted	Age weighting (not mandatory), future life discounted	Not relevant

Reproduced from North America, 1976; Ghana Ministry of Health, 1981; World Bank Development Report, 1993; World Health Organization Report, 2000.

specific diseases, but rather to the average extent of disability among that proportion of each age group that is disabled. The lack of correlation between a condition or disease entity and the measure made it less valuable for resource allocation and cost-effectiveness calculations.

Summary

TABLE 1-5 summarizes these summary measures (original versions) in terms of origins, purposes, level of use, sources of data, and disciplinary background of originators.

Valuing Life: Social Value Issues

The very idea of valuing some lives more than others is jarring, yet these notions are regularly reflected in our actions. The value of life is often implicit in the way resource allocation decisions are made; therefore, as much as possible such decisions should be explicit, open, and transparent. Many thoughtful people have serious reservations about assigning a single number to such a complex multidimensional phenomenon as health. But what is the alternative for use as a measure of utility or effectiveness in economic analyses? Outcome measures must be expressed as a unidimensional measure to be comparable to unidimensional monetary expenditure units for costs. (However, decisions about allocation should not be made on a mechanical basis; other factors, including the effect on equity, may need to be considered in decisions in addition to the goal of maximizing healthy life per unit expenditure.)

To construct composite measures of population health, important social value choices must be made. Choices about which expectation for life should be used and about valuing life lived at different ages, valuing future life as compared with the present, valuing life in terms of economic and social productivity, and valuing equity in relation to efficiency—all raise major ethical concerns.

Expectation of Life

Years of life lost due to death and to chronic disability are based on life expected had the disease not occurred. To estimate the expectation of life in a population, a choice must be made between using a local, national, or model life table. This choice should be determined by the purpose of the study. For example, for assisting in national and local decision making, it may be more suitable to use national life tables based on the mortality and fertility of the population in question than to use model life tables. Conversely, a model life table might be selected to reflect the best health state possible in the world, as this selection may allow a fair comparison with other countries. For example, from a global perspective it would be unfair to use national life tables to compare gains that could be achieved from a particular intervention in Ghana with those in the United Kingdom, even if both costs and lives saved were the same in each country. The reason is that those lives saved in Ghana would have a lower life expectancy than those in the United Kingdom, resulting in less healthy life saved for the same expenditure. From the global viewpoint in this example, the priority would be to fund the intervention in the United Kingdom because it would produce more healthy life per expenditure than for Ghana.

Valuing Life Lived at Different Ages

Age weighting refers to the valuing of a year of life according to the age at which it is lived. This practice immediately raises questions about the basis for valuing human life. Does a day of one person's life have the same value as a day of anyone else's life? Does the value vary with age, economic productivity, or social status? Should life itself be valued separately from what is done with that life?

The Ghana Health Assessment Team (1981) judged that all human life was intrinsically valuable and that a given duration of any life was equal in value to that of any other life. The valuing of a year of life equally, irrespective of age, has been considered egalitarian (Busschbach, Hesing, & de Charro, 1993; Morrow & Bryant, 1995). This choice was incorporated into the development of the HeaLY approach: A year of life lived at any age is equally valued.

The original 1990 DALY formulation assigned an exponential function to provide a value chosen so that life lived as a dependent (e.g., infants, children, the elderly) was given less value than life lived during the productive years. With this approach, the intrinsic value of life increased from zero at birth to a maximum at age 25 and declined thereafter, so that a day of life of a 50-year-old was worth about 25% less than that of a 25-year-old. Paradoxically, the age weighting used in the original DALY formulation led to higher valuation of life lived before age 15 than did the HeaLY formulation, in which life lived at all ages has equal value (Barendregt et al., 1996; Hyder et al., 1998). Current formulations of the DALY leave age weighting as an option, and such weighting is not used with the HALE.

Age-related valuing has been justified by studies showing that individuals value their own life lived at different ages differently. Such values have been reported in the literature, and classic studies have reported that they are consistent across respondents of different ages (Busschbach et al., 1993). In addition, studies from many countries reveal a preference for saving younger lives as compared with older ones (Murray & Lopez, 1994). Nevertheless, it was not clear how much of the differential valuing of life at different ages is related to an underlying appreciation that economic and social productivity varies at different ages. If it is decided at the local or national level that healthy life should be valued according to economic and social productivity, then an alternative to age weighting might be to explicitly add a productivity factor or subtract for the societal costs of dependents, such as education (see the section "Valuing Life for Its Economic and Social Productivity").

Valuing Future Life Compared with Present Life: Discounting

Discounting is the process for determining the present value of future events. *Social time preference* takes into account the phenomenon that people value events at present more highly than those in the future (independent of inflation and of uncertainty). For investments in other sectors, time preference is normally taken into account by discounting future returns and costs by some appropriate discount rate. Thus, the discount rate can be considered the inverse of an interest rate. The main issue concerning discounting in relation to summary measures is whether discounting life itself is appropriate. There seems to be little question about the usefulness of discounting the future value of what is produced by healthy life, but should the life itself be discounted (Morrow & Bryant, 1995)?

Discounting has been applied in the health sector because both the losses from a disease and the benefits from a health intervention often occur in the future. An intervention today may not produce immediate benefits (such as in immunization), or it may result in benefits being sustained over a long time (such as in supplementary nutrition). The costs for these activities must be borne now, but the benefits are realized in the future and are less valuable than if they could

occur now. This is equivalent to investing money now so as to obtain more in the future. Thus, a HeaLY now has greater intrinsic value to an individual or community than one in the future (Gold, 1996; Weinstein, Siegel, Gold, Kamlet, & Russell, 1996).

The rate at which society is supposed to discount benefits, termed the social discount rate (SDR), is a numeric reflection of societal values regarding intertemporal allocation of current resources. There is no consensus about the most appropriate choice of a discount rate in health, but most agree that it should be lower than that used in the private commercial sector. Historically, the 1993 World Development Report and the first GBD studies used a discount rate of 3% per year (more recent iterations of the GBD studies do not use discounting in the default calculations, but instead have it as an optional parameter that can be set by the user); in lieu of other information, this rate has been adopted in many international public health cost-effectiveness studies. Nevertheless, the impact of using a range of discount rates, including zero, should be explored with each study.

Valuing Life for Its Economic and Social Productivity

Whether and how to value economic and social productivity for purposes of healthcare decision making is highly contentious; to a large extent, the age weighting incorporated in the original DALY formulation was considered by many to be a proxy for productivity. Such valuations should be considered separately, made explicit, and very much dependent on the purpose of the valuations. In general, productivity may be attributed to adults aged 15 to 64, and persons in these age groups could be given a higher value. Persons younger than age 15 and older than age 65 may be considered as dependents and given a lower value. Many variations for differential valuing are possible, including type of employment. People at different socioeconomic levels in a society are expected to have different capacities for productivity—yet, to value life according to income levels or social class would not seem fair and generally would not be acceptable. In LMICs, the value of marginal wages for subsistence agriculture is negligible, but the value of the workers' lives certainly is not.

A fundamental question is whether to consider adding a productivity component to the summary measures. Health issues do not readily conform to the requirements of market economics; information is inadequate, and misinformation is rife on the part of the providers as well as the public. Externalities from good health are generally large; demand for costly services is largely determined by the healthcare providers rather than by the consumers, and competitive market forces have not worked well for those in greatest need. In the private sector, demand for services is clearly related to productivity and willingness (and ability) to pay. However, if left to market forces alone, inequitable distribution would be inevitable.

Economic arguments have been put forward for valuing life according to productivity, but counter-claims have been made that human life cannot and should not be expressed in economic terms for decision-making purposes. Nevertheless, efforts to avoid such expression result in implicit valuation of life. More than 30 years ago, Barnum (1987) argued for adding productivity to the valuing of human life, stating that it has been ignored in health policy, is readily quantifiable, and does not ignore the welfare of children because the whole population is dependent on adult productivity for quality and sustenance. Such an economic appraisal of human life is often based on the net transfer of resources from the "producers" to the "consumers" and the consequent interdependence of people. In relation to this issue, in the *Report of the Commission on Macroeconomics and Health* (WHO, 2001), a DALY was stated to be worth at least an average annual income per head, although the basis for such a valuation was not adequately justified. More work on explicit valuations of human life and what that life produces is needed, and will certainly affect health-related cost-effectiveness decisions.

Valuing Equity in Relation to Efficiency

A child born in Malawi or Uganda will likely live only half (approximately) as long as one born in Sweden or Singapore; many babies born in Niger or Sierra Leone will not live to see their fifth birthday. These inequalities are unfair and harmful and, therefore, qualify as inequities. In terms of social justice, equity has to do with a fair distribution of benefits from social and economic development. However, the term *equity* is used in different conceptual senses: equal access to health services for all (opportunity equality), equal resources expended for each individual (supply equality), equal resources expended on each case of a particular condition (equality of resource use to meet biological need), equal healthy life gained per dollar expended (cost-effectiveness), care according to willingness to pay (economic-demand equality), care according to biological or socioeconomic need, and equal health states for all.

Decisions based on cost-effectiveness (e.g., cost per HeaLY), therefore, may not accord well with concerns about equity. These calculations are generally indifferent to equity; they are designed to steer interventions to what is efficient, whatever the differential need may be. To meet the requirements of equity,

health system planners need to go beyond ensuring equality of access to health care and require a balance so that health system responses are in accord with equity as well as efficiency.

Provided that health information is available according to socioeconomic and vulnerable groups, use of these summary indicators as tools for determining equity by calculating healthy life per dollar to be gained by all socioeconomic and vulnerable groups could readily be undertaken. It would be important to assess the impact of specific health decisions to ensure that they enhance equity. Summary measures such as HeaLYs and DALYs can be used to guide allocation of resources to ensure equitable distribution of those resources to reach those most in need. Cost-effectiveness by itself does not provide adequate guidance; equity should also be an associated criterion to govern the distribution of societal benefits.

Data for Composite Measures

Types of Data

The data needs for estimating the burden of disease in a region or country are extensive, and obtaining even reasonable estimates in LMICs has been a source of concern (Anand & Hanson, 1997; Barker & Green, 1996; Bobadilla, 1998; Murray et al., 2002). Brief descriptions of the types of data required follow. Note that any available data need to be carefully reviewed and optimally utilized.

Demographic Data. Population data are integral to burden of disease estimations and are needed both as denominators and for consistency checks. In a national setting, a recent census is useful for providing population counts by age, sex, and geographic location. Particularly helpful, when there is inadequate death registration, is to have a one-year post-census follow-up on a sample of enumeration areas to obtain robust age, sex, and place mortality. The age and sex distribution of the population is critical, and often is a major factor that determines the nature of the disease burden. A good vital registration system is a key asset that will provide both birth and death numbers. Underreporting, age misreporting, and other biases in data may have to be addressed (using standard demographic methods) prior to use of these data in burden of disease estimation.

Mortality. Mortality data are required for any burden of disease analysis. Specifically, age, sex, and place mortality rates greatly assist the analysis by defining the contribution of mortality to the pattern of disease burden. They also serve as an essential framework that constrains estimates obtained from a variety of special studies that fill important information gaps but may be incomplete or biased in the populations covered. Reporting errors, such as underreporting of deaths and failure to report age at death, need to be carefully examined. In particular, information must be evaluated for deficiencies in the under-5 group and older age groups. For the youngest ages, the probabilities of deaths in the first year (1q0) and in the next four years (4q1) provide better estimates of the risk of death than do overall mortality rates. Methods such as the Brass method for indirect estimates of mortality provide useful ways to assess age-specific mortality data for potential errors (Hill, 2001).

For burden of disease studies, cause of death data are required for all ages (except stillbirths, which are not included—a limitation of composite measures), but reliable cause of death records are often not available in LMICs, especially for deaths that do not occur in healthcare facilities. Even if available, the classification system used may be outdated and not based on the ICD, and the reliability of coding may vary by the type and location of the hospital. Young-adult deaths may be better recorded than deaths of infants and the elderly. Especially in low-income countries, it can be helpful to cross-check death records with other information, using postmortem interviews and hospital registers to assist in defining causes of death or to extrapolate from other data or other regions to assist in making estimates.

Morbidity. Meaningful data on disability are even more difficult to find and interpret than mortality data. Often morbidity information is institution-based or restricted to one or two sources, such as hospital inpatient and clinic outpatient records. The representativeness of small studies and the range and types of morbidity covered in any survey need careful evaluation. National disability surveys or regional studies conducted for the evaluation of disabled people may be available; such research is useful in providing some estimate of the prevalence of serious disabilities and their age and sex distribution. However, linkage between disability and disease is often not available, and attributing one type of disability to specific causes is difficult. For example, because many conditions can lead to blindness—for example, diabetes, hypertension, injuries, trachoma, and cataracts—the attribution of proportions of blindness in a population to its cause can be problematic. Information on the duration of disability may be found in specialized studies and the experience of institutions. The severity of disability will have to be rated on a scale; the various methods used in the literature were described earlier in this chapter.

Variables

The types of data just described need to be processed in the form of specific disease-based estimates. The key variables are defined in Table 1-4. The *incidence* rate (usually expressed per 1,000 general population per year) is central to the natural history of disease concept. Although incidence is a basic epidemiologic indicator, it is usually not found in routine data collection systems. Special studies, prospective surveys, or calculations based on the prevalence (which is more commonly available than the incidence) and knowledge of the average duration of the disease can be helpful in developing this measure.

The case fatality rate is the proportion of those developing the disease who die from it at any time. It is expressed as a decimal value between 0 (for nonfatal conditions) and 1 (for universally lethal conditions). The *case disability ratio* (analogous to the CFR) is the proportion of those diagnosed with a disease who have signs or symptoms, and is usually 1 (as discussed earlier). *Age* is required in various formats. Age at onset is when disease onset occurs in a population; age at fatality denotes the age at death as a result of the disease. The *expectation of life* at age of onset is the years of life expected at that age had the disease not occurred. Similarly, expectation of life at fatality is the years of life expected at that age had the death not occurred.

Checking Data

Data used for generation of indicators need to be evaluated for validity, reliability, and consistency, using defined qualitative and quantitative criteria. Large population-based studies may be given preference over smaller sample-based work if both are available and the quality of their data is comparable. Better conclusions may be possible by cross-checking different sources of data. Community-based studies, which may be representative of the population but have limited diagnostic validity, may be compared with hospital-based work, in which diagnosis may be valid but would come from a biased population sample. The following subsections profile simple types of checks for data quality.

Comparison of Total Numbers. Cross-checks should be done to compare total numbers. It is essential to check that the number of deaths in a year in a region is the same as the sum of all deaths from all causes in the same region. Similarly, program-based data can be compared with data from other sources to ensure better estimates of causes of death. The comparison of totals allows one to work within a frame of mortality and avoids double counting of one death. However, it does not assist in the distribution of deaths within that frame.

Relationship Between Variables. Checks based on the epidemiologic relationship between parameters refer to the application of simple, yet vital, relationships such as the following:

- Prevalence (point) = incidence × average duration of disease
- Cause-specific mortality rate = incidence × case fatality rate

These checks allow estimates from different sources to be compared for internal consistency. Such relationships can also be used to derive one of the estimates in the equations when the others are known.

Sensitivity Analysis. Sensitivity analysis is a useful tool to determine whether data that are more precise are required for the purposes of a particular decision. A one-way sensitivity analysis (Petiti, 1994) evaluates the effect of manipulating one variable at a time on the dependent variable. If the outcome is sensitive to one or more variables, their precision is more important in the estimation.

Disease Groups: Classification

Murray and Chen (1992) introduced a disease group system based on WHO's ICD classification system. Group I includes conditions characteristic of low-income countries: communicable diseases, maternal and prenatal conditions, and nutritional deficiencies. These conditions decline at rates faster than overall mortality rates as socioeconomic conditions improve; thus, group I contributes to a relatively small percentage of deaths in the high-income world. Group II, which consists of noncommunicable and chronic diseases, accounts for most loss of healthy life in high-income countries and proportionately increases with the epidemiologic transition in LMICs. Group III consists of injuries, both intentional and unintentional (including violence).

The distribution of the disease burden among these three groups is one indicator of the type of disease burden and the level of epidemiologic transition in a country. It is important to distinguish between the proportions of deaths attributed to these groups, as opposed to the risk of dying from the conditions in these groups. For example, the proportion of deaths attributable to group II causes increases from high- to low-mortality countries (or to an older age structure of the population); however, the risk of death from group II conditions is higher in high-mortality countries.

Implementing a Burden of Disease Study

Knowing how to conduct a burden of disease analysis is important for all countries. Generic steps for a national burden of disease study include the following:

- Assess demographic information, including a census with age, sex, geographic (urban/rural), and selected socioeconomic status information, and vital statistics with births and deaths.
- Collect cause-of-death information for all deaths in a year by age, sex, geographic location, and socioeconomic status as possible, according to the ICD system.
- Define disability by cause/disease, and develop a severity scale using expert and community input.
- Collate information by disease from all sources and assess these data's reliability and validity, using expert opinion when needed to define variables for a spreadsheet.
- Decide whether social value preferences such as age weighting, discounting, economic and social productivity, and expectation of life will be used and what their values will be.
- Estimate healthy life lost for each disease condition and by disease groups.
- Perform sensitivity analyses to check the robustness of results relative to critical variables and assumptions.
- Consider other variations, including assessment of losses by risk factors; regional, age, and sex breakdowns; and future projections.
- Review the policy implications for overall mortality and morbidity in the country and by cause; feed data into cost-effectiveness analysis and further research.
- Include other modifications as appropriate to the country setting.

To use summary measures to assist in health planning and resource allocation decisions, additional steps include the following:

- Estimate the effectiveness (gains of healthy life) of each intervention under consideration in terms of expected coverage and reductions in incidence and/or case fatality or case disability ratios.
- Work out the costs of the proposed interventions.
- Develop cost-effectiveness ratios to plan which combination of interventions targeted to which groups will provide a maximum return of healthy life per expenditure for the funds allocated to health.
- Review expected gains of healthy life according to age, sex, geographic area, and socioeconomic and vulnerable groups to ensure that all are better off (or at least none is worse off) and adjust as necessary.

Another very important consideration in this process is time. The conduct and analysis of such studies must be timely to assure its appropriate use by policy makers and useful for resource allocation decisions. The precision and comprehensive nature of the study must be balanced by the need for timely results.

The steps described previously may be carried out simultaneously or in some sequence, depending on the specific national situation. Modifications will likely be needed depending on the availability of data (**EXHIBIT 1-5**). An actual study requires careful planning on the part of those responsible for its conduct and may include many additional steps that are beyond the scope of this chapter. Even so, these generic steps summarize the essentials of applying the burden of disease methods to a country. Increasingly countries are obtaining, refining, and using these data on an ongoing fashion.

EXHIBIT 1-5 The Burden of Disease in United Arab Emirates, 2010

The United Arab Emirates (UAE) is one of the seven Gulf Cooperation Council (GCC) countries and consists of seven emirates: Abu Dhabi, Dubai, Sharjah, Ajman, Uum Al-Quwain, Fujairah, and Ras Al-Khaimah. An academic study was undertaken to estimate the burden of disease in UAE in 2010 and to develop a set of 2010 population estimates as well as a mortality and disability profile by age group and gender. This analysis was conducted for the UAE national and migrant populations. A burden of disease approach was used to calculate standard expected years of life lost (YLL). Data from the Global Burden of Disease 2010 study and vital registration data from the UAE's Ministry of Health were utilized for the study. Forty-five disease and injury conditions were included in the analysis. Disability estimates were calculated in terms of years of life lived with disability (YLD) using health data from the Emirate of Abu Dhabi; 150 conditions were included in the final analysis.

In 2010, the crude death rate per 1,000 population was 2.47 for UAE nationals and 1.00 for migrants. For all residents of the UAE, regardless of nationality, the leading causes of death were ischemic heart disease (14.5%), other circulatory system diseases (12.4%), motor vehicle crashes (11.3%), malignant neoplasm (6.5%), and other respiratory tract infections (5.5%). For UAE nationals, the crude death rate was estimated at 2.84 for males and 2.08 for females; for migrants, the crude death rate was estimated at 0.99 for males and 1.01 for females. Age-specific mortality rates (ASMRs) for most age groups were lower in migrant males compared to males who were UAE nationals, except for the age groups of 10–14 and 25–29 years.

(continues)

EXHIBIT 1-5 The Burden of Disease in United Arab Emirates, 2010 *(continued)*

The rates were also lower in migrant females compared to nationals, except for migrant females in the age groups of 30–39, 45–49, and 55–64 years.

A total of 291,834 YLL was lost in 2010 due to premature mortality; nationals accounted for 27% of this total and migrants for 73%. YLL lost for migrants (males: 48,514; females: 28,986) were greater than that for nationals (males: 169,864; females: 44,470), reflecting the population imbalance in the UAE wherein migrants vastly outnumber migrants. The leading causes of premature mortality for migrants were motor vehicle crashes, ischemic heart disease, and other circulatory system diseases. Drowning, falls, fires, and unintentional poisoning were some of the important causes of injuries among migrant children between 5 and 14 years of age. Occupational injuries were also an important cause of premature deaths among migrant males between 15 and 44 years.

Overall, the estimated YLD lost was 373,051; nationals accounted for 37% of YLD losses, and the losses were distributed equally among males and females. The top five conditions causing disability among nationals were low back and neck pain (26%), other musculoskeletal diseases (8.3%), injuries (7%), and diabetes (5.3%). Noncommunicable diseases and injuries were responsible for the top 20 causes of YLDs among both nationals and migrants.

FIGURE 1-5 shows the main causes of death in the UAE, and **TABLE 1-6** reviews the top conditions responsible for disease burden in the UAE.

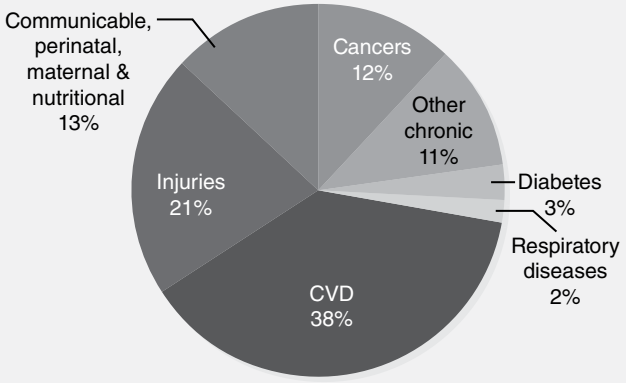


FIGURE 1-5 Deaths in the United Arab Emirates by cause, 2008.

Reprinted from World Health Organization, Noncommunicable Diseases Country Profiles 2011, page 196, Copyright 2011.

TABLE 1-6 Burden of Disease, United Arab Emirates, 2010

	Premature Mortality Only	Disability (YLD) Only	Disability-Adjusted Life Years (DALY)
Rank	Disease	Disease	Disease
1	Motor vehicle crashes	Low back and neck pain	Injuries
2	Ischemic heart disease	Other musculoskeletal disorders	Diseases of the musculoskeletal
3	Other circulatory disease	Injury requiring urgent care	Diseases of the circulatory system
4	Congenital anomalies	Fracture of sternum, rib, or bone	Diseases of the respiratory system
5	Perinatal conditions	Diabetes	Endocrine, nutritional, and metabolic diseases, and immunity disorders
6	Other respiratory tract diseases	Chronic obstructive pulmonary disease	Neoplasm
7	Other malignant neoplasm	Open wounds, superficial injuries	Congenital anomalies
8	Cerebrovascular disease	Fracture of patella, tibia, fibula	Perinatal conditions
9	Suicide	Other neurologic disorders	Diseases of the nervous system
10	Diabetes mellitus	Eczema	Diseases of the digestive system

Data from Allen, K. (2010). A national burden of disease study for the United Arab Emirates (UAE): Quantifying mortality and morbidity differentials between nationals and migrants, 2010.

Data from Allen, K. (2010). A national burden of disease study for the United Arab Emirates (UAE): Quantifying mortality and morbidity differentials between nationals and migrants, 2010.

Comparisons and Trends in Disease Burden

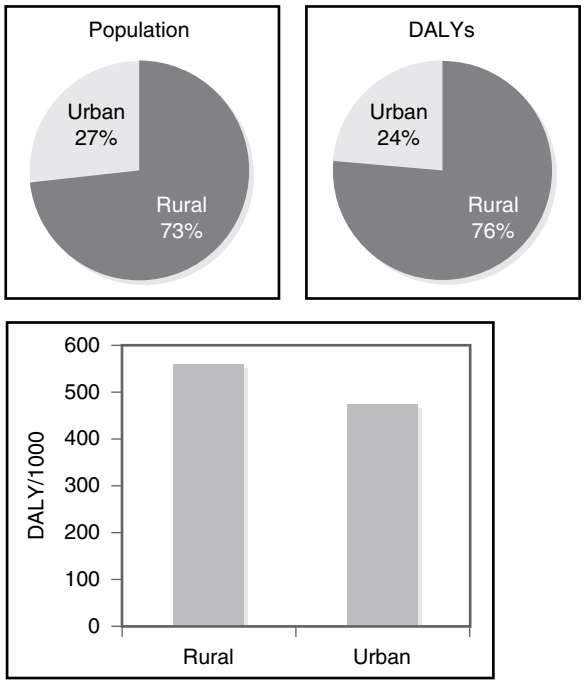
This section reviews a number of country-based and historical burden of disease studies to compare and assess trends in disease burden from place to place and over time.

National Disease Burden Assessments

Comparing the burden of disease across populations, time, and place is an important aspect of national burden of disease studies. This subsection uses examples from burden of disease studies over the past two decades to illustrate how disaggregated data can help in understanding the distribution of ill health in a country.

The Andhra Pradesh Burden of Disease Study, 2001

The regional distribution (urban/rural, state, district) of the disease burden is important to explore in a national burden of disease study. Andhra Pradesh, a state in India, was the focus of one of the most meticulous burden of disease studies, conducted between 1994 and 2001. It had a population of 76 million in 2001, 27% of whom lived in urban areas (20.8 million people); a 1:3 ratio of urban-to-rural disease burden in terms of DALYs lost was identified (Mahapatra, 2001). The burden of disease rates were 19% higher in rural areas than in urban areas, as measured by DALYs lost per 1,000 population (**FIGURE 1-6**).



Note: Total DALYs lost in Andhra Pradesh = 5 million.

FIGURE 1-6 Burden of disease in Andhra Pradesh, 2001, by region.

Data from Mahapatra, P. (2001). *Estimating national burden of disease: The burden of disease in Andhra Pradesh 1990s.* Hyderabad: Institute of Health Systems.

The Burden of Disease and Injury in New Zealand, 2013

According to the latest burden of disease estimates from New Zealand, there has been a reduction in DALYs by 1.2% per year after adjusting for population changes and age structures. The main burden of disease is due to losses from noncommunicable diseases (88%), followed by injuries (8%), and communicable diseases (4%). Approximately 52% of the DALYs are due to nonfatal outcomes. Aging of the population and longer life expectancy is an area of focus for New Zealand to develop and implement preventive and management programs. Coronary heart disease in males and back pain among females are main causes of health loss (Tobias, 2016) (**FIGURE 1-7**).

Burden of Disease Estimates for South Africa, 2000

HIV/AIDS is ravaging Africa; thus, the impact of HIV/AIDS on the burden of disease in African countries can be significant. In South Africa, 30% of the 15 million DALYs lost in 2000 were attributed to HIV/AIDS (**FIGURE 1-8**) (Bradshaw et al., 2003); for a population of 45 million, this means 0.33 DALY lost per capita. Such data are important for national decision making.

The Burden of Disease and Injury in Australia, 2003

In 2003, 2.63 million DALYs lost were attributed to the disease and injury burden in Australia, with 62% of this burden caused by noncommunicable diseases such as cancers and cardiovascular diseases and mental and neurologic disorders. Approximately 49% of the disease burden was caused by deaths from various diseases and injuries, while 51% was due to nonfatal disease and injury outcomes. The top three causes of fatal outcomes were cancers (32.0%), cardiovascular disease (29.0%), and injuries (11.0%); the leading causes of nonfatal outcomes were mental disorders (24%) and neurologic and sense disorders (19%) (Begg et al., 2007) (**FIGURE 1-9**).

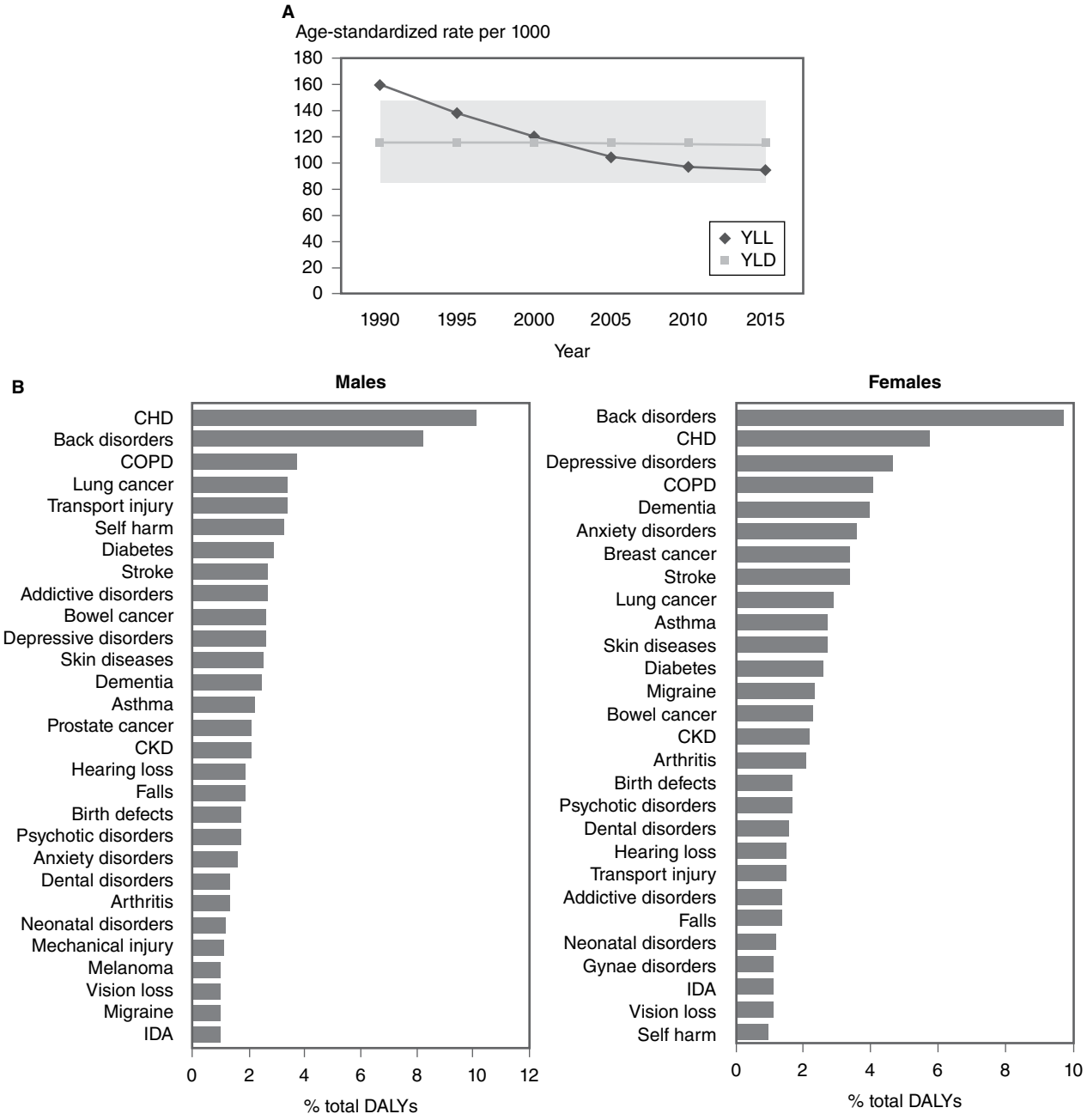
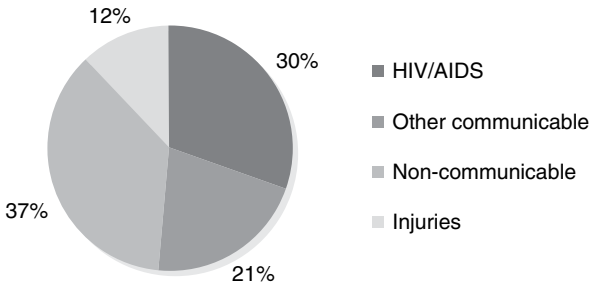


FIGURE 1-7 (a) All-cause YLL and YLD age-standardized rates per 1,000, whole population,1990–2013, in New Zealand. (b) Contribution of leading major specific conditions to health loss (% total DALYs), by gender, 2013, in New Zealand.

(a) New Zealand Ministry of Health. (2016). *Health loss in New Zealand, 1990–2013*. (b) New Zealand Ministry of Health. (2016). *Health loss in New Zealand, 1990–2013*.



Note: Total DALYs lost in South Africa for 2000 = 15 million.

FIGURE 1-8 Burden of disease in South Africa, 2000, by disease groups.

Data from Burden of Disease Research Unit. (2003). *Initial burden of disease estimates for South Africa, 2000*. South Africa: South African Medical Research Council.

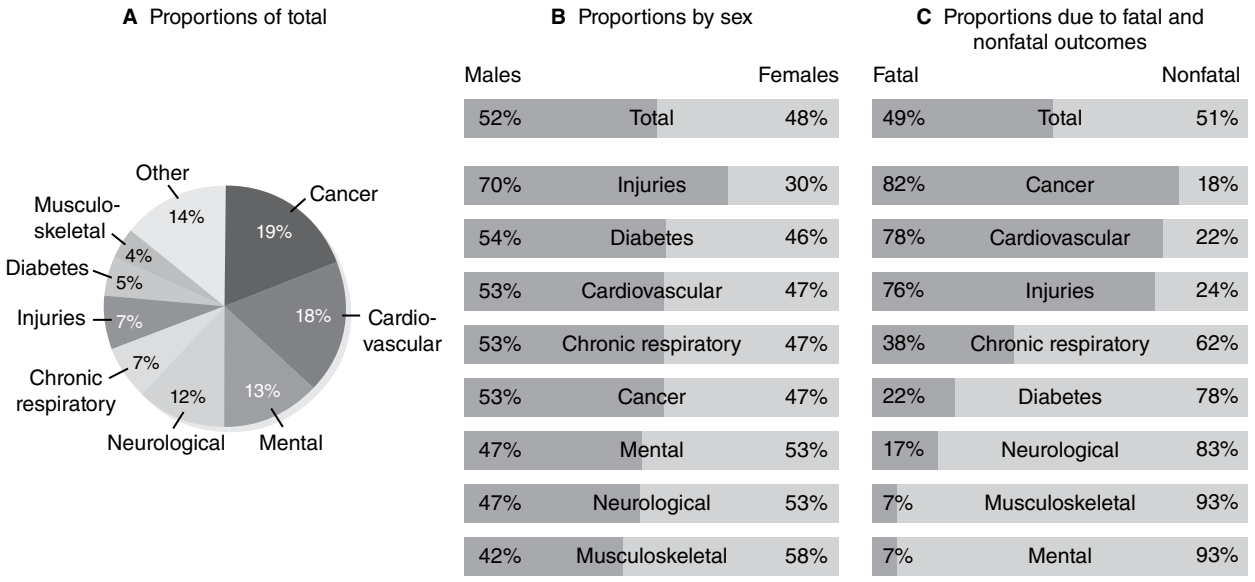


FIGURE 1-9 Burden (DALYs) by broad cause group expressed as (a) proportions of total, (b) proportions by sex, and (c) proportions due to fatal and nonfatal outcomes, Australia, 2003.

The Burden of Disease Attributable to 11 Risk Factors in Hubei Province, China, 2016

Health losses caused by 11 risk factors—namely, smoking, secondhand smoke, alcohol use, diet low in vegetables, diet low in fruits, diet high in sodium, low physical activity, high body mass index, high fasting plasma glucose, high total cholesterol, and high blood pressure—were estimated using multiple data sources in China. They included data related to population demographics, mortality, DALYs, and distribution of exposure to risk factors. Approximately 53% of all deaths were caused by these 11 risk factors, with the mortality rate being 330 per 100,000 people. The DALY rate for the 11 risk factors was 9624 per 100,000 people and accounted for 36% of all DALYs. The leading risk factor for death and DALYs was high blood pressure, which accounted for 15% of all deaths and 9% of total DALYs (**FIGURE 1-10**).

WHO Estimates of Global Burden of Disease

WHO has categorized its member states by income levels into high-, middle-, and low-income nations. The population of the world in 2015 totaled slightly more than 7.3 billion people, with 84% residing in LMICs (**FIGURE 1-11**). As may be expected, approximately 88% of the global burden is found in LMICs, reflecting the double challenge faced by the majority of people in the world: They are relatively poor and they are unhealthy. This relationship between ill health and poverty has long been recognized as complex and has been the object of much research and inquiry.

Burden of Disease by Disease Groups

Another way to disaggregate data is to explore the disease burden based on three disease groups: group I (communicable, infectious, maternal, and perinatal), group II (noncommunicable, chronic), and group III (injuries and violence). There is great variation in the portions allocated to these groups; for example, group I conditions may be responsible for anywhere from 12% to 70% of the burden of disease. When the countries are stratified by gross national income (GNI) per capita as a measure of development, an important trend can be seen from historical data (**TABLE 1-7**): As income rises, the proportion of the burden attributable to group I conditions decreases, while the share attributable to group II conditions increases. This effect is progressive, although countries such Turkmenistan (a middle-income country) still retain a high group I burden. This finding is consistent with the theory of epidemiologic transition, which predicts a change in a country's disease profile with economic development.

Global Assessments of Disease Burden

Information regarding health and disease for all countries of the world can be collated to provide a picture of global health status. In addition, global health assessments may be completed as a separate activity, and these data can then be disaggregated into regional

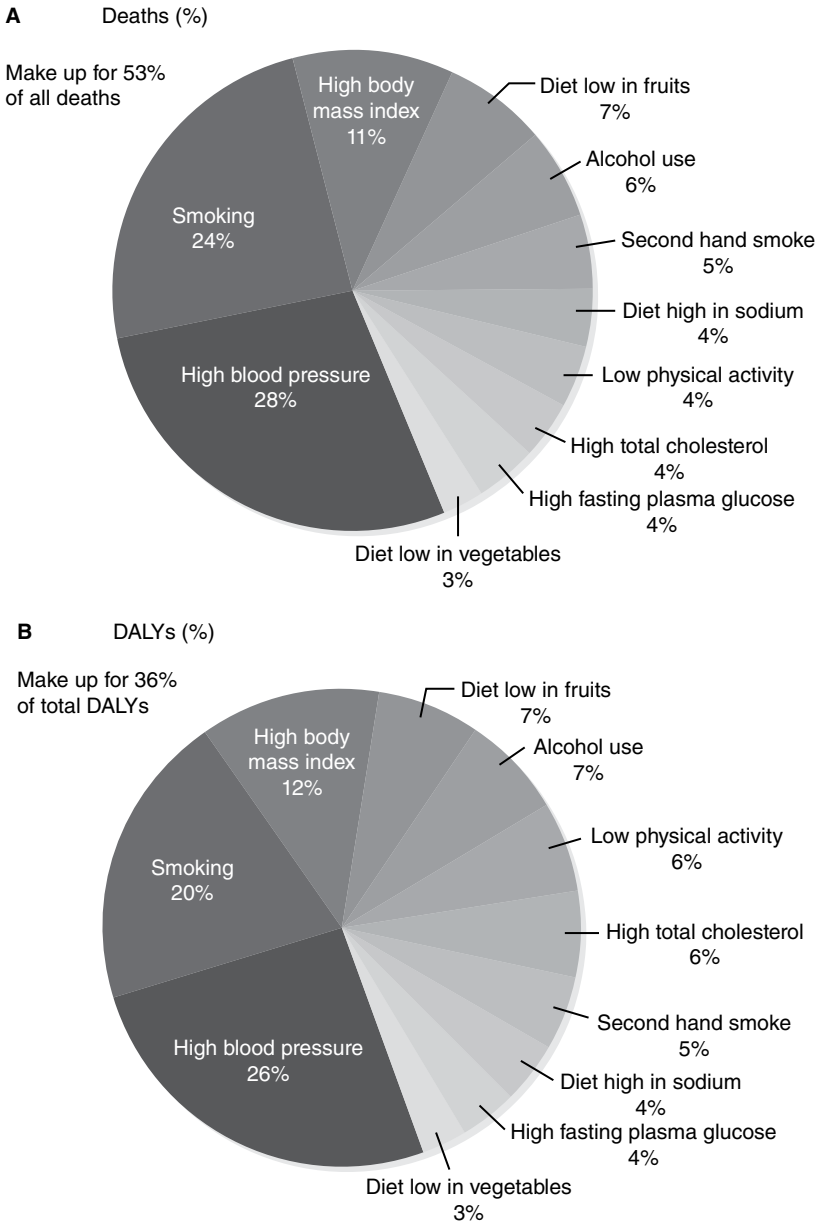


FIGURE 1-10 Burden of disease attributed to 11 risk factors in Hubei Province, China, 2013.

Data from Cui F, Zhang L, Yu C, Hu S, Zhang Y. (2016). Estimation of the disease burden attributable to 11 risk factors in Hubei Province, China: A comparative risk assessment. *International Journal of Environmental Research and Public Health*, *13*(10), 944.

information. Global assessments serve to highlight major challenges facing the world community, and trends in such assessments indicate progress, if any, in improving the health of people worldwide. Such information is critical to the work of organizations such as WHO and UNICEF in their efforts to combat ill health and disease worldwide. This section highlights results of some of these global exercises for assessment of the disease burden and projections for the future.

The Global Burden of Disease Study

The 2015 Global Burden of Disease study constructed estimates of mortality, disability, and DALYs by cause for regions of the world. Demographic estimates of deaths in 2015 by age and sex form the basis of this work. It provides all-cause and cause-specific estimates for 195 countries for 249 causes from the year 1990 to 2015. The results were based on a variety of sources, including vital registrations systems, special studies, surveys, and expert opinion. This section summarizes some of the 2015 GDB data.

Mortality. Globally, in 2015, an estimated 55.8 million deaths occurred, 55% of whom were males. Approximately 20% of global deaths in 2015 were due to group I causes (communicable, maternal, neonatal, and

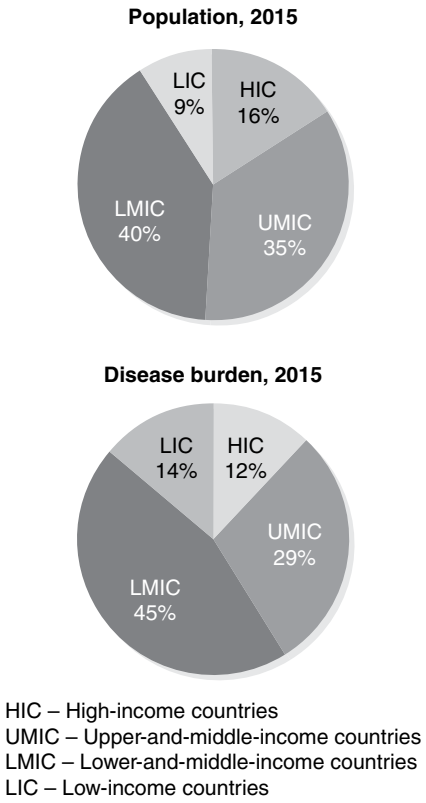


FIGURE 1-11 Global burden of disease, 2015, by income level of countries.

Data from World Health Organization. (2015). *Global burden of disease estimates*. http://www.who.int/healthinfo/global_burden_disease/en/

nutritional diseases), 71% were due to noncommunicable diseases, and 9% were due to injuries. Ischemic heart disease, cerebrovascular disease, and chronic obstructive pulmonary disease were the top three causes of death. Road traffic injuries, self-harm, falls, interpersonal violence, and drowning were the top five causes of injury-related deaths in 2015. Approximately 5.8 million deaths occurred in children younger than 5 years of age globally, with more than 80% of these fatalities being caused by group I causes. Of those under-5 deaths, 79% occurred in the African and Southeast Asia regions. An inordinate share of the mortality burden at the beginning of this century was found in LMICs, even among adults.

TABLE 1-8 shows the differences in the 10 leading causes of deaths in 2015 for high- and low-income countries. The presence of perinatal conditions, tuberculosis, HIV/AIDS, and malaria in the low-income world is indicative of the high impact of these conditions on premature mortality. These conditions are absent from the top 10 causes in the high-income countries, reflecting the success in combating these infectious conditions in the modern era. It is important to note that noncommunicable diseases such as cerebrovascular diseases and ischemic heart disease had already become prominent causes of premature deaths in the low-income world in 2015.

TABLE 1-7 Historical Distribution of Disease Burden Within Countries

Disease Burden in Disease Categories (of 100%)			
Country	**Group I**	**Group II**	**Group III**
Low-Income Nations (GNI per capita: $1,045 or less)*			
Benin	60	30	10
Guinea	65	26	9
Haiti	40	45	15
Mozambique	64	27	9
Lower Middle-Income Nations (GNI per capita: $1,046–$4,125)			
Bangladesh	32	59	9
Cambodia	33	54	13
Egypt	21	72	7
Uzbekistan	26	64	10
Upper Middle-Income Nations (GNI per capita: $4,126–$12,736)			
Colombia	13	67	20
Mauritius	9	84	7
Mexico	13	75	13
Turkmenistan	24	65	11

Note: Disease classification system: Group I: Communicable, infectious, maternal, and perinatal; Group II: Noncommunicable and chronic; Group III: Injuries and accidents.
* Gross national income (GNI) per capita calculated using the World Bank Atlas method.
Data from WHO and World Bank, 2015.

Disability. The GBD study 2015 update also provides an evaluation of the contributions made by specific conditions to disability in the world. Leading causes of disability in 2015 worldwide are shown in

TABLE 1-8 Leading Causes of Deaths in High-Income and Low-Income Countries, 2015

High-Income Countries		Low-Income Countries	
Rank	**Cause**	**Rank**	**Cause**
1	Ischemic heart disease	1	Lower respiratory infections
2	Cerebrovascular disease	2	HIV/AIDS
3	Alzheimer's and other dementias	3	Diarrheal diseases
4	Lung cancer	4	Malaria
5	Lower respiratory tract infection	5	Cerebrovascular disease
6	Chronic obstructive pulmonary disease	6	Ischemic heart disease
7	Colorectal cancers	7	Tuberculosis
8	Chronic kidney disease	8	Congenital defects
9	Diabetes	9	Neonatal preterm birth complications
10	Breast cancer	10	Neonatal encephalopathy

Data from GBD, 2015 (Based on the World Bank income categories).

TABLE 1-9. Neuropsychiatric and behavioral conditions, musculoskeletal disorders, and diabetes dominate the causes of disability, accounting for 5 of the top 10 conditions. However, a diverse spectrum of conditions, such as iron-deficiency anemia, skin diseases, and oral disorders, also appear on the list. A unique contribution of the GBD work has been its placement of nonfatal health outcomes in the center of international health policy in recent years. The important, and yet often ignored, effects of these conditions are obvious once disability is counted in these estimates of disease burden.

Disease Burden. Based on the estimation of deaths and disability presented in the preceding subsection, the global disease burden for 2015 was estimated using DALYs. Leading causes of the global burden in 2015 (**TABLE 1-10**) indicate how those conditions affect the global disease burden. The top 10 list is a mixture of communicable and perinatal conditions, noncommunicable diseases, and road traffic injuries. This situation highlights the challenge facing the global health community as it simultaneously continues to fight infectious diseases, seeks to improve the response to chronic conditions, and prepares to meet the increasing impact of injuries.

Age and Disease Distributions. **FIGURE 1-12** illustrates the distribution of the global burden in 2015 by disease groups and demonstrates the growing relative impact of chronic diseases (group II) over infectious conditions (group I). **FIGURE 1-13** provides comparable figures for loss of healthy life in seven major regions of the world. Note that communicable diseases still represent a considerable portion of the disease burden in 2015, especially in sub-Saharan Africa.

As the figures demonstrate, various subregions within middle- and low-income countries are at different stages of the epidemiologic transition. The influx of chronic diseases has added another layer of problems in some areas, while the burden of communicable diseases has not yet been eradicated. This "double burden" poses a major challenge for the health systems in these nations. In addition, the scarcity of resources in many of these countries makes the situation even more critical, and it becomes imperative to define interventions that are cost-effective and able to reduce the burden of disease.

Other Ways Burden Can Be Measured

Mortality and morbidity alone have been used for decades for international comparisons of disease

TABLE 1-9 Leading Causes of Disability Losses Globally, 2015

Rank	Cause
1	Lower back and neck pain
2	Sense organ diseases
3	Depressive disorders
4	Iron-deficiency anemia
5	Skin diseases
6	Diabetes
7	Migraine
8	Other musculoskeletal disorders
9	Anxiety disorders
10	Oral disorders

Note: Disability losses are defined by years of life lived with disability (YLD). Ranking is based on YLD rates for all age groups, both sex and level 3 cause (http://ihmeuw.org/44lj).
Source: Institute for Health Metrics and Evaluation, 2015.
Data from GBD, 2015.

TABLE 1-10 Leading Cause of Global Burden of Disease, 2015

Rank	Cause
1	Ischemic heart disease
2	Cerebrovascular disease
3	Lower respiratory infections
4	Low back and neck pain
5	Neonatal preterm birth complications
6	Diarrheal diseases
7	Sense organ disease
8	Neonatal encephalopathy
9	Road traffic injuries
10	HIV/AIDS

Data from Institute for Health Metrics and Evaluation, 2015. Ranking based on DALYs rates for all age groups, both sex for 2015 and level 3 cause. Retrieved from http://ihmeuw.org/44lk.

burden. Mortality among children younger than 5 years is considered a sensitive indicator of the overall health of nations, but especially of the health of women and children. UNICEF publishes an annual *State of the World's Children* report (UNICEF, 2016) that includes a ranking of nations based on this indicator (**TABLE 1-11**).

Gross national income (GNI) per capita is an indicator of national wealth, and the relationship between these variables usually follows an expected sequence, such that the country with the lowest GNI per capita has the worst indicators of health. However, as Table 1-11 indicates, even countries that have relatively higher per capita income can have poor indicators of health service accessibility (e.g., proportion of newborns protected against tetanus). For example, the per capita GNI for Mongolia is higher than that for Nicaragua, yet both the countries rank 84th for under-5 mortality, with lower life expectancy in Mongolia compared to Nicaragua. Such examples demonstrate that the relationship between health and poverty is complex and needs in-depth investigation. When seeking to improve the health of nations, both absolute poverty and the disparities within societies serve as impediments to empowerment of the poor and needy, especially women and children.

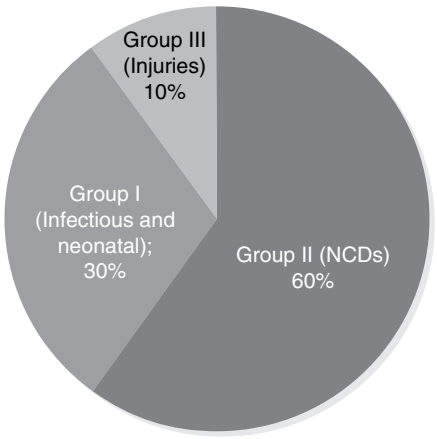


FIGURE 1-12 Global burden of disease 2015 by disease groups.
Data from GBD, 2015.

Projections

Forecasts of disease burden have been attempted with the intent of providing some basis for health planning. Making such projections is a challenging task

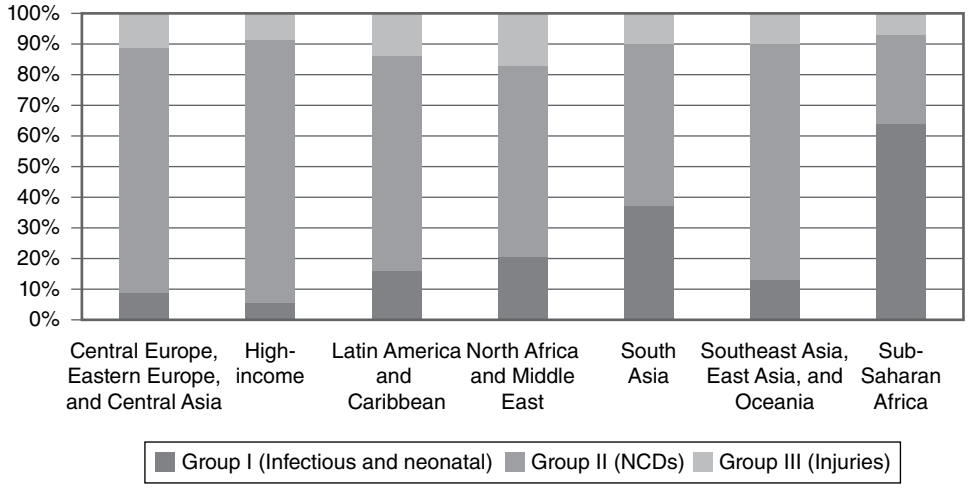


FIGURE 1-13 Proportion of disease burden by disease groups in selected regions, 2015.

Data from GBD, 2015.

TABLE 1-11 Health Status Indicators and National Income for Selected Low- and Middle-Income Countries

Country	Ranking by Child Mortality (<5 years)	Life Expectation at Birth, 2015 (years)	Stunted Children <5 Years (%)	Newborns Protected Against Tetanus *(%)	Gross National Income per Capita, 2014 (U.S. dollars)
Afghanistan	16	61	41	70	680
Angola	1	53	29	78	4,126 to 12,735
Bhutan	67	70	34	83	2,370
Guatemala	71	72	47	85	3,430
Mongolia	84	70	11	—	4,280
Nicaragua	84	75	23	81	1,870
Niger	10	62	43	81	410
Pakistan	22	66	45	75	1,400
Peru	99	75	15	85	6,360
Sierra Leone	5	51	38	85	700

* Modeling is done based on pregnant women who were given two doses or more of tetanus toxoid vaccine to calculate the percentage of births that can be considered as protected against tetanus.
Data from UNICEF, 2016.

that requires further data manipulations and the use of assumptions. These assumptions must predict changes in disease prevalence and incidence over time, the effects of interventions, and other factors. As a result, all projections are estimates with substantial variations that are highly dependent on the data used to derive them.

The GBD 2015 study was conducted to provide burden of disease estimates for more than 200 health conditions. These estimates are available based on gender, age groups, countries, and regions, allowing for comparisons across different years. Estimates are available since 1990, and the current

leading causes of disease burden are presented in Table 1-10.

The dominance of chronic diseases on this list is obvious, although respiratory conditions still appear to be important. Injuries from road traffic crashes are the ninth leading cause of the global disease burden and are expected to rise in prominence.

The growing importance of noncommunicable diseases is a global phenomenon, and these conditions' impact on low- and middle-income countries and regions needs to be assessed. However, there is a persistent burden of respiratory infections and diarrheal diseases in these regions. The situation in the low- and middle-income world is one where the "triple burden" of persistent communicable diseases, prevalent noncommunicable conditions, and increasing injuries will call for an appropriate response from public health officials.

Burden of Disease Attributed to Risk Factors

An analysis of risk factors that underlie many important disease conditions can help inform policy decisions concerning interventions directed toward health promotion and disease reduction. For example, smoking, alcohol, hypertension, and malnutrition are risk factors for a variety of diseases, and specific interventions have been developed that may reduce their prevalence. Risk factors include an array of human behaviors, nutritional deficiencies and excesses, substance abuse, and certain characteristics such as hypertension. Some factors are both an outcome and a risk factor (hypertension), some are challenging to measure (violence), and yet others (smoking and alcohol) lead to many possible disease outcomes. The linkage between an identified risk factor and the set of associated disease outcomes may be difficult to directly quantify, and the portion of specific disease prevalence attributable to any one factor may be problematic.

Relationships such as those shown in **FIGURE 1-14** require careful assessment to determine the proportion of heart disease to be attributed to hypertension in relation to other interacting causal factors. The best way to determine the portion of disease that may be ascribed to hypertension is through randomized trials with careful assessment of disease outcomes over time: Results from studies that control hypertension have shown a reduction of death and disability from not only cardiac disease, but also from cerebrovascular and renal diseases.

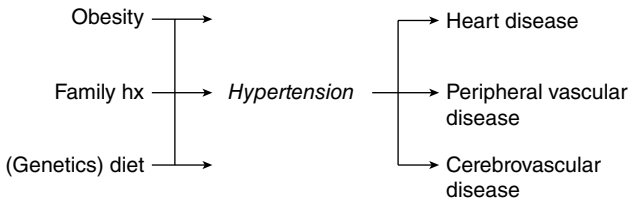


FIGURE 1-14 Flowchart of a "disease" and "risk factor" – Hypertension.

Because the most important purpose of risk factor analysis is to assist in decision making about the allocation of resources, the link between the various risk factors and potential interventions to reduce those risks should be clear. The effectiveness of interventions against risk factors ultimately should be judged by their ability to reduce the amount of healthy life lost attributed to the diseases that the risk factor affects. For the evaluation of an intervention that reduces hypertension, for example, the healthy life losses from the entire range of diseases that hypertension influences are required.

Although understanding the underlying factors that lead to disease and the complex interrelations in the web of causation has long been a major focus of epidemiology (Rothman, 2012), most analyses of the relationships of risk factors to specific diseases have been done in the context of individual risk factors in limited settings and with wide variations in the criteria for risk assessment. As a consequence, comparisons of risk factors as determinants of disease on a population health level are problematic.

The Burden of Selected Major Risk Factors

The model used in recent GBD studies for causal attribution of health outcomes is based on *counterfactual analysis* that focuses on risk factor compared to a counterfactual risk distribution (Ezzati et al., 2002). Within this analysis, the contribution of one risk factor, or a group of risk factors, to disease or mortality was estimated by comparing the current or future disease burden with the levels that would be expected under an alternative hypothetical scenario. This involves an evaluation of the effect that a risk factor has on a disease and its consequences, by setting the risk factor to its minimum while keeping all other factors constant. This method has the advantage of showing the potential gains by risk reduction from all levels of suboptimal exposure in a consistent way across risk factors (Ezzati et al., 2002).

The Comparative Risk Assessment (CRA) part of GBD 2015, for example, carried out a systematic evaluation of 79 risk factors divided into three main groups (referred to as "level 1")—behavioral, environmental

and occupational, and metabolic risks—relative to global and regional burdens of disease using a specific model for analysis (WHO, 2002a). Recently, the analysis was conducted from 1990 to 2015 for 17 level 2 risk factors (Global Burden of Disease Risk Factors Collaborators, 2016). As shown in part (a) of **FIGURE 1-15**, the five leading risks for mortality on a global level are dietary risks (19%), high systolic blood pressure (17%), tobacco smoke (11%), air pollution (10%), and high fasting plasma glucose (8%); together these account for 65% of global mortality. These five factors especially increase risks for heart disease, diabetes, and cancer and have major consequences for countries across all income groups. Part (b) of Figure 1-15

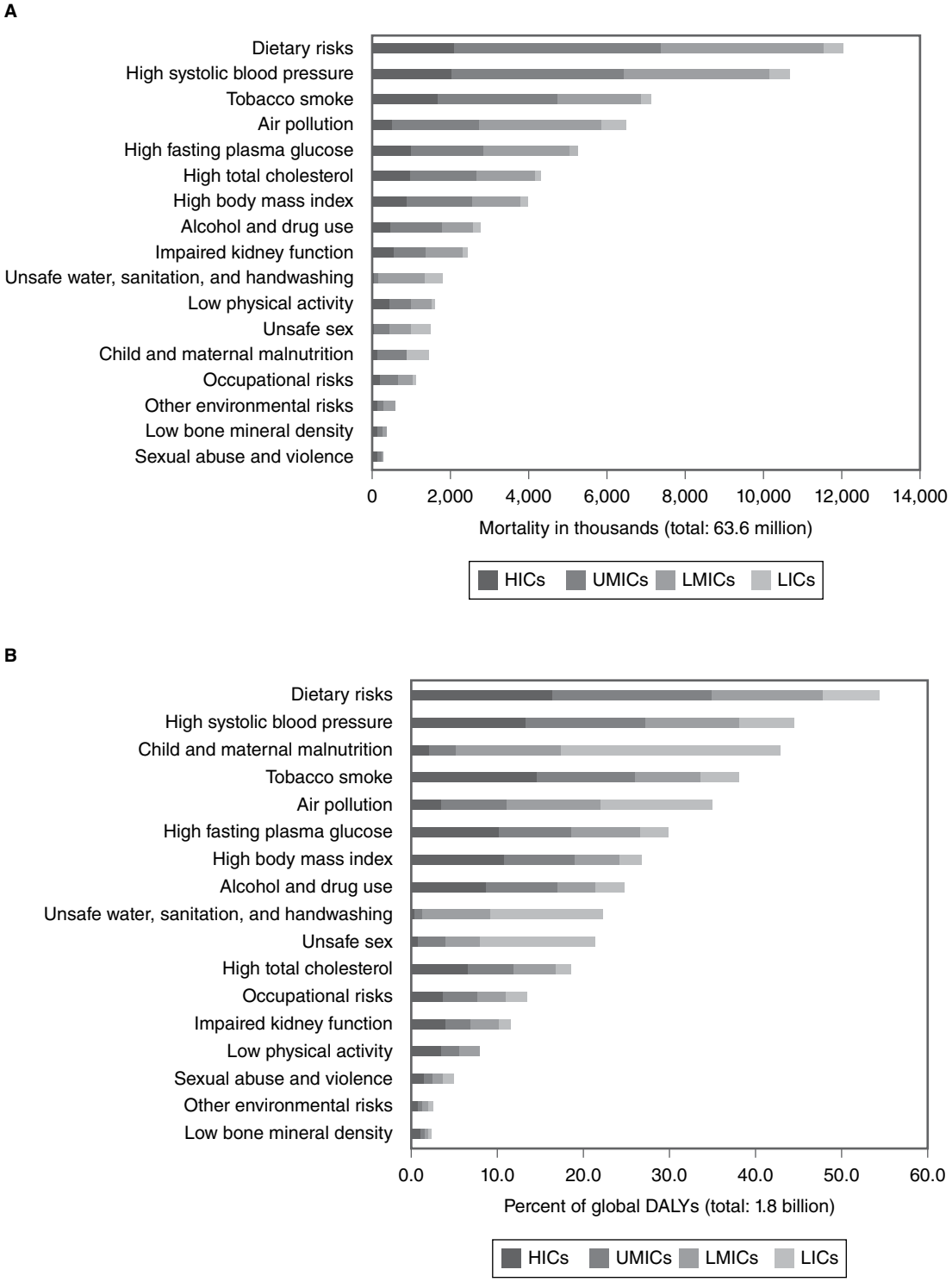


FIGURE 1-15 Mortality (a) and DALYs (b) due to 19 leading risk factors by country income level, 2015.

Data from GBD, 2015.

shows the main risks for burden of disease (DALYs) on a global level, with dietary risks (15%), high systolic blood pressure (12%), child and maternal malnutrition (9%), tobacco smoke (9%), and air pollution (9%) accounting for more than half of the global DALYs. These estimates show the contribution of behavioral and environmental risk factors toward development of group II diseases.

The recent reductions in environmental risk factors mostly stem from reductions in unsafe water, unsafe sanitation, poor hygiene, and household air pollution. In terms of behavioral risk factors, smoking is among five top risk factors contributing to DALYs at the regional level. Alcohol and drug abuse is largely a problem for men, contributing to 7% of burden of disease in 2015, but those mainly affected vary greatly by geographic region: This factor has its greatest impact on men in Southeast Asia, South Asia, and Central and Eastern European countries (Institute for Health Metrics and Evaluation, 2015).

Eight level 2 risk factors—dietary risks, high blood pressure, high total cholesterol, air pollution, tobacco, high fasting plasma glucose, high body mass index, and low physical activity—account for more than 90% of ischemic heart disease (the leading cause of death worldwide). Although these major risk factors are associated with high-income countries, more than 88% of the total global burden of disease that they cause occurs in LMICs (Institute for Health Metrics and Evaluation, 2015). Some risk factors may have few effects on the total global burden of disease, yet be very important locally within certain populations and regions. For example, iodine deficiency affects certain LMICs and results in substantial disability in those populations.

The demographic and epidemiologic transitions discussed earlier in this chapter affect shifts in these major risk factors by changing social, economic, and political trends and their complex interactions. Low-income countries continue to struggle against the high burdens of infectious diseases, malnutrition (including undernutrition and micronutrient deficiencies), and maternal and child health problems; at the same time, they must deal with the additional burdens of high levels of noncommunicable disease and injuries. WHO has estimated that if such types of risks did not exist, global life expectancy would be 10 years longer (Mathers, Stevens, & Mascarenhas, 2009).

Conclusion

The health of populations is the fundamental concern of global public health. The first step in the pursuit of population health improvement is the measurement of health and disease. Measurement is required to establish the magnitude of disease problems, define causal factors, explore potential solutions, and determine the impact of interventions. Measuring the impact of diseases on populations in terms of mortality and morbidity and their consequences is essential for planning effective ways to reduce the burden of illness and for setting priorities.

The burden of disease in populations has been gauged in many ways. Examples include measures of mortality, such as infant mortality rates; demographic measures, such as expectation of life at birth; and measures of morbidity, such as days away from work. However, for purposes of comparison among populations and for assisting in health planning and resource allocation, a common denominator combining these factors is needed. Summary measures of population health based on the amount of healthy life lost from disability and from death have been developed to serve that purpose. Composite indicators (such as HeaLYs and DALYs) use duration of time (years, weeks, days) to measure the loss of healthy life from disease and the gain from interventions. These metrics are evolving into important tools for assisting health-related decision making. Nevertheless, to avoid misuse of such indicators, it is critical for those using them to understand the underlying assumptions and limitations and to meet the rather formidable data requirements. These summary measures also could be used to examine the burden of disease among subpopulations defined by sociocultural economic attributes and especially on vulnerable groups. Thus, they could be used for ensuring that health-related decisions consider equity as well as cost-effective criteria.

Trends in disease burden provide important clues about the success of ongoing health programs and the need for development of new interventions. At the same time, they reflect non-health-related factors that are important to the production or maintenance of health in populations. Inter-country and inter-regional comparisons allow for measuring progress among nations; they can highlight inequalities in health status and examine these disparities in relation to social, economic, educational, and other factors.

Health systems across the world are greatly affected by ongoing changes in disease profiles and population dynamics. These systems must develop the capacity to respond to such changes effectively within the resources of each nation. Decisions must be based on evidence about the patterns of diseases, their risk factors, and the effectiveness of alternative interventions. Timely collection and analysis of appropriate, high-quality data to support such evidence are

a prerequisite for improving equitable global health development.

Acknowledgments

The authors would like to express their gratitude to Nukhba Zia for the editorial support provided in the preparation of this chapter. The authors would also like to acknowledge the contributions of the late Dr. Richard H. Morrow, MD, MPH, who made significant contributions to previous iterations of this chapter.

Discussion Questions

1. How can data help achieve the main purpose of a health system in any country? Give examples.
2. What are the essential elements of a burden of disease assessment, and which types of data are the most challenging to obtain in a low-income country?
3. What are the relative strengths and weaknesses of summary measures such as HeaLYs and DALYs compared to more traditional indicators of disease burden such as infant or maternal mortality?
4. In your country or city, what would be the most appropriate set of indicators to assess the impact of chronic diseases on the population? Why?

Appendix 1

Some of the Health-Related SDG Indicators

Target 2.1: By 2030, end hunger and ensure access by all people, in particular the poor and people in vulnerable situations, including infants, to safe, nutritious, and sufficient food all year round

- 2.1.1 Prevalence of undernourishment

Target 2.2: By 2030, end all forms of malnutrition, including achieving, by 2025, the internationally agreed targets on stunting and wasting in children under 5 years of age, and address the nutritional needs of adolescent girls, pregnant and lactating women, and older persons

- 2.2.1 Prevalence of stunting
- 2.2.2 Prevalence of malnutrition

Target 3.1: By 2030, reduce the global maternal mortality ratio to less than 70 per 100,000 live births

- 3.1.1 Maternal mortality ratio
- 3.1.2 Proportion of births attended by skilled health personnel

Target 3.2: By 2030, end preventable deaths of newborns and children under 5 years of age, with all countries aiming to reduce neonatal mortality to at least as low as 12 per 1,000 live births and under-5 mortality to at least as low as 25 per 1,000 live births

- 3.2.1 Under-5 mortality rate
- 3.2.2 Neonatal mortality rate

Target 3.3: By 2030, end the epidemics of AIDS, tuberculosis, malaria, and neglected tropical diseases and combat hepatitis, water-borne diseases, and other communicable diseases

- 3.3.1 Number of new HIV infections per 1,000 uninfected population, by sex, age, and key populations
- 3.3.2 Tuberculosis incidence per 1,000 population
- 3.3.3 Malaria incidence per 1,000 population
- 3.3.4 Hepatitis B incidence per 100,000 population
- 3.3.5 Number of people requiring interventions against neglected tropical diseases

Target 3.4: By 2030, reduce by one third premature mortality from noncommunicable diseases through prevention and treatment and promote mental health and well-being

- 3.4.1 Mortality rate attributed to cardiovascular disease, cancer, diabetes, or chronic respiratory disease
- 3.4.2 Suicide mortality rate

Target 3.5: Strengthen the prevention and treatment of substance abuse, including narcotic drug abuse and harmful use of alcohol

- 3.5.1 Coverage of treatment interventions (pharmacological, psychosocial, and rehabilitation and aftercare services) for substance use disorders
- 3.5.2 Harmful use of alcohol, defined according to the national context as alcohol per capita consumption (aged 15 years and older) within a calendar year in liters of pure alcohol

Target 3.6: By 2020, halve the number of global deaths and injuries from road traffic accidents

- 3.6.1 Death rate due to road traffic injuries

Target 3.7: By 2030, ensure universal access to sexual and reproductive healthcare services, including for family planning, information and education, and the integration of reproductive health into national strategies and programs

- 3.7.1 Proportion of women of reproductive age (aged 15–49 years) who have their need for family planning satisfied with modern methods
- 3.7.2 Adolescent birth rate (aged 10–14 years; aged 15–19 years) per 1,000 women in that age group

Target 3.8 Achieve universal health coverage, including financial risk protection, access to quality essential healthcare services, and access to safe, effective, quality, and affordable essential medicines and vaccines for all

- 3.8.1 Coverage of essential health services
- 3.8.2 Number of people covered by health insurance or a public health system per 1,000 population

Target 3.9: By 2030, substantially reduce the number of deaths and illnesses from hazardous chemicals and air, water, and soil pollution and contamination

- 3.9.1 Mortality rate attributed to household and ambient air pollution
- 3.9.2 Mortality rate attributed to unsafe water, unsafe sanitation and lack of hygiene (exposure to unsafe Water, Sanitation and Hygiene for All [WASH] services)
- 3.9.3 Mortality rate attributed to unintentional poisoning

Target 4.2: By 2030, ensure that all girls and boys have access to quality early childhood development, care, and pre-primary education so that they are ready for primary education

- 4.2.1 Proportion of children under 5 years of age who are developmentally on track in health, learning, and psychosocial well-being, by sex

Target 5.2: Eliminate all forms of violence against all women and girls in the public and private spheres, including trafficking and sexual and other types of exploitation

- 5.2.1 Proportion of ever-partnered women and girls aged 15 years and older subjected to physical, sexual, or psychological violence by a current or former intimate partner in the previous 12 months, by form of violence and by age
- 5.2.2 Proportion of women and girls aged 15 years and older subjected to sexual violence by persons other than an intimate partner in the previous 12 months, by age and place of occurrence

Target 5.6: Ensure universal access to sexual and reproductive health and reproductive rights

- 5.6.2 Number of countries with laws and regulations that guarantee women aged 15–49 years access to sexual and reproductive health care, information, and education

Target 6.1: By 2030, achieve universal and equitable access to safe and affordable drinking water for all

- 6.1.1 Proportion of population using safely managed drinking water services

Target 6.2: By 2030, achieve access to adequate and equitable sanitation and hygiene for all and end open defecation, paying special attention to the needs of women and girls and those in vulnerable situations

- 6.2.1 Proportion of population using safely managed sanitation services, including a hand-washing facility with soap and water

Target 8.8 Protect labor rights and promote safe and secure working environments for all workers, including migrant workers, in particular women migrants, and those in precarious employment

- 8.8.1 Frequency rates of fatal and nonfatal occupational injuries, by sex and migrant status

Target 11.5: By 2030, significantly reduce the number of deaths and the number of people affected and substantially decrease the direct economic losses relative to global gross domestic product caused by disasters, including water-related disasters, with a focus on protecting the poor and people in vulnerable situations

- 11.5.1 Number of deaths, missing persons, and persons affected by disaster per 100,000 people

Target 11.7 By 2030, provide universal access to safe, inclusive, and accessible green and public spaces, in particular for women and children, older persons, and persons with disabilities

- 11.7.2 Proportion of persons who are victims of physical or sexual harassment, by sex, age, disability status, and place of occurrence, in the previous 12 months

Target 16.1: Significantly reduce all forms of violence and related death rates everywhere

- 16.1.1 Number of victims of intentional homicide per 100,000 population, by sex and age
- 16.1.2 Conflict-related deaths per 100,000 population, by sex, age, and cause
- 16.1.3 Proportion of population subjected to physical, psychological, or sexual violence in the previous 12 months
- 16.1.4 Proportion of population that feel safe walking alone around the area they live

Target 16.2 End abuse, exploitation, trafficking, and all forms of violence against and torture of children

- 16.2.1 Proportion of children aged 1–17 years who experienced any physical punishment and/or psychological aggression by caregivers in the past month
- 16.2.3 Proportion of young women and men aged 18–29 years who experienced sexual violence by age 18

Reproduced from United Nations. Sustainable Development Goals. SDG Indicators. Retrieved from: https://unstats.un.org/sdgs/metadata/

References

Allen, K. (2010). A national burden of disease study for the United Arab Emirates (UAE): Quantifying mortality and morbidity differentials between nationals and migrants, 2010. Unpublished doctoral thesis. Baltimore, MD: Johns Hopkins Bloomberg School of Public Health.

Anand, S., & Hanson, K. (1997). Disability-adjusted life years: A critical review. *Journal of Health Economics, 16*(6), 685–702.

Anker, M., Black, R. E., Coldham, C., Kalter, H. D., Quigley, M. A., Ross, D., & Snow, R. W. (1999). A standard verbal autopsy method for investigating causes of death in infants and children. WHO/CDS/CSR/ISR/99.4. Geneva, Switzerland: World Health Organization.

Barendregt, J. J., Bonneux, L., & Van der Maas, P. J. (1996). DALYs: The age-weights on balance. *Bulletin of the World Health Organization, 74*(4), 439–443.

Barker, C., & Green, A. (1996). Opening the debate on DALYs (disability-adjusted life years). *Health Policy and Planning, 11*(2), 179–183.

Barnum, H. (1987). Evaluating healthy days of life gained from health projects. *Social Science and Medicine, 24*(10), 833–841.

Begg, S., Vos, T., Barker, B., Stevenson, C., Stanley, L., & Lopez, L. D. (2007). *The burden of disease and injury in Australia 2003.* Canberra, Australia: Australian Institute of Health and Welfare.

Blumstein, J. F. (1997). The Oregon experiment: The role of cost–benefit analysis in the allocation of Medicaid funds. *Social Science and Medicine, 45*(4), 545–554.

Bobadilla, J. L. (1998). *Searching for essential health services in low and middle income countries: A review of recent studies on health priorities.* Washington, DC: Inter American Development Bank.

Bradshaw, D., Groenewald, P., Laubscher, R., Nannan, N., Nojilana, B., Norman, R., . . . Johnson, L. (2003). Initial burden of disease estimates for South Africa, 2000. *South African Medical Journal*, *93*(9), 682–688.

Busschbach, J. J. V., Hesing, D. J., & de Charro, F. T. (1993). The utility of health at different stages of life: A qualitative approach. *Social Science and Medicine, 37*(2), 153–158.

Coale, A. J., Demeny, P., & Vaughan, B. (1983). *Regional model life tables and stable populations. Studies in population*. New York, NY/London, UK: Academic Press.

Coale, A. J., & Guo, G. (1989). Revised regional model life tables at very low levels of mortality. *Population Index, 55*(4), 613–643.

Crooper, M., & Kopits, E. (2003). *Traffic fatalities and economic growth. World Bank Policy Research Working Paper 3035*. Washington, DC: World Bank.

Eddy, D. M. (1991). Oregon's methods. Did cost-effectiveness analysis fail? *Journal of the American Medical Association, 266*(15), 2135–2141.

Evans, J. R., Hall, K. L., & Warford, J. (1981). Shattuck Lecture: Health care in the developing world: problems of scarcity and choice. *New England Journal of Medicine, 305*(19), 1117–1127. doi: 10.1056/NEJM198111053051904

Ezzati, M., Lopez, A. D., Rodgers, A., Vander Hoorn, S., Murray, C. J., & Comparative Risk Assessment Collaborating Group. (2002). Selected major risk factors and global and regional burden of disease. *Lancet, 360*(9343), 1347–1360. doi: 10.1016/S0140-6736(02)11403-6

Figgis, P., & Walters, C. (2015–2017). Depleting resources adding pressure to healthcare. Retrieved from http://www.pwc.com/gx/en/industries/healthcare/emerging-trends-pwc-healthcare/depleting-resources.html

GBD 2015 Mortality and Causes of Death Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease study 2015. *Lancet, 388*(10053), 1459–1544. doi: 10.1016/S0140-6736(16)31012-1

Ghana Health Assessment Team. (1981). A quantitative method of assessing the health impact of different diseases in less developed countries. *International Journal of Epidemiology, 10*(1), 73–80.

Global burden of disease and risk factors. (2006). Washington, DC/New York, NY: International Bank for Reconstruction and Development, World Bank, & Oxford University Press.

Global Burden of Disease Risk Factors Collaborators. (2016). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: A systematic analysis for the Global Burden of Disease study 2015. *Lancet, 388*(10053), 1659–1724. doi: 10.1016/S0140-6736(16)31679-8

Global Burden of Disease Risk Factors Collaborators, Forouzanfar, M. H., Alexander, L., Anderson, H. R., Bachman, V. F., Biryukov, S., . . . Murray, C. J. (2015). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: A systematic analysis for the Global Burden of Disease study 2013. *Lancet, 386*(10010), 2287–2323. doi: 10.1016/S0140-6736(15)00128-2

Gold, M. R. (1996). *Cost-effectiveness in health and medicine*. New York, NY: Oxford University Press.

Hill, K. (2001). *Demographic techniques: Indirect estimation*. Oxford, UK: Elsevier Science.

Hill, K. (2003). *Adult mortality in the developing world: What we know and how we know it.* Paper presented at the Training Workshop on HIV/AIDS and Adult Mortality in Developing Countries, New York, NY.

Hyder, A. A., & Morrow, R. H. (1999). Steady state assumptions in DALYs: Effect on estimates of HIV impact. *Journal of Epidemiology and Community Health, 53*(1), 43–45.

Hyder, A. A., Rotllant, G., & Morrow, R. H. (1998). Measuring the burden of disease: healthy life-years. *American Journal of Public Health, 88*(2), 196–202.

Institute for Health Metrics and Evaluation. (2015). GBD Compare. Retrieved from http://www.healthdata.org/data-visualization/gbd-compare

Jamison, D. T., Breman, J. G., Measham, A. R., Alleyne, G., Claeson, M., Evans, D. B., . . . Musgrove, P. (Eds). (2006). *Disease control priorities in developing countries*. Washington, DC/New York, NY: International Bank for Reconstruction and Development, World Bank, & Oxford University Press.

Kaplan, R. M., & Anderson, J. P. (1996). The general health policy model: An integrated approach. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (2nd ed., Chapter 32, pp. 309–322). Philadelphia, PA: Lippincott-Raven Publishers.

Last, J. M., Spasoff, R. A., & Harris, S. S. (Eds). (2000). *A dictionary of epidemiology* (4th ed). New York, NY: Oxford University Press.

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., . . . Memish, Z. A. (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease study 2010. *Lancet, 380*(9859), 2224–2260. doi: 10.1016/S0140-6736(12)61766-8

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., . . . Memish, Z. A. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease study 2010. *Lancet, 380*(9859), 2095–2128. doi: 10.1016/S0140-6736(12)61728-0

Mahapatra, P. (2001). *Estimating national burden of disease: The burden of disease in Andhra Pradesh 1990s*. Hyderabad: Institute of Health Systems.

Mathers, C., Stevens, G., & Mascarenhas, M. (2009). *Global health risks: Mortality and burden of disease attributable to selected major risks.* Geneva, Switzerland: World Health Organization.

Mathers, C., Vos, T., Lopez, A., Salomon, J., Lozano, R., & Ezzati, M. (Eds). (2001). *National burden of disease studies: A practical guide* (2nd ed.). Geneva, Switzerland: World Health Organization.

Ministry of Health (2016). *Health loss in New Zealand 1990–2013: A report from the New Zealand burden of diseases, injuries and risk factors study.* Wellington, NZ: Ministry of Health.

Morrow, R. H., & Bryant, J. H. (1995). Health policy approaches to measuring and valuing human life: conceptual and ethical issues. *American Journal of Public Health, 85*(10), 1356–1360.

Murray, C. J. L., & Chen, L. C. (1992). Understanding morbidity change. *Population & Development Review, third edition, 18*.

Murray, C., Lopez, A., Ahmad, O. B., Inoue, M., Fergusson, B., Salomon, J. A., . . . Hill, K. H. (2002). *World mortality in 2000: Life tables for 191 countries.* Geneva, Switzerland: World Health Organization.

Murray, C. J. L., & Lopez, A. D. (Eds). (1994). *Global comparative assessments in the health sector: Disease burden, expenditures and intervention packages.* Geneva, Switzerland: World Health Organization.

Murray, C. J. L., Salomon, J., Mathers, C., & Lopez, A. (2002). *Summary measures of population health: Concepts, ethics, measurement and applications*. Geneva, Switzerland: World Health Organization.

Nord, E. (1992). Methods for quality adjustment of life years. *Social Science and Medicine, 34*(5), 559–569.

Nord, E. (1993). Unjustified use of the Quality of Well-Being Scale in priority setting in Oregon. *Health Policy, 24*(1), 45–53.

Omran, A. (1971). The epidemiological transition: A theory of the epidemiology of population change. *Milbank Memorial Fund Quarterly, 49*, 509–538.

Petiti, D. B. (1994). *Meta-analysis, decision analysis and cost-effectiveness analysis: Methods for quantitative synthesis in medicine, second edition*. New York, NY: Oxford University Press.

Robine, J. M. (1994). *Disability free life expectancy trends in France: 1981–1991: International comparisons.* Paper presented at the Advances in Health Expectancies conference, Canberra, Australia.

Rothman, K. J. (2012). *Epidemiology: An introduction* (2nd ed.). Oxford, UK: Oxford University Press.

Salomon, J. A., Haagsma, J. A., Davis, A., de Noordhout, C. M., Polinder, S., Havelaar, A. H., . . . Vos, T. (2015). Disability weights for the Global Burden of Disease 2013 study. *Lancet Global Health, 3*(11), e712–e723. doi: 10.1016/S2214-109X(15)00069-8

Salomon, J. A., Vos, T., Hogan, D. R., Gagnon, M., Naghavi, M., Mokdad, A., . . . Jonas, J. B. (2012). Common values in assessing health outcomes from disease and injury: Disability weights measurement study for the Global Burden of Disease study 2010. *Lancet, 380*(9859), 2129–2143. doi: 10.1016/S0140-6736(12)61680-8

Thatte, N., Kalter, H. D., Baqui, A. H., Williams, E. M., & Darmstadt, G. L. (2009). Ascertaining causes of neonatal deaths using verbal autopsy: Current methods and challenges. *Journal of Perinatology, 29*(3), 187–194. doi: 10.1038/jp.2008.138

Torrance, G. W. (1986). Measurement of health state utilities for economic appraisal. *Journal of Health Economics, 5*(1), 1–30.

UNAIDS. (2017). *UNAIDS data 2017*. Geneva, Switzerland: Joint United Nations Programme on HIV/AIDS.

UNICEF. (2015). *Levels and trends in child mortality.* New York, NY: Author.

UNICEF. (2016). *The state of the world's children 2016: A fair chance for every child.* (2016). New York, NY: Author.

Weinstein, M. C., Siegel, J. E., Gold, M. R., Kamlet, M. S., & Russell, L. B. (1996). Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. *Journal of the American Medical Association, 276*(15), 1253–1258.

Woolf, S. H., Aron, L., Dubay, L., Simon, S. M., Zimmerman, E., & Luk, K. X. (2015). *How are income and wealth linked to health and longevity?* Washington, DC: Urban Institute & Center on Society and Health.

World Bank. (1993). *World development report 1993: Investing in health.* New York, NY: Oxford University Press.

World Bank. (2015). *New country classifications*. Washington, DC: The World Bank Group. Retrieved from https://blogs.worldbank.org/opendata/new-country-classifications

World Health Organization (WHO). (1980). *International classification of impairments, disabilities and handicaps: A manual of classification relating to the consequences of disease.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2000). *The world health report 2000*. Geneva, Switzerland: Author.

World Health Organization (WHO), Commission on Macroeconomics and Health. (2001). *Macroeconomics and health: Investing in health for economic development*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2002a). *The Global Burden of Disease 2000 project: Aims, methods and data sources*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2002b). *World report on violence and health*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2002c). *The international classification of functioning, disability and health: Introduction.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2003). *The world health report 2003: Shaping the future*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2015). *Global status report on road safety.* (2015). Geneva, Switzerland: Author.

World Health Organization (WHO). (2016). International classification of diseases. Retrieved from http://www.who.int/classifications/icd/en/

World Health Organization (WHO). (2017). Global Health Observatory (GHO) data: HIV/AIDS. Retrieved from http://www.who.int/gho/hiv/en/

Zeckhauser, R., & Shephard, D. (1976). Where now for saving lives? *Law and Contemporary Problems, 40*(b), 5–45.

CHAPTER 2

Culture, Behavior, and Health

Susan C. Scrimshaw and Sandra D. Lane

"If you wish to help a community improve its health, you must learn to think like the people of that community. Before asking a group of people to assume new health habits, it is wise to ascertain the existing habits, how these habits are linked to one another, what functions they perform, and what they mean to those who practice them" (Paul, 1955, p. 1).

People around the world have beliefs and behaviors related to health and illness that stem from cultural forces as well as individual experiences and perceptions. A 16-country study of community perceptions of health, illness, and primary health care found that in all 42 communities studied, people used both the Western biomedical system and indigenous practices, including indigenous practitioners. Also, there were discrepancies between which services the governmental agencies said existed in the community and what was really available. Due to positive experiences with alternative healing systems and shortcomings in the Western biomedical system, people relied on both (Nichter, 2008; Scrimshaw, 1992). In recognition of the widespread use of nonbiomedical therapies, the U.S. National Institutes of Health established a center for the study of alternative and complementary medicine, which is now called the National Center for Complementary and Integrative Health (NCCIH, 2014). With a yearly budget of $124.1 million, the NCCIH's mandate is to provide research evidence of efficacy for nonbiomedical treatments, such as acupuncture, herbal preparations, meditation, and spiritual healing. The popularity of complementary and integrative approaches was demonstrated by a U.S. nationwide survey, demonstrating that 33.2% of adults used such treatments (NCCIH, 2017). The global importance of those approaches is evidenced by the World Health Organization's (WHO) major report, *WHO Traditional Medicine Strategy 2014–2023*, and the establishment of 21 regional centers for the study of traditional medicine to strengthen the evidence base and promote quality.

In the past three decades, we have moved from rejecting traditional health beliefs and practices to understanding that culturally rooted, traditional healing approaches are valued and used, often in combination with biomedicine. At this time there is increasing recognition that such approaches can and should be rigorously evaluated. Nevertheless, appreciation for cultural knowledge and behaviors in the planning and execution of health programs often falls short of the mark. Experience has shown that when health programs fail to recognize and work with indigenous beliefs and practices, they also fail to reach their goals. Similarly, research to plan and evaluate health programs must take cultural beliefs and behaviors into account if researchers expect to understand why programs are not working and determine what to do about it.

This chapter draws on the social sciences—particularly, anthropology, psychology, and sociology—to examine the cultural and behavioral parameters that are essential to understanding global health efforts. It complements the *Understanding and Acting on Social Determinants of Health and Health Equity*

chapter, which covers social, political, and economic forces that affect health, but does not go deeply into the cultural components of health. This chapter begins with some key concepts from the field of anthropology and the subfield of medical anthropology. It continues with brief descriptions of the various types of health belief systems and healers around the world. Next, some key theories of health behavior and behavioral and cultural change are described and discussed. Issues of health literacy and health communication are then addressed, along with the myriad health promotion strategies available. Methodological issues are presented, followed by a case study of acquired immunodeficiency syndrome (AIDS) and a commentary on the international efforts to curb the Ebola virus pandemic in West Africa. Another case study examines the use of rapid assessment methods to guide the introduction of an improved nutritional cereal for infants and children in Ghana. The chapter concludes by summarizing how all of these areas need to be considered in global health efforts.

Basic Concepts from Medical Anthropology

Health and illness are defined, labeled, evaluated, and acted upon in the context of culture. In the nineteenth century, anthropologist Edmund Tyler (1871) defined culture as "that complex whole which includes knowledge, belief, art, morals, law, custom, and any other capabilities acquired by man as a member of society." Since those early days of anthropology, there have been literally hundreds of definitions of culture, but most have the following concepts in common (Board on Neuroscience and Behavioral Health, 2002):

- Shared ideas meanings and values
- Socially learned, not genetically transmitted
- Patterns of behavior that are guided by these shared ideas, meanings, and values
- Often exists at an unconscious level
- Constantly modified through "lived experiences"

The last of these concepts—*lived experiences*—comprises the experiences that people (and sometimes groups of people) go through as they live their lives. These experiences modify their culturally influenced beliefs and behaviors (Garro, 2000; Mattingly & Garro, 2000). As a consequence, culture is not static on either the group or the individual level; rather, people are constantly changing. This concept allows for cultural change as people migrate to a new setting (community, region, or country), as people acquire additional education and experiences, and as conditions change around them (e.g., armed conflicts, economic changes in a country or region, political changes). This is a helpful viewpoint when looking at cultural change on both individual and group levels.

Medical anthropologists observe different cultures and their perspectives on disease and illness. For example, they look at the biological and the ecological aspects of disease, the cultural perspectives, and the ways in which cultures approach prevention and treatment.

Insider Versus Outsider Perspectives

To understand the cultural context of health, it is essential to work with several key concepts. First, the concepts of insider and outsider perspectives are useful for examining when we are seeing things from our point of view and when we are trying to understand someone else's view of things. The insider perspective (*emic,* in anthropological terminology) shows the culture as viewed from within. It refers to the meaning that people attach to things from their cultural perspective. For example, the view that worms (*Ascaris*) in children are normal and are caused by eating sweets is a perspective found within some cultures. The outsider perspective (*etic,* in anthropological terminology) refers to the same thing as seen from the outside. Rather than meaning, it conveys a structural approach, or something as seen without understanding its meaning for a given culture. The outsider perspective can also convey an outsider's meaning attached to the same phenomenon. For example, this view might hold that *Ascaris* infection is contracted through eggs in contaminated soil or foods contaminated by contact with that soil; the eggs get into the soil through fecal wastes from infected individuals. The concepts of insider and outsider perspectives allow us to look at health, illness, and prevention and treatment systems from several vantage points; to analyze the differences between these perspectives; and to develop approaches that will work within a cultural context (Scrimshaw & Hurtado, 1987).

To continue the example, in Guatemalan villages where the previously mentioned insider beliefs about *Ascaris* prevailed, researchers learned that some mothers believed that worms are normal and are not a problem unless they become agitated. In their view, worms live in a bag or sac in the stomach and are fine while so confined. Agitated worms get out and appear in the feces or may be coughed up. Mothers also believed that worms are more likely to become agitated during the rainy season, because the thunder and lightning frighten them. From an outsider perspective, this relationship makes sense: Sanitation is

more likely to break down in the rainy season, so there is more chance of infection and more diarrheal disease, which will reveal the worms.

The dilemma for the health workers, in the Guatemalan worms example, was to get the mothers to accept deworming medication for their children, because most of the time worms were perceived as normal. If the health workers tried to tell the mothers that their beliefs were wrong, the mothers would reason that the health workers did not understand illness in a Guatemalan village and would reject their proposal. The compromise was to suggest that the children be dewormed just before the rainy season, so as to avoid the problem of agitated worms. It worked.*

The insider–outsider approach leads to another set of concepts. According to the Western biomedical definition, disease is the outsider perspective—that is, disease is an undesirable deviation from a measurable norm. Deviations in temperature, white blood cell count, red blood cell count, bone density, and many others are, therefore, seen as indicators of disease. Illness, in contrast, means "not feeling well." Thus, it is a subjective, insider view. This sets up some immediate dissonances between the two views. It is possible to have an undesirable deviation from a Western biomedical norm and to feel fine. Hypertension, early stages of cancer, human immunodeficiency virus (HIV) infection, and early stages of diabetes are all instances where people may feel well, yet have a disease. Thus, healthcare providers must communicate the need for behaviors to "fix" something that people may not realize is wrong.

This potential conflict becomes even more critical when we think about the role of risk factors and prevention in contemporary biomedical practice. Laboratory tests can reveal the potential risk of future disease—for example, elevated low-density lipoprotein cholesterol (LDL) level, which may be predictive of an increased possibility of future cardiac or arterial disease. To reduce the possibility of such future disease, the patient might be labeled as "high risk" and prescribed medication to take daily for years. Even among those persons with substantial education, it is difficult to understand how to interpret risk and probability when one feels well and does not have a disease. The widespread rejection of vaccines by some well-educated parents represents another example of the power of insider (*emic*) understandings. The failure of public health advocacy for vaccines to convince such parents of the outsider (*etic*) view of the safety and value of vaccines demonstrates the power of cultural beliefs even in the face of scientific evidence.

It is also possible for someone to feel ill and for the Western biomedical system not to identify a disease. When this occurs, there is a tendency for Western-trained healthcare providers to say that nothing is wrong or that the person has a "psychosomatic" problem. Although both of these statements can be correct, there are several other explanations for this occurrence. One possibility is that Western biomedical science has not yet figured out how to measure a disease or disorder. Recent examples of this phenomenon include chronic fatigue syndrome and fibromyalgia: These conditions were labeled "psychosomatic" at one time, and remain incompletely understood by biomedicine, but now are defined by measurable deviations from a biological norm. Similarly, painful menstruation was labeled "subconscious rejection of femininity" in the past, but is now associated with elevated prostaglandin levels and can be helped by administration of a prostaglandin inhibitor.

A more intriguing set of conditions are what anthropologists have called "culture-bound syndromes" (Hughes, 1990; Simons, 2001; Simons & Hughes, 1985), but that might be better described as "culturally defined syndromes." *Culturally defined syndromes* are an insider way of describing and attributing a set of symptoms. They often refer to symptoms of a mental or psychological problem, but a physiological disease may also exist, posing a challenge to the health practitioner. For example, Rubel, O'Nell, and Collado-Ardon (1984) found that an illness called *susto* ("fright") in Mexico corresponded with symptoms of tuberculosis in adults. If people were told there was no such thing as *susto* and that they had tuberculosis, they rejected the diagnosis and the treatment on the grounds that the doctors obviously knew nothing about *susto*. This situation was complicated by the fact that tuberculosis was viewed as serious and stigmatizing. The solution was to discuss the symptoms with people and mention that Western biomedicine has a treatment for those symptoms (Rubel et al., 1984). *Susto* may also be used to describe other sets of symptoms—for example, those of diarrheal disease in children (Scrimshaw & Hurtado, 1988). Other examples of culture-bound syndromes include evil eye (Latin America, the Mediterranean), zar (the Middle East and North Africa), brain fag or brain fog (West Africa), amok (running amok) or *mata elap* (Indonesia, Malaysia, and the Philippines), *latah* (Malaysia and Indonesia), *p'a leng* (China), and *ataque de nervios* (Puerto Rico) (Guarnaccia et al., 2010; Simons & Hughes, 1985).

* We are indebted to Elena Hurtado of Guatemala for this example.

A recent example of what appears to be a new culture-bound syndrome has emerged among refugee children in Sweden whose families faced deportation (Aviv, 2017). *Uppgivenhetssyndrom*, or resignation syndrome, afflicting hundreds of children on the eve of their families' deportation, involves loss of speech and voluntary movement. The children exhibit no neurologic pathology in clinical tests, yet they must be fed, washed, and turned in bed by caregivers, as if they are comatose.

Not all individuals within a group will necessarily have the same beliefs and behaviors. With culturally defined syndromes, it is essential for an outsider to ask about the symptoms associated with the illness and to proceed with diagnosis and treatment on the basis of those symptoms. This is good practice in any event, because people often make a distinction between the cause of a disease or illness and its symptoms. Even if the perceived cause is inconsistent with the Western biomedical system, a disease can be diagnosed and treated based on the symptoms without challenging people's beliefs about the cause. When people's beliefs about the cause are denied, they may reject prevention or treatment measures entirely (Nichter, 2008).

The term *Western biomedicine* is used throughout this chapter because a term like *modern medicine* would deny the fact that there are other medical systems, such as Chinese and Ayurvedic medicine, that have modern forms. *Indigenous medical system* is used to refer to an insider—"within the culture"—system. Thus, Western biomedicine is an indigenous medical system in some countries, such as the United States and Canada, but it may exist side by side with other indigenous systems, even in the United States and Western Europe. In most of the world, Western biomedicine now coexists with, and often dominates, local or indigenous systems. Because of this multiplicity of systems, and because of class differences, physicians and policy makers in a country may not accept or even be aware of the extent to which indigenous systems exist or recognize their importance (Cameron, 2010). Also, many countries are home to peoples with multiple cultures and multiple languages. The cross-cultural principles discussed in this chapter may be just as important to work within a country as it is to work in multiple countries or cultures.

Ethnocentrism

Another key concept from medical anthropology is that of ethnocentrism. *Ethnocentric* refers to seeing your own culture as "best." Ethnocentrism is a natural tendency, because the survival and perpetuation of a culture depend on its teaching its children to accept the culture and on its members feeling that it is a good thing. In the context of cross-cultural understanding, ethnocentrism poses a barrier if people approach a culture with the attitude that it is inferior to their own culture. One of anthropology's key contributions to how we view the world is *cultural relativism*, which refers to the idea that the beliefs, behaviors, and values of each culture make sense within that culture. From this perspective, each culture has developed its own ways of solving the problems of how to live together; how to obtain the essentials of life, such as food and shelter; how to explain phenomena; and so on. Cultural relativism has been misunderstood in contemporary discourse as implying that we should not hold beliefs about "right" and "wrong," especially about culturally patterned practices that may cause harm. For example, in parts of Africa and the Middle East, female infants or young girls may have their genitals partially or completely amputated, which is sometimes called "female genital mutilation" by Western activists seeking to end the practice. Lane and Rubinstein (1996), in their analysis of the practice, advocated that cultural relativism means that we should try to understand the cultural explanations driving the practice and the context in which it occurs. At the same time, they emphasize that understanding why and how this behavior occurs does *not* mean that we condone the practice. Indeed, familiarity with the cultural rationale of a potentially harmful practice can be the basis of effective intervention against it.

The importance of seeking to understand a cultural practice can be a major challenge when global health is considered. What if a behavior is "wrong" from an epidemiologic perspective? How does one distinguish between a "dangerous" behavior (e.g., using an HIV-contaminated needle, swimming in a river with snails known to carry schistosomiasis, ingesting a powder with lead in it as part of a healing ritual) and behaviors that are merely different and, therefore, seem odd? For example, Bolivian peasants traditionally used very fine clay in a drink believed to be good for digestion and stomach ailments. Health workers succeeded in discouraging this practice in some communities because "eating dirt" seemed like a bad thing. The health workers then found themselves faced with increased caries and other symptoms of calcium deficiency in these same communities. Analysis revealed that the clay was a key source of calcium for these communities. It turns out that biomedicine also uses clay—but we color it pink or give it a mint flavor and put it in a bottle with a fancy label (S. Scrimshaw, personal observation).

Thus, there is a delicate balance between being judgmental without good reason and seeking to introduce behavior change because there is real harm from existing behaviors. In general, it is best to leave harmless practices alone and focus on understanding and changing harmful behaviors. This task is more difficult than it might seem, because the concept of cultural relativism also applies to perceptions of quality of life. A culture in which people believe in reincarnation may approach death with more equanimity, and may not embrace drastic procedures that prolong life only briefly. In some cultures, loss of a body organ is viewed as impeding the ability to go to an afterlife or the next life, and such surgery may be refused. Thus, it is important in global health for cultural outsiders to be cautious about making statements about what is good for someone else.

Holism

The concept of holism is also useful in looking at health and disease cross-culturally. *Holism* is an approach used by anthropologists that looks at the broad context of whatever phenomenon is being studied. Holism involves staying alert for unexpected influences, because you never know what may have a bearing on the program you are trying to implement. For public health, this consideration is crucial because diverse factors may influence health and health behavior (Nichter, 2008).

A classic example of this situation is the detective work that went into discovering the etiology of the New Guinea degenerative nerve disease, kuru. Epidemiologists could not figure out how people contracted the disease, which appeared to have a long incubation period and to occur more frequently in women and children than in men. Many hypotheses were advanced, including inheritance (genetic), infection (bacterial, parasitic), and psychosomatic explanations.

By the early 1960s, the most widely accepted of the prevailing hypotheses was that kuru was genetically transmitted. Nevertheless, this proposal did not explain the sex differences in infection rates in adults but not in children, nor how such a lethal gene could persist. Working with Gadjusek of the National Institutes of Health (NIH), cultural anthropologists Glasse and Lindenbaum used in-depth ethnographic interviews to establish that kuru was relatively new to that region of New Guinea, as was the practice of cannibalism. Women and children were more likely to engage in the ritual consumption of the brains of dead relatives as a way of paying tribute to them, which was culturally less acceptable for men. Also, this tissue was cooked, but women, who did the cooking, and children, who were around during cooking, were more likely to eat it when it was partially cooked and, therefore, still infectious. Lindenbaum and Glasse suggested the disease was transmitted by cannibalism. To confirm their hypothesis, Gadjusek's team inoculated chimpanzees with brain material from women who had died of kuru; the animals subsequently developed the disease. The disease, initially thought to be a slow virus, was subsequently identified as caused by prions, which are mis-folded proteins transmitted through the ingestion of brain tissue. Since then, the practice of cannibalism has declined and the disease has now virtually disappeared (Gadjusek, Gibbs, & Alpers, 1967; Lindenbaum, 1971). The research on kuru led scientists to discover a similar pattern of disease caused by the ingestion of beef that had been fed neurologic tissue of deceased cattle, called "mad cow disease" or variant Creutzfeldt-Jakob disease (Collinge et al., 2006).

Health Literacy

In recent years, increasing attention has focused on another area that intersects with culture in people's ability to understand and access health care—the concept of health literacy. *Health literacy* is defined as "the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions" (Ratzan & Parker, 2000). Health literacy has been most thoroughly explored in the United States, and until recently was seen more as a literacy issue than a cultural issue. A 2004 Institute of Medicine report notes the importance of considering cultural issues such as many of those discussed in this chapter, and of taking a more global look at the problem and needed interventions (Nielsen-Bohlman, Panzer, & Kindig, 2004). Many National Academy of Medicine (NAM) reports have followed, along with an NAM Roundtable and discussion papers on the topic (e.g., Allen, Auld, Logan, Montes, & Rosen, 2017).

In looking at culture and health literacy, several categories for misunderstandings between provider and patient emerge.

First, there is a difference between medical terminology and lay terminology, which can occur in any language or culture. What is "diastolic" or a "bronchodilator"? What are HDL and LDL? What are T cells?

Second, individual and cultural differences surround concepts. What does it mean to maintain a "moderate" weight? To an anxious teen who wants to become a model, moderate weight might mean

something clinically dangerously low (from the outsider, health practitioner perspective). To some women from Latin America or the Middle East, moderate weight will be heavier than U.S. norms, whereas a U.S. woman who fears she weighs too much might be viewed as dangerously thin in those cultures.

Third, meanings may differ. While working with prenatal care programs in Mexico, Scrimshaw's team struggled with communicating the concept of risk in pregnancy as they developed materials to help women identify symptoms that meant that they should seek care. It turned out that the direct translation of "risk" into Spanish, or *riesgo*, did not carry the same meaning. When they explained the concept to women, they said, "Oh, you mean *peligro*." *Peligro* translates directly as "danger" (Alcalay, Ghee, & Scrimshaw, 1993; Scrimshaw et al., 1990).

Finally, language issues may affect understanding. While researchers were investigating seizure disorders in adolescents from three cultures, it became clear that the word "trauma" has two different meanings. It can mean psychological shock, or it can mean physical trauma, such as a blow to the head. The exact same word *trauma* is used in Spanish, with the same two potential meanings. When neurologists talked with patients who had epilepsy and their parents from Latino cultures, the neurologists used the word "trauma" as a cause of seizures to mean a blow to the head. The Latino parents heard the psychological meaning and thought their child had been traumatized psychologically by some fright or shock (Long, Scrimshaw, & Hernandez, 1992).

Lane and her colleagues (2017) found that health literacy can be compromised at times even for well-educated individuals. It is particularly important to note that health literacy is as much a problem of the healthcare provider and health communication staff as it is of a patient or the people in a community. If medical "jargon" is used, no amount of education short of experience in medicine or nursing will help someone understand. Terms such as "oncology," "nephrology," and "gastroenterology" have meaning for the medical world, but not for patients. Healthcare providers outside the United States often have a better understanding of this issue than their U.S. counterparts.

Cultural Competence

A concept related to health literacy is that of cultural competence. *Cultural competence* in health care describes "the ability of systems to provide care to patients with diverse values, beliefs and behaviors, including tailoring delivery to meet patients' social, cultural, and linguistic needs" (Betancourt, Green, & Carrillo, 2002). The concept of cultural competence emerged, in part, from the U.S. federal government's elaboration of the *Cultural and Linguistically Appropriate Services* (CLAS) guidelines to improve the care of diverse populations (Office of Minority Health, 2016). Healthcare institutions began using the CLAS standards to design cultural competence trainings for their staff. It became clear, however, that in-service training was insufficient to make participants truly competent in cultural issues. From this critique, scholars have suggested that cultural humility is a more realistic framework, because it promotes the understanding and appreciation of health beliefs and behaviors in their cultural contexts and respectful strategies to negotiate optimal health in the context of these beliefs and behaviors (Tervalon & Murray-García, 1998). To achieve this goal, we must understand our own biases.

Cultures vary in their definitions of health and of illness. A condition that is endemic in a population may be seen as normal and may not be defined as illness. *Ascaris* infection in young children was previously mentioned as a perceived "normal" condition in many populations. Similarly, malaria is seen as normal in some parts of Africa, because everyone has it or has had it. In Egypt, where schistosomiasis was common and affected the blood vessels around the bladder, blood in the urine was referred to as "male menstruation" and was seen as normal. These definitions may also vary by age and by gender. In most cultures, symptoms such as fever in children are seen as more serious than the same symptoms in adults. Men may deny symptoms more than women in some cultures, but women may do the same in others. Often, adult denial of symptoms is due to the need to continue working.

Sociologist Talcott Parsons (1948) first discussed the concept of the sick role, wherein an individual must "agree" to be considered ill and to take actions (or allow others to take actions) to define the state of his or her health, discover a remedy, and do what is necessary to become well. Individuals who adopt the sick role neglect their usual duties, may indulge in dependent behaviors, and seek treatment to get well. By adopting the sick role, they are viewed as having "permission" to be exempted from usual obligations, but they are also under an obligation to try to restore health. The process of seeking to remain healthy or to restore health is discussed in more detail later in this chapter.

Belief Systems

EXHIBIT 2-1 depicts types of insider cultural explanations of disease causation. Based on the literature, it attempts to be as comprehensive as possible for cultures around the world. The exhibit consists of generalizations about culture-specific health beliefs and behaviors; these generalizations cannot, however, be assumed to apply

EXHIBIT 2-1 Types of Insider Cultural Explanations of Disease Causation

Body Balances

- Temperature: Hot, cold
- Energy
- Blood: Loss of blood; properties of blood reflect imbalance; pollution from menstrual blood
- Dislocation: Fallen fontanel
- Organs: Swollen stomach; heart; uterus; liver; umbilicus; others
- Incompatibility of horoscopes

Emotional

- Fright
- Sorrow
- Envy
- Stress

Weather

- Winds
- Change of weather
- Seasonal disbalance

Vectors or Organisms

- Worms
- Flies
- Parasites
- Germs

Supernatural

- Bewitching
- Demons
- Spirit possession
- Evil eye
- Offending God or gods
- Soul loss

Food

- Properties: Hot, cold, heavy (rich), light
- Spoiled foods
- Dirty foods
- Sweets
- Raw foods
- Combining the "wrong" foods (incompatible foods)
- Mud/clay

Sexual

- Sex with forbidden person
- Overindulgence in sex

Heredity

Old Age

to every individual from a given culture. We can learn about the hot/cold balance system of Latinos, Asians, and Middle Easterners, explained in the next section, but the details of the system will vary from country to country, from village to village, and from individual to individual. When someone walks in the door of a clinic, you cannot know whether he or she as an individual adheres to the beliefs described for his or her culture and what shape the individual's belief system takes. This makes the task of the culturally proficient healthcare provider both easier and harder. It means a practitioner working with a Mexican population does not have to memorize which foods are hot and which are cold in Mexico, but the practitioner does need to know that the hot/cold belief system is important in Mexican culture and be able to be understanding and responsive when people bring up the topic.

The beliefs held by cultures around the world are classified into various categories, which are discussed here. These categories are used for diagnosis and treatment and for explaining the etiology or origin of the illness. Often, multiple categories are used. For example, emotions may be seen as causing a "hot" illness.

Body Balances

Within body balances (opposites) belief systems, the concepts of "hot" and "cold" are among the most pervasive around the world. The hot/cold balance is particularly important in Asian, Latin American, and Mediterranean cultures. Hot and cold beliefs are part of what is referred to as "humoral medicine," which is thought to have derived from Greek, Arabic, and East Indian pre-Christian traditions (Foster, 1953; Logan, 1972; Weller, 1983). The concept of opposites (e.g., hot and cold, wet and dry) also may have developed independently in other cultures (Rubel & Haas, 1990). For example, in the Chinese medical tradition, hot is referred to as *yin* and cold as *yang* (Topley, 1976).

In the hot and cold belief system, a healthy body is seen as in balance between the two extremes. Illness may be brought on by violating the balance, such as washing the hair too soon after childbirth (cold may enter the body, which is still "hot" from the birth), eating hot or heavy foods at night, or breastfeeding while upset (the milk will be hot from the emotions and make the baby ill). "Hot" does not always refer to temperature, however. Often foods such as beef and pork

are classified as hot regardless of temperature, whereas fish may be seen as cold regardless of temperature.

When illness has been diagnosed, the system is used to attempt to restore balance. Thus, in Central America, some diarrheas in children are viewed as hot, and protein-rich "hot" foods such as meats are withheld, aggravating the malnutrition that may be present and may be exacerbated by the diarrheal disease (Scrimshaw & Hurtado, 1988). An extensive literature exists on the topic of hot and cold illness classifications and treatments for them advocated by many of the world's cultures.

Energy balance is particularly important in Chinese medicine, where it is referred to as *chi*. When this balance is disturbed, it creates internal problems of homeostasis. Foods (often following the hot/cold theories) and acupuncture are among the strategies used to restore balance (Topley, 1976).

Blood beliefs include the concept that blood is irreplaceable; thus loss of blood—even small amounts—is perceived as a major risk. Adams (1955) describes a nutritional research project in a Guatemalan village where this belief inhibited the researcher's ability to obtain blood samples until the phlebotomists were instructed to draw as little blood as possible. Also, villagers were told that the blood would be examined to see if it was "sick" or "well" (another belief about blood) and they would be informed and given medicines if it were sick, which in fact did occur.

Menstrual blood is regarded as dangerous, especially to men, in many cultures, and elaborate precautions are taken to avoid contamination with it (Buckley & Gottlieb, 1988). As seen in the Guatemalan example, blood may have many properties that both diagnose and explain illness. Bad blood is seen as causing scabies in South India (Beals, 1976, p. 189). Haitians have a particularly elaborate blood belief system, which includes concepts such as *mauvais sang* (literally, "bad blood," when blood rises in the body and is dirty), *saisissement* (rapid heartbeat and cool blood, due to trauma), and *faiblesses* (too little blood). Blood qualities may also be seen as "opposites," such as clean–unclean, sweet–normal, bitter–normal, high–normal, heavy–weak, clotted–thin, and quiet–turbulent (C. Scott, personal communication, 1976). It is easy to see how these concepts could be used in a current program to prevent HIV infection in a Haitian community, because the culture already has ways of describing problems with blood.

Dislocation of body parts may occur with organs, but also with a physical aspect, such as the fontanel or "soft spot" in a baby's head where the bones have not yet come together in the first year or so to allow for growth. From the outsider perspective, a depression in this spot can be indicative of dehydration, often due to diarrheal disease. From the insider perspective, it is referred to as a cause of the disease (*caida de mollera*) in Mexico and Central America.

Many cultures associate illness with problems in specific organs. Good and Good (1981) talk about the importance of the heart for both Chinese and Iranian cultures. They discuss a case in which problems with cardiac medication were wrongly diagnosed for a Chinese woman who kept complaining about pain in her heart. In fact, she was referring to her grief over the loss of her son. The Hmong people of Laos link many problems to the liver, referring to "ugly liver," "difficult liver," "broken liver," "short liver," "murmuring liver," and "rotten liver." These terms are said to refer to mental and emotional problems, and so are idiomatic rather than literal (O 'Connor, 1995, p. 92; Thao, 1986).

Topley (1976) mentions incompatibility of horoscopes between mother and child in Chinese explanations for some children's illnesses.

Emotional Illnesses

Illnesses of emotional origin are important in many cultures. Sorrow (as in the case of the Chinese woman mentioned previously), envy, fright, and stress are often seen as causing illnesses. In a Bolivian village in 1965, for example, Scrimshaw was told that a young girl's smallpox infection was attributed to her sorrow over the death of her father.

Envy can cause illness because people with envy could cast the "evil eye" on someone they envy, even unwittingly, or the envious person can become ill from the emotion (Reichel-Dolmatoff & Reichel-Dolmatoff, 1961). Fright, called *susto* in Latin America, has already been mentioned. In addition to the case of tuberculosis in adults discussed previously, *susto* is a common explanation for illness in children. It is also mentioned in Chinese culture (Topley, 1976).

Weather

Everything from the change of seasons to unusual variations within seasons (too warm, too cold, too wet, too dry) can be blamed for causing illness. Winds, such as the Santa Ana in California or the Scirocco in the North African desert, are also implicated as sources of illness in many cultures. From the outsider perspective, changes in seasons can be associated with increases in risk and incidence of disease. For example, in many areas people have a greater chance of contracting malaria, cholera, and gastrointestinal infections during the rainy season.

Vectors or Organisms

Vectors or organisms are blamed for illness in some cultures and represent a blend of Western biomedical and indigenous concepts. "Germs" is a catch-all category, as is "parasites." Worms are seen as causing diarrhea, whereas flies are seen as causing illness and, sometimes, as carrying germs.

The Supernatural

The supernatural is another frequently viewed source of illness, especially in Africa and Asia, though this belief system is certainly not confined to those regions. In fact, the evil eye is a widespread concept—someone deliberately or unwittingly brings on illness by looking at someone with envy, malice, or too hot a gaze. In cultures where most people have dark eyes, strangers with light eyes may be seen as dangerous. In Latin America, a light-eyed person who admires a child can risk bringing evil eye to that child, but can counter it by touching the child. In other cultures, touching the child can be unlucky, so it is important to learn about local customs. Frequently, amulets and other protective devices, such as small eyes of glass, red hats, and a red string around the wrist, are worn to prevent evil eye. These objects can be viewed as an opportunity to discuss preventive health measures, because they are an indication that people are thinking about prevention.

Bewitching is deliberate malice, done either by the individual who wishes someone ill (literally) or by a practitioner at someone else's request. Bewitching can be countered by another practitioner or by specific measures taken by an individual. In some regions of Africa, epidemics are blamed on "too many witches," and people disperse to get away from them, thereby reducing the critical population density that had previously sustained the epidemic (Alland, 1970).

Belief in soul loss is widespread throughout the world. Soul loss can be caused by sources such as fright, bewitching, evil eye, and demons. It can occur in adults and children. Soul loss is serious and can lead to death. It must be treated through rituals to retrieve the soul. In Bolivia, for example, a village priest complained to Scrimshaw that his attempt to visit a sick child was thwarted when the family would not allow him to enter the house. The family later reported that an indigenous healer was performing a curing ritual at the time, and the soul was flying around the house as they were trying to persuade it to reenter the child. Opening the door to the priest would have allowed the soul to escape. In the Western biomedical system, this child's symptoms would have been attributed to severe malnutrition.

Spirit possession is also a worldwide belief, and one that is found especially frequently in African and Asian cultures. One of the best-known accounts of this phenomenon is *A Spirit Catches You and You Fall Down* (Fadiman, 1997), a moving story of seizure disorders in a Hmong community and the misunderstandings between the family and physicians. In another example, from South India, Beals (1976) mentions spirit possession in a daughter-in-law whose symptoms included refusing to work and speaking insultingly to her mother-in-law. He suggests that spirit possession is a "culturally sanctioned means of psychological release for oppressed daughters-in-law" (p. 188). Freed and Freed (1967) discuss similar cases in other regions of India. In Tanzania, malaria in children is sometimes blamed on possession by a bird spirit (Kamat, 2008). In Haiti, spirit possession is seen as a mark of favor by the spirits and is actively sought out. One of the drawbacks, however, is that the possessing spirits object to the presence of foreign objects in the body; as a consequence, some women do not want to use intrauterine devices as a means of birth control.

Demons are viewed as causing illness in Chinese culture, while offending God or gods is a problem in other cultures (Topley, 1976). In South India, epidemic diseases such as chickenpox and cholera (and, formerly, smallpox) are believed to be caused by disease goddesses. These goddesses bring the diseases to punish communities that become sinful (Beals, 1976, p. 187). The concept of punishment from God is seen in a case study from Mexico, where onchocerciasis (river blindness), which is caused by a parasite transmitted by the bite of a fly that lives near streams, is often thought to be due to sins committed either by the victim or by relatives of the victim. These transgressions against God are punished by God closing the victim's eyes (Gwaltney, 1970).

Food

In many cultures, food is perceived as being able to cause illness through its role in the hot and cold belief system; through spoiled foods, dirty foods, or raw foods; and by combining the wrong foods. Sweets are implicated as a cause of worms in children, and children who eat mud or dirt may become ill. Foods may also cause problems if eaten at the wrong time of day, such as "heavy" foods at night. An extensive literature describes food beliefs and practices worldwide, which has important implications for public health practice.

Sexual Illnesses

In Ecuador in the early 1970s, children's illnesses were sometimes blamed on affairs between one of the

child's parents and a *compadre* or *comadre*—one of the child's godparents (Scrimshaw, 1974). Such a relationship was viewed as incestuous and dangerous to the child. In India, sex is sometimes viewed as weakening to the man, so overindulgence is considered a cause of weakness. To return to the concept of blood beliefs, it is thought that 30 drops of blood are needed to make one drop of semen, so blood loss weakens a man.

Heredity and Old Age

Heredity is sometimes blamed for illness, early death, or some types of death. Similarly, old age may be the simple explanation given for illness or death. Heredity as noted here is expressed as an insider view, but is also part of the Western biomedical body of evidence on the causation or risk for many diseases.

Illness in Various Forms

TABLE 2-1 illustrates the way in which some of these beliefs are used to explain a particular illness—in this case, diarrheal disease in Central America. It is typical of the way in which an illness may be seen as having different forms, or manifestations, with different etiologies. It is also typical of the way in which several different explanations may be put forth for one set of symptoms.

In this case, Table 2-1 and **FIGURE 2-1** (the diagram of treatments) were key in expanding the orientation of the Central American diarrheal disease program. The program had intended to emphasize the distribution of oral rehydration solutions (ORS) in the clinics, but the insider perception was that a child should be taken to the clinic only for the worst form of diarrhea, dysentery. Instead, the most common treatment for diarrhea consisted of fluids in the form of herbal teas or sodas with medicines added. Often, storekeepers and pharmacists were consulted. It made sense to provide the ORS at stores and pharmacies as well as at clinics, so that all diarrheas were more likely to be treated (Scrimshaw & Hurtado, 1988).

In a related situation, Kendall, Foote, and Martorell (1983) found that, when the government of Honduras did not include indigenous or "folk" terminology for diarrheal disease in its mass-media messages regarding oral rehydration, people did not use ORS for diarrheas attributed to indigenously defined causes.

Healers

EXHIBIT 2-2 lists types of healers, which range from indigenous practitioners to Western biomedical providers. Pluralistic healers are those who mix the two traditions, although some Western biomedical healers and those from other medical systems may also mix traditions in their practices.

As with explanations of disease, the types of healers listed in Exhibit 2-2 are found in different combinations in different cultures. There is always more than one type of healer available to a community, even if members have to travel to seek care. The 16-country study of health-seeking behavior described earlier found that in all communities, people used more than one healing tradition, and usually more than one type of healer (Scrimshaw, 1992). The process of diagnosing

TABLE 2-1 Taxonomy of Diarrhea

CAUSE			SYMPTOMS All types have watery and frequent stools	TREATMENT
Mother's milk	Physical activity	**Hot**		Not breastfeeding when hot
	Hot foods			Mother changes diet
	Pregnancy			Breastfeeding stops
	Anger	**Emotional**	Very dangerous	Home, drugstore, injectionist, witch, spiritualist
	Sadness			
	Fright			

Food		Bad food	Flatulence, feeling of fullness	Home, folk curer
		Excess		
		Does *not* eat on time		
	Hot	Quality		
	Cold			
Tooth Eruption			Tooth eruption	None
Fallen fontanel, fallen stomach	Fallen stomach		Green with mucus Sunken fontanel; vomiting; green in color	Folk curer
	Fallen fontanel			
Evil eye			Fever	Folk curer
Stomach Worms			Worms	Drugstore, home, folk curer
Cold enters Stomach	From feet		White in color	Folk curer
	From head			
Dysentery			Blood in stools, "urgency;" color is red or black	Home, drugstore, health post

Courtesy of Scrimshaw, S. C. M., & Hurtado, E. (1987). *Rapid assessment procedures for nutrition and primary health care: Anthropological approaches to improving program effectiveness (RAP)*. Los Angeles: UCLA Latin America Center, p. 26. Reprinted with permission of the Regents of the University of California.

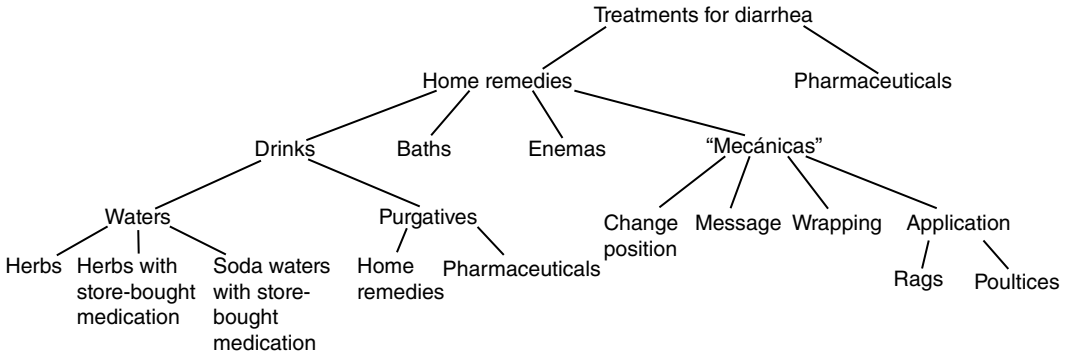


FIGURE 2-1 Taxonomy of treatments for diarrhea.

Courtesy of Scrimshaw, S. C. M., & Hurtado, E. (1987). *Rapid assessment procedures for nutrition and primary health care: Anthropological approaches to improving program effectiveness (RAP)*. Los Angeles: UCLA Latin America Center, p. 26. Reprinted with permission of the Regents of the University of California.

EXHIBIT 2-2 Types of Healers

Indigenous

- Midwives
- Doulas
- Shamans
- Curers
- Spiritualists
- Witches
- Sorcerers
- Priests
- Diviners
- Herbalists
- Bonesetters
- Massagers

Pluralistic

- Injectionists
- Indigenous health workers
- Western-trained birth attendants
- Traditional chemists/herbalists
- Storekeepers and vendors

Western Biomedical

- Pharmacists
- Nurse–midwives
- Nurses
- Nurse practitioners
- Physicians
- Dentists
- Other health professionals

Other Medical Systems

- Chinese medical system
  - Practitioners
  - Chemists/herbalists
  - Acupuncturists
- Ayurvedic practitioners
- Taoist priests

illness and seeking a cure has been referred to as "patterns of resort," a descriptor that is now favored over the older term "hierarchy of resort" (Scrimshaw & Hurtado, 1987). People may zigzag from one practitioner to another, crossing from one type to another type of healer, and not always starting with the simplest and cheapest, but with the one they can best afford and who they believe will be most effective, given the severity of the problem. Even middle- and upper-class individuals, who can afford Western biomedical care, may use other types of practitioners and practices.

Indigenous practitioners are usually members of the culture and follow traditional practices. Today, they often mix elements of Western biomedicine and other traditional systems. In many instances, they are "called" to their profession through dreams, omens, or an illness, which usually can be cured only by their agreement to become a practitioner. Most learn through apprenticeship to other healers, although some are taught by dreams. Often, they will take courses in Western practices in programs such as those developed to train Chinese "barefoot doctors" or community-based health promoters. In some instances, they must conceal their role as traditional healer from those running the training programs. The incorporation of some Western biomedical knowledge and skills often enhances a practitioner's prestige in the community.

Some indigenous practitioners charge for their services, but many do not, accepting gifts instead. In a few traditions (including some Chinese cultures), practitioners are paid as long as family members are well, but they are not paid for illness treatment. The duty of the practitioner in those cases is to keep people well, which argues for the acceptability of prevention programs in those cultures.

For the most part, indigenous practitioners do "good," meaning healing. Some can do both good and evil (e.g., shamans, sorcerers, and witches in many cultures). A few in these categories practice only evil or negative rituals, which aim to cause harm—for example, for revenge or to counteract the good fortune of a rival. Their work must then be countered by someone who does "good" magic. The power of belief is such that if individuals believe they have been bewitched, they may need a counteractive ritual, even if the Western biomedical system detects and treats a specific disease. In Guayaquil, Ecuador, one woman believed that she had been *maleada* (cursed) by a woman who was jealous of her, and that this curse was making her and her children ill. A *curandera* (curer) was brought in to do a *limpia* (ritual cleansing) of the house and family to remove the curse (Scrimshaw, 1974).

The importance of the power of belief is not confined only to bewitching. One anthropologist working with a Haitian population discovered that a Haitian burn patient made no progress until she went to a *Houngan* (voodoo priest) on the patient's behalf and had the appropriate healing ritual conducted (J. Halifax-Groff, personal communication, 1976).

In some cultures, healers are seen as diagnosticians, while others do the treatment (Alland, 1970). Other healers may handle both tasks, but refer some

kinds of illness to other practitioners. In Haiti, both midwives and voodoo priests refer some cases to the Western biomedical system. Healers who combine healing practices or who combine the ability to do both diagnosis and treatment are viewed as more powerful than other types. Topley (1976, pp. 259–260) discusses this issue in the setting of Hong Kong, noting that Taoist priest healers are particularly respected in that area. They are seen as both priest and doctor and "claim to combine the ethics of Confucianism, the hygiene and meditation of Taoism, and the prayers and self-cultivation of the Buddhist monk."

Pluralistic healers combine Western biomedical and indigenous practices. Injectionists will give an injection of antibiotics, vitamins, or other drugs purchased at pharmacies or stores. Sometimes these injections are suggested by the pharmacist or storekeeper; at other times they are self-prescribed. Because antibiotics proved so dramatically effective in curing infections when Western biomedicine was first introduced in many cultures, injections are often seen as conveying greater healing than the same substance taken orally. Thus, many antibiotics now available orally and vitamins are injected. In today's environment, this practice increases the risk of contracting HIV or hepatitis if sterile or new needles and syringes are not used.

Traditional chemists and herbalists, as well as storekeepers and vendors (many communities are too small to have a pharmacy), often sell Western biomedical medications, including those that require a prescription in the United States and Western Europe. While prescriptions may be "legally" required in many countries, the laws are not always rigorously enforced. This is also true for pharmacies, which are very important—sometimes the most important—sources of diagnosis and treatment in many communities around the world.

For more than 50 years, countries around the world have enlisted and trained indigenous health practitioners to function as part of the national or regional health system. These programs have ranged from China's "barefoot doctors" to the education of community members in Latin America, Africa, and Asia to provide preventive care and triage. These efforts have nearly disappeared in some areas (e.g., China) and reemerged in others. For example, in Australia, indigenous people are now involved as indigenous health outreach workers to their communities (2010, healthinfonet@ecu.edu.au). In Nepal, indigenous health workers have been enlisted in programs to address diarrheal disease and acute respiratory infections (Ghimire, Pradhan, & Mahesh, 2010), and female Ayurvedic doctors are important resources for women's health (Cameron, 2010).

Western biomedical practitioners are an important source of care, but they may also be expensive or difficult to access in remote areas. As mentioned earlier, if an individual believes that an illness is due to a cause explained by the indigenous system and a Western biomedical practitioner denies that cause, the individual may not return to that practitioner but rather seek help elsewhere (Kamat, 2008; Nichter, 2008).

As noted, there are other medical systems with long traditions, systematic ways of training practitioners, and well-established diagnostic and treatment procedures. Until recently, Western biomedical practitioners totally rejected both these and indigenous systems, often failing to recognize how many practices and medicines that Western biomedicine has "borrowed" from other systems (e.g., quinine, digitalis, many anesthetics, aspirin, and estrogen). Elements of these systems that were derided in the past, such as acupuncture, have now found their way into Western biomedical practice and are being "legitimized" by Western research (Baer, 2008).

Theories of Health Behavior and Behavior Change

The fields of sociology, psychology, and anthropology have developed many theories to explain health beliefs and behaviors and behavior change (Schumacher, Ockene, & Riekert, 2009). Some theories developed by sociologists and psychologists in the United States were developed first for U.S. populations and only later applied internationally. Others were developed with international and multicultural populations in mind from the beginning. Only a few of the many theories of health and illness beliefs and behavior are covered in this section; those included here have been quite influential in general or are applicable for international work in particular.

Health Belief Model

The health belief model suggests that decision making about health behaviors is influenced by four basic premises—perceived susceptibility to the illness, perceived severity of the illness, perceived benefits of the prevention behavior, and perceived barriers to that behavior—as well as by other variables, such as sociodemographic factors (Rosenstock, Strecher, & Becker, 1974). In general, people are seen as weighing perceived susceptibility (how likely they are to get the disease) and perceived severity (how serious the disease is) against their belief in the benefits and effectiveness of the prevention behavior they must undertake and the costs of that behavior in terms of barriers such as time, money, and aggravation. The more serious the disease is believed to be, and the more effective

the prevention, the more likely people are to incur the costs of engaging in the prevention behavior.

The health belief model has been extensively studied, critiqued, modified, and expanded to explain people's responses to symptoms and compliance with healthcare regimens for diagnosed illnesses. One concern has been that this model does not work as well for chronic problems or habitual behaviors because people learn to manage their behaviors or the healthcare system. Also, it has been accused of failing to take environmental and social forces into account, which in turn increases the potential for blaming the individual. The difficulty in quantifying the model for research and evaluation purposes is also a problem.

Work by Bandura led to the inclusion of self-efficacy in the model. *Self-efficacy* has been defined as "the conviction that one can successfully execute the behavior required to produce the desired outcome" (Bandura, 1977, 1989). The concept of *locus of control*, or belief in the ability to control one's life, also has been incorporated into this model. An example of this concept can be found in a comparison of migrant Yugoslavian and Swedish females with diabetes. The study revealed a stronger locus of control in the Swedish women and more passivity toward self-care in the Yugoslavian women, who also had a lower self-efficacy. The authors attributed these findings to the different political systems in the two countries—collectivism in Yugoslavia versus individualism in Sweden (Hjelm, Nyberg, Isacsson, & Apelqvist, 1999). A more recent example, among impoverished HIV-positive women in Kenya, found that the women preferred to use indigenous treatments to manage their symptoms, rather than taking available antiretroviral therapy. The women's shared cultural model of self-management enhanced the women's confidence in their own ability to care for themselves rather than take the Western biomedical drugs (Copeland, 2017).

The value of the four basic premises of the health belief model has held up well under scrutiny. Perceived barriers have the strongest predictive value of the four dimensions, followed by perceived susceptibility and perceived benefits. Perceived susceptibility is most frequently associated with compliance with health screening exams. Perceived severity of risk has been noted to have a weaker predictive value for protective health behaviors, but is strongly associated with sick-role behaviors.

In *Medical Choice in a Mexican Village*, Young (1981) describes a health decision-making process very similar to that found in the health belief model. In choosing between home remedies, pharmacy, or store, and between indigenous healer or doctor, the villagers weigh the perceived severity of the illness, the potential efficacy of the cure to be sought, the cost (money, time, and so on) of the cure, and their own resources to seek treatment and pay the cost as they make their decision. The simplest, least costly treatment is always the first choice, but the severity of illness and efficacy issues may force adoption of a more costly option. Other studies of health-seeking behavior have found similar patterns throughout the world (e.g., Kamat, 2008).

Theory of Reasoned Action

The theory of reasoned action was first proposed by Ajzen and Fishbein (1972) to predict an individual's intention to engage in a behavior in a specific time and place. This theory was intended to explain virtually all behaviors over which people have the ability to exert self-control. Five basic constructs precede the performance of a behavior: (1) behavioral intent, (2) attitudes and beliefs, (3) evaluations of behavioral outcomes, (4) subjective norms, and (5) normative beliefs. Behavioral intent is seen as the immediate predictor of behavior. Factors that influence behavioral choices are mediated through this variable. To maximize the predictive ability of an intention to perform a specific behavior, the measurement of the intent must closely reflect the measurement of the behavior. For example, measurement of the intention to begin to take oral contraceptives must include questions about when a woman plans to visit a clinic and which clinic she plans to attend. The failure to address action, target, context, and time in the measurement of behavioral intention will undermine the predictive value of the model.

In a test of this theory's ability to predict condom use intentions in a national sample of young people in England, measures of past behavior were found to be the best predictors of intentions and attenuated the effects of attitude and subjective norms (Sutton, McVey, & Glanz, 1999).

Diffusion of Health Innovations Model

The diffusion of health innovations model proposes that communication is essential for social change, and that diffusion is the process by which an innovation is communicated through certain channels over time among members of a social system (Rogers, 1983; Rogers & Shoemaker, 1972). An *innovation* is an idea, practice, service, or other object that is perceived as new by an individual or group. Ideally, the development of a diffusion strategy for a specific health behavior change goal will proceed through six stages:

1. Recognition of a problem or need
2. Conduct of basic and applied research to address the specific problem

3. Development of strategies and materials that will put the innovative concept into a form that will meet the needs of the target population
4. Commercialization of the innovation, which will involve production, marketing, and distribution efforts
5. Diffusion and adoption of the innovation
6. Consequences associated with adoption of the innovation

According to classic diffusion theory, a population targeted by an intervention to promote acceptance of an innovation includes six groups: Innovators, early adopters, early majority, late majority, late adopters, and laggards. The rapidity and extent to which health innovations are adopted by a target population are mediated by a number of factors, including relative advantage, compatibility, complexity, communicability, observability, trialability, cost-efficiency, time, commitment, risk and uncertainty, reversibility, modifiability, and emergence (see the *Innovation, Technology and Design* chapter for more information on scaling up health innovations).

Relative advantage refers to the extent to which a health innovation is better (faster, cheaper, more beneficial) than an existing behavior or practice. Antibiotics, for example, were quickly accepted in most of the world because they were dramatically faster and more effective than traditional practices.

Compatibility is the degree to which the innovation is congruent with the target population's existing set of practices and values. Polgar and Marshall (1976) point out that injectable contraceptives were acceptable in the village in India where Marshall worked because injections were viewed so positively due to the success of antibiotics.

The degree to which an innovation is easy to incorporate into existing health regimens may also affect rates of diffusion. Iodized salt is easier to use than taking an iodine pill, because consuming salt is already a habit. Health innovations are also more likely to be adopted quickly and by larger numbers of individuals if the innovation itself can be easily communicated.

The concept of trialability involves the ease of trying out a new behavior. For example, it is easier to try a condom than to be fitted for a diaphragm. Observability refers to role models, such as village leaders volunteering to be the first recipients in a vaccination campaign.

A health innovation is also more likely to be adopted if it is seen as cost-efficient. A famous case study of water boiling in a Peruvian town demonstrated that the cost in time and energy of gathering wood and making a fire to boil the water far outweighed any perceived benefits, so water boiling was seldom adopted (Wellin, 1955). Successful health innovations are likely to be those that do not require expenditure of much additional time, energy, or other resources.

One of the overall messages regarding communicating health education and promotion stated by Rogers (1973) is that mass media and interpersonal communication channels should both be used. Implementing both methods is of particular importance in low- and middle-income countries (LMICs), especially in rural communities. Rogers emphasizes that mass media deliver information to a large population and add knowledge to the general knowledge base, but interpersonal contacts are needed to persuade people to adopt new behaviors (thereby using the knowledge function, the persuasion function, and the innovation-decision process). In Rogers's work and other work cited by him, "family planning diffusion is almost entirely via interpersonal channels" (p. 263). Notably, Rogers presents five examples in different countries (including India, Taiwan, and Hong Kong), wherein interpersonal channels were the primary source for family planning information and were the motivating factors to seeking services.

The limitations of using mass media to disseminate health messages include the following issues:

- *Limited exposure.* In LMICs, smaller audiences have access to mass media. Radio continues to be an important mass-media tool, but social media via cell phones and other modes of Internet access are now a key way to reach people. Low literacy levels are another barrier.
- *Message irrelevancy.* The content of mass-media messages may be of no practical use for many rural and "non-elite" populations. Often instrumental information—"how to" information—is not included in the messages (e.g., where to receive services or the positive and negative consequences of adapting a particular health behavior).
- *Low credibility.* For people to accept and believe the messages being diffused, trustworthiness needs to exist between the sender and the receiver. In many LMICs, radio and TV stations are run by a government monopoly and their content may be considered to be government propaganda by the receivers. Radio and TV in Nigeria, Pakistan, and other African and Asian countries, for example, are controlled by the government (Rogers, 1973).

The diffusion of innovations model focuses solely on the processes and determinants of adoption of a new behavior and does not help to understand or

explain the maintenance of behavior change. Many health behaviors require permanent or long-term changes. Also, it is important to understand whether a new behavior is being carried out appropriately, consistently, or at all. One salient example involves condom use, which healthcare practitioners demonstrated to a population by unrolling the condom over a banana. Women who became pregnant while they reported using condoms had been faithfully putting them on bananas.

The rapid development of information technology—in particular, the use of smartphones—is revolutionizing diffusion of information and communication. For example, a smartphone innovation among men who have sex with men (MSM) consists of a computer app designed to facilitate finding sexual partners of men infected with a sexually transmitted disease. This app was used by the Monroe County Health Department in Rochester, New York, to map the spread of several sexually transmitted infections among this group, to document their sexual networks, and to advise sexual partners that they may have been exposed to an infection (Pennise et al., 2015).

Wearable devices range from sensors embedded in clothing to glasses to computerized watches and health-related wrist units. These devices have multiple applications, including fall risk assessment, quantifying sports exercise, studying people's habits, and monitoring the elderly (Hagthi, Thurow, & Stoll, 2017). Smartphone apps, in some cases, are motivating people to engage in more physical activity or eat healthier—for example, by tracking their daily steps or completing a food diary (Karpman, 2016). Others allow patients to upload their blood glucose readings, heart rates and other health data that they and their providers can monitor over time. Hagthi et al. (2017) note that "Based on consultation with expert scientists in environmental engineering and medicine, we believe that, motion trackers, gas detectors, and vital signs are the most important elements in health monitoring. . . ." Globally, smartphones and wrist or belt units are the most likely to be relevant for health-related applications in the near future.

Transtheoretical Model

Theories around the concept of stages of change have been evolving since the early 1950s. Currently, the most widely accepted stage change model is the transtheoretical model of behavior change developed by Prochaska, DiClemente, and Norcross (1992) (**FIGURE 2-2**). This model includes five core constructs: (1) stages of change, (2) decisional balance, (3) self-efficacy, (4) processes of change, and (5) maintenance. Interventions relying on this model are expected to include all four constructs in the development of strategies to communicate, promote, and maintain behavior change.

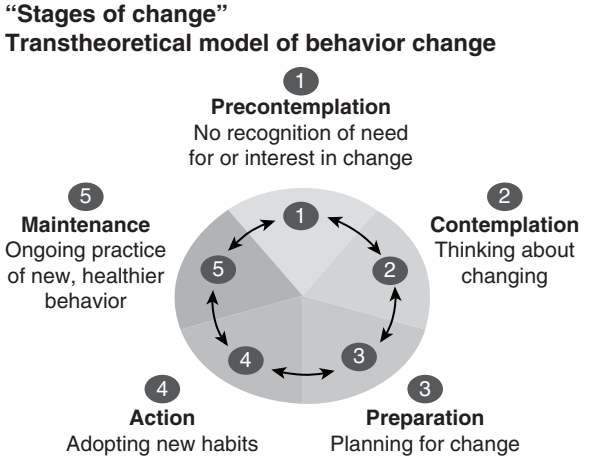


FIGURE 2-2 "Stages of change" transtheoretical model of behavior change.

Reproduced from Prochaska, J., DiClemente, C., & Norcross, J. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47,* 1102–1104.

The transtheoretical model identifies five stages of change. The first is precontemplation, in which individuals have no intention to take action within the next 6 months. In the contemplation stage, individuals express an intention to take some action to change a negative health behavior or adopt a positive one within the next 6 months. The preparation stage refers to the intent to make a change within the next 30 days. The action stage involves the demonstration of an overt behavior change for an interval of less than 6 months. In the fifth stage, known as maintenance, a person will have sustained a change for at least 6 months. The model helps healthcare providers, including those working in health education and communication, to develop appropriate strategies for reaching people at the various stages of readiness for change.

Decisional balance is an assessment of the costs and benefits of changing, which will vary with the stage of change. Self-efficacy is divided into two concepts within the transtheoretical model. First, confidence exists that one can engage in the new behavior. Second, the temptation aspect of self-efficacy refers to factors that can tempt one to engage in unhealthy behaviors across different settings.

The fourth construct of the transtheoretical model deals with the process of change. It includes 10 factors that can influence the progression of individuals from the precontemplation stage to the maintenance stage.

The Harm Reduction Model

Harm reduction is a concept that emerged from chemical dependency treatment programs, in which researchers realized that expecting those persons who were addicted to substances to become abstinent in one single step, without any relapse, was not possible for the vast majority. It drew on the ideas of the transtheoretical model that changing health behavior involves several steps (contemplation, planning, action) and may also involve relapse. The important insights of the harm reduction model are that (1) relapse is not failure and (2) making even a small step in improving health behavior often leads to additional positive changes (Lane, Lurie, Bowser, Kahn, & Chen 1999). The harm reduction model has been integrated into obesity-reduction programs, smoking cessation, and teen pregnancy prevention.

Explanatory Models

Explanatory models were initially proposed by the physician-anthropologist Kleinman (1980, 1986, 1988). They differ from some of the theories described earlier in this section in that they are designed for multicultural settings. They include models such as the meaning-centered approach to staff–patient negotiation described by Good and Good (1981). Although such models focus on individual interactions between physician or other staff and patients, the concepts underlying them—such as Kleinman's negotiation model—have proved useful for research and for behavioral interventions for larger populations. An explanatory model is seen as dynamic, and can change based on individual experiences with health, with health information, or with the illness in question (McSweeney, Allan, & Mayo, 1997).

EXHIBIT 2-3 adapts and summarizes concepts from Good and Good's (1981) description of the meaning-centered approach. This approach involves mutual interpretations across systems of meaning. The interpretive goal is understanding the patient's perspective. The underlying premise is that disorders vary profoundly in their psychodynamics, cultural influences in interpretation, behavioral expression, severity, and duration. As noted earlier, it is difficult to apply "codes" to culture and symptoms due to factors such as individual variations, groups assimilating or changing, and groups adding beliefs and behaviors from other cultures. For example, belief in *espiritismo* (spiritism) was traditionally strongest among Puerto Rican groups in the United States, but this belief has now been adopted by other cultures of Latin American origin as well. Thus, instead of trying to provide "formulas" for understanding health and illness belief systems within different cultures, the focus with the meaning-centered approach is on the meaning of symptoms. The medical encounter must involve the interpretation of symptoms and other relevant information.

Other Theories

A number of other theories can be useful in looking at culture and behavior. For example, the multi-attribute utility theory predicts behavior directly from an individual's evaluation of the consequences or outcomes

EXHIBIT 2-3 Meaning-Centered Approach to Clinical Practice

Primary Principles

- Groups vary in the specificity of their medical complaints.
- Groups vary in their style of medical complaining.
- Groups vary in the nature of their anxiety about the meaning of symptoms.
- Groups vary in their focus on organ systems.
- Groups vary in their response to therapeutic strategies. Human illness is fundamentally semantic or meaningful (it may have a biological base, but is a human experience).

Actions

Practitioners must:

- Elicit patients' requests, questions, and other communications.
- Elicit and decode patients' semantic networks.
- Distinguish disease and illness and develop plans for managing problems.
- Elicit explanatory models of patients and families, analyze conflict with biomedical models, and negotiate alternatives.

Corollary

- Clinical practice is inherently interpretive.

associated with both performing and not performing a given behavior. Some models, such as social learning theory, have been criticized by anthropologists who argue against the notion that people are like a "black box" into which you can pour information and expect a specific behavior change.

Common Features of Successful Health Communication and Health Promotion Programs

When applied in practice, many of the principles discussed in this chapter can help increase the success of health communication and health promotion programs. In particular, understanding and incorporating people's insider cultural values, beliefs, and behaviors; a community-based approach with strong community participation; recognition of gender issues (Zamen & Underwood, 2003); peer group education, including use of community-based outreach workers; and multilevel intervention approaches have proved essential to program success.

The Agita Sao Paulo Program in Brazil provides a case study in using local culture to design both the content and the delivery system for a program to use physical activity to promote health (Matsudo et al., 2002). Just the word *agita* (which means to move the body—to "agitate" in the sense of "stirring," but also to change) is more culturally understood and internalized than a literal translation of "exercise." In addition to representing careful work on culturally acceptable ways of delivering the message, this project provides multiple culturally valued ways to increase physical activity, and "tailors" these options to the age, gender, and lifestyles of community members.

In a very different project, work in three townships in South Africa focused on identifying where AIDS prevention would be most effective from the culturally appropriate, insider perspective (Weir et al., 2003). Among other things, researchers learned that ideal prevention intervention sites varied depending on whether the central business district or the township was the most popular location for initiating new sexual encounters. The type of sex (commercial versus casual) as well as the availability of condoms varied with the site. The age of people engaging in risky behaviors and risk behaviors by gender also varied by site. Again, prevention programs needed to be tailored.

In another HIV/AIDS prevention project, this time in Vietnam, paying attention to culture and religion was essential to program strategies (Rekart, 2002). In Belize, understanding adolescents and making sure the program met their needs in both cultural- and age-appropriate ways was key (Martiniuk, O'Connor, & King, 2003). In Nepal, the use of indigenous workers and attention to cultural practices helped lower the incidence and severity of diarrheal and respiratory infections in the districts targeted for interventions (Ghimire et al., 2010). Evaluation of programs addressing family planning and HIV prevention shows that behavior-change communication increases knowledge and interpersonal communication among audience members and motivates positive changes in behavior (Salem, Bernstein, Sullivan, & Lande, 2008).

Another example of focusing on understanding and changing cultural values around unhealthy behaviors is found in the area of smoking cessation. Abdullah and Husten (2004) set forth a framework for public health intervention in this area that addresses multiple levels of society.

The need for the involvement of communities is also clearly demonstrated in the literature, as literally hundreds of references exist on this topic. A recent summary article outlines many of the broad principles underlying this approach, including community analysis with community participation, action plans designed with community input, and community involvement in implementation (Nguyen-Truong, Tang, & Hsiao, 2017). Community involvement may take the form of ongoing oversight and evaluation as well as the more usual modes of using community outreach workers (e.g., Thevos, Quick, & Yanduli, 2000), working through community organizations, and getting individuals involved (Bhuyan, 2004). A report from a recent project in Bolivia documents the success of involving community members in everything from mapping the villages to priority setting for the program (Perry, Shanklin, & Schroeder, 2003). The former head of the United Kingdom's National Health Service, Nigel Crisp, argues strongly that the quest for global health in the twenty-first century must involve a paradigm shift in which nations, communities, and indigenous peoples around the world have a much greater voice in the design and implementation of health services (Crisp, 2010).

Two projects in Chicago demonstrate the success of the community outreach worker approach. In one case, the project focuses on intravenous drug abusers, helping them to reduce their HIV/AIDS-related risk behaviors and to initiate drug abuse treatment programs. This work simply could not have been accomplished without the efforts of community outreach workers, all of whom are former addicts who know how and when to reach

current addicts. Also, the outreach workers come from the predominant cultural/ethnic group in each community (Booth & Wiebel, 1992; Wiebel, 1993; Wiebel et al., 1996). Similarly, the Chicago Project for Violence Prevention involves ex-gang members as outreach workers. Both programs have been adopted internationally as well as in other cities in the United States. A similar focus on peer group education in Botswana led to increased knowledge and prevention behaviors among women at risk for HIV/AIDS infection (Norr, Norr, McElmurray, Tlou, & Moeti, 2004).

Methodologies for Understanding Culture and Behavior

Many of the research methodologies developed in the United States do not translate easily, literally, or figuratively to international settings. Differences in linguistic nuances, in the meanings of words and concepts, in what people would reveal to a stranger, and in what they would reveal to someone from their community have all complicated the application of the quantitative methodologies used by sociologists, psychologists, and epidemiologists. The realization of these problems came about gradually, through failed projects and missed interpretations, and especially once AIDS appeared. As a disease whose prevention is still largely behavioral, and with which many hidden or taboo behaviors are involved, AIDS highlighted the need for qualitative research and for research conducted by individuals from the cultures being studied. More recently, the Ebola pandemic of 2013–2016 reinforced the need to understand and address behaviors and beliefs as part of health interventions.

The field of global health has now moved from an almost exclusively quantitative orientation to the recognition that a toolbox of methodologies is available. Some of these tools may be more valuable than others for some situations or questions; at other times, a mix of several methodologies may offer the best approach. These methodologies derive from epidemiology, survey research, psychology, anthropology, marketing (including social marketing), and other fields. The biggest disagreement has been over the relative value of quantitative and qualitative methods.

The debate on the scientific value of qualitative versus quantitative research is well summarized by Pelto and Pelto (1978). They define *science* as the "accumulation of systematic and reliable knowledge about an aspect of the universe, carried out by empirical observation and interpreted in terms of the interrelating of concepts referable to empirical observations" (p. 22). The Peltos add that "if the 'personal factor' in anthropology makes it automatically unscientific, then much of medical science, psychology, geography, and significant parts of all disciplines (including chemistry and physics) are unscientific" (p. 23).

In fact, scientific research is not truly objective, but rather is governed by the cultural framework and theoretical orientation of the researcher. One example of this bias can be found in the past tendency of biomedical researchers in the United States to focus on adult men for many health problems that also occur in women (such as heart disease). The earlier example of kuru also demonstrates the limitations of cultural bias.

Qualitative research techniques include the following:

- Observation: Behaviors are observed and recorded.
- Participant observation: The researcher learns by participating in cultural events and practices.
- Interviews: Both open-ended and semi-structured queries are possible, usually based on interview guides or checklists.
- Focus groups: A group of people are asked to discuss specific questions and topics.
- Document analysis: Existing documents and prior research are evaluated.
- Systematic data collection: This technique ranges widely, from photography and videotaping to asking informants to draw maps; sort cards with pictures, words, or objects; answer questions based on scales; and many more (Bernard, 2013; Pelto & Pelto, 1978; Scrimshaw & Hurtado, 1987).

A key feature of qualitative research is the use of these multiple methods to triangulate, or compare, data so as to ensure accuracy. With these approaches, the researcher does not simply rely on what is said, but can observe what is actually done. Another feature is that the researcher spends enough time in the community to be able to interview, observe, or otherwise evaluate the same individuals or behaviors multiple times, thereby further ensuring the depth and accuracy of the resulting data.

These techniques yield data that are descriptive and exploratory, and that serve to investigate little-understood phenomena, identify or discover important variables, and generate hypotheses for further research. Results are often explanatory, helping researchers to understand the social and cultural forces causing the phenomenon and to identify plausible causal networks. They also present the "voices" of the participants, and introduce context and meaning into the findings. They yield themes, patterns, concepts, and insights related to cultural phenomena. They can be particularly

valuable for behaviors that are often hidden, such as sexual risk taking and drug abuse (Dickson-Gomez, 2010; Wiebel, 1993) In evaluations, they help practitioners make judgments about a program, improve its effectiveness, and inform decisions about future programming, as illustrated in the case study on acceptability of an infant cereal found later in this chapter.

The methodological concepts of validity and reliability provide a common foundation for the integration of quantitative and qualitative techniques. *Validity* refers to the accuracy of scientific measurement—"the degree to which scientific observations measure what they purport to measure" (Pelto & Pelto, 1978, p. 33). For example, in Spanish Harlem in New York City, a study using the question "¿Sabe como evitar los hijos?" ("Do you know how to avoid [having] children?") elicited responses on contraceptive methods and was used as the first in a series of questions on family planning. By not using family planning terminology at the outset, the study was able to avoid biasing respondents (Scrimshaw & Pasquariella, 1970). The same phrase in Ecuador, however, produced reactions like "I would never take out [abort] a child!" If the New York questionnaire had been applied in Ecuador without first testing it through semi-structured ethnographic interviews, the same words would have produced answers to what was, in fact, a different question (Scrimshaw, 1974). Qualitative methods often provide greater validity than quantitative methods because they rely on multiple data sources, including direct observation of behavior and multiple contacts with people over time. Thus, they can be used to increase the validity of survey research.

Reliability refers to replicability—the extent to which scientific observations can be repeated and the same results obtained. In general, this goal is best accomplished through survey research or other quantitative means. Surveys can test hypotheses and examine questions generated through qualitative data. Qualitative methods may help us discover a behavior or learn how to ask questions about it, while quantitative data can tell us how extensive the behavior is in a population and which other variables are associated with it.

Murray (1976) describes just such a discovery during qualitative research in a Haitian community, where a simple question—"Are you pregnant?"—had two meanings. Women could be pregnant with *gros ventre* ("big belly") or could be pregnant and in *perdition*. Perdition meant a state where a woman was pregnant, but the baby was "stuck" in utero and refused to grow. Perdition was attributed to causes such as "cold," spirits, or ancestors. Women may be in perdition for years, and may be separated, divorced, or widowed, but the pregnancy is attributed to her partner when it commenced. Murray subsequently included questions about perdition in a later survey, which revealed that it was apparently a cultural way of making infertility or subfecundity socially acceptable, as many women in perdition fell into these categories.

Surveys are effective tools for collecting data from a large sample, particularly when the distribution of a variable in a population is needed (e.g., the percentage of women who obtain prenatal care) or when rarely occurring events must be assessed (e.g., neonatal deaths). Surveys are also used to record people's answers to questions about their behavior, motivations, perception of an event, and similar topics. Although surveys are carefully designed to collect data in the most objective manner possible, they often suffer inaccuracies based on respondents' perceptions of their own behavior, their differing interpretations of the meaning of the question, or their desire to please the interviewer with their answers. Surveys also can encounter difficulty in uncovering motives (i.e., why individuals behave as they do), and they are not apt to uncover behaviors that may be consciously or unconsciously concealed. In "Truths and Untruths in Village Haiti: An Experiment in Third World Survey Research," Chen and Murray (1976) describe some of these problems.

The traditional anthropological approach involves one person or a small team who remain at the research site for at least a year. This practice is intended to ensure that the findings take into account the changes in people's lifestyles with the changes in seasons, activities, available food, and so on. Also, the anthropologist often needs time to learn a language or dialect and learn enough about the culture to provide a context for questions and observations. More recently, a subset of anthropological tools (ethnographic interview, participant observation, conversation, and observation) plus the market researchers' tool of focus groups have been combined in a rapid anthropological assessment process known as the Rapid Assessment Procedure (RAP) (Scrimshaw, Carballo, Carael, Ramos, & Parker, 1992; Scrimshaw, Carballo, Ramos, & Blair, 1991; Scrimshaw & Hurtado, 1987).

RAP evolved around the same time as Rapid Rural Appraisal was developed by rural sociologists (Chambers, 1992). Both methods made listening to community voices easier for program planners and healthcare providers and became frequently used tools for program development and evaluation. RAP is designed to involve local researchers who already know the language and much of the cultural context. Such procedures have been developed for many topics, including AIDS, women's health, diarrheal disease, seizure disorders, water and health, and childhood

obesity prevention. RAP has become a generic concept, and has been modified for many uses. Modified titles include RARE, ERAP, and FES (focused ethnographic study). In the past 20 years, the RAP methodologies have been embraced by community members, researchers, and funders alike and have been broadly used in community participatory research.

The case study on the use of focused ethnographic methods to assess the feasibility of introducing a fortified infant cereal in an African country, which appears in the next section of this chapter, is a good example of the use and value of this approach. With a relatively small number of interviews, researchers were able to establish that the cereal as constituted and packaged would be unlikely to succeed. Minor modifications (i.e., a cereal that did not require cooking and was packaged in small amounts) were recommended to change the product's likely acceptability.

In community participatory research, community members become involved in the design, conduct, and interpretation of research. This approach has been used most often for health intervention and behavior change programs where community acceptance of such interventions and programs is essential for success. It also has been found to increase the validity (accuracy) of the data, as community members are invested in developing programs that work.

A final comment on methodology is that as the social sciences are increasingly combining methodologies and sharing each other's tools, it is also important to share theoretical approaches. Where methodology is concerned, this leads to using multilevel approaches to research, in which environment, biological factors, cognitive issues, societal and cultural context, and political and economic forces all can contribute to the analyses. This should take place at least to the extent that an examination is made of data one step above and one step below the phenomenon being explained (Rubenstein, Scrimshaw, & Morrissey, 2000).

An example of a logic framework using this approach can be found in the work of the Centers for Disease Control and Prevention (CDC) task force that developed and maintains the *Guide to Community Preventive Services*—a series of evidence-based recommendations for community public health practice based on a systematic and critical review of the evidence. Topics considered in the guide include major risk behaviors (e.g., tobacco use, alcohol abuse and misuse, other substance abuse, nutrition, physical activity, healthy sexual behavior), specific illnesses (e.g., cancer, diabetes), and one overarching topic, the sociocultural environment. **FIGURE 2-3** presents the logic framework for this topic. The outcomes of community health (on the right side of the figure) stem

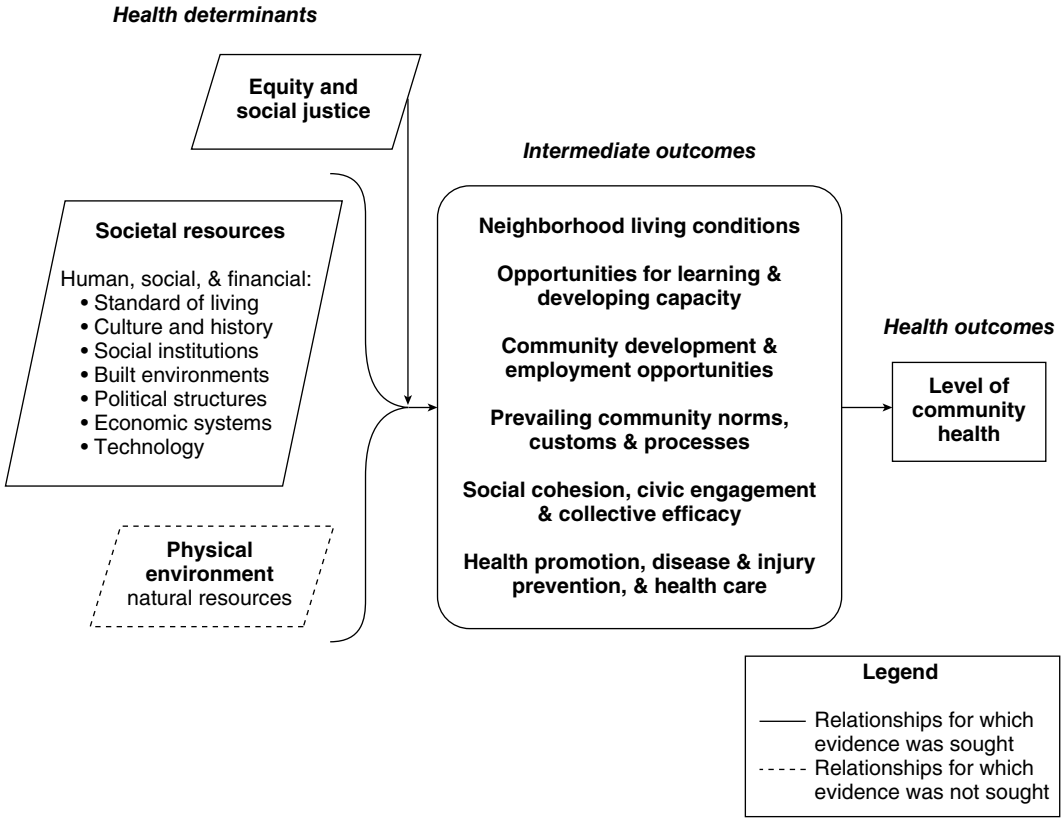


FIGURE 2-3 The *Guide to Community Preventive Services'* social environment and health model.

Reproduced from Anderson L. M., Scrimshaw S. C., Fuillilove M. T., et al. (2003). The community guide's model for linking the social environment to health. *American Journal of Preventive Medicine, 24*(35), p. 13. Reprinted with permission from Elsevier.

from factors in the physical environment and societal resources; outcomes related to equity and social justice issues derive from factors on the left side of the figure. The immediate outcomes, which are listed in the middle of the figure, range from neighborhood living conditions to prevailing community norms regarding prevention and health care (Anderson, Fielding, et al., 2003; Anderson, Scrimshaw, et al., 2003). This approach greatly broadens the context for understanding and addressing the health of individuals and of communities.

The CDC's *Community Guide* can be accessed through the following website: www.thecommunityguide.org. That website and related publications listed there provide evidence-based guidelines for improving community health, many of which have global relevance.

Case Study: Use of a Focused Ethnographic Study to Assess the Acceptability of a Fortified Infant Cereal in Africa†*

Earlier in this chapter, qualitative methods derived largely from anthropology were described as being important tools for ascertaining cultural facilitators and impediments to behavior changes that lead to improved health. As discussed then, these tools have been adapted for use in rapid assessment. This case study involves a rapid (focused) ethnographic study that was conducted before a new infant cereal was introduced to assess the potential success of this cereal by investigating household and local market behaviors.

Many infants and young children (IYC) in Africa continue to suffer from malnutrition or undernutrition. Where mothers breastfeed exclusively for at least six months, quality foods that complement the nutrients in human milk are important after six months when human milk alone is not adequate to meet nutritional needs.

In one African country, a project was devised to introduce a fortified cereal-based food that could help contribute to improved nutrition for infants and children consuming it. The aim of the focused ethnographic study was to determine whether families would purchase the food if available and, if the product was purchased, who actually consumed it. Why each of these steps did or did not occur was also important to assess. The price of the product was pegged at the amount that most urban families, even the poor, could afford to spend per day for a cereal.

It was important to the researchers to make the study framework broad enough to include key variables, so that they could avoid mistakes made by intervention projects that fail to take important cultural, economic, access, and other factors into account. A household perspective was chosen because the purchase and preparation of the cereal would be done at the household level.

Research questions included the following items:

- How much are households currently spending on food for their infants?
- Are their current expenditures providing a nutritionally adequate diet?
- Do households that are spending less than *x* amount (e.g., cents/day) have the potential to shift their expenditures?
- Which other factors determine household buying and/or preparation of foods for the infant or young child, and would these other factors interfere with making a switch, even if the family could afford it monetarily?

It was also important to take into account the possible foods for infants and young children in this environment: (1) human milk, (2) home-prepared foods that are made for family members and are also given to the infant or young child, (3) home-prepared foods that are made exclusively for the infant or young child, (4) commercial products that are marketed and purchased for household consumption, (5) commercial products that are marketed and purchased exclusively for the infant or young child, and (6) commercial products that are marketed for household consumption, but are purchased exclusively for the infant or young child.

The focused ethnographic study interviewed key informants—people who had personal knowledge and experience in an area of concern to the project. Four different techniques were used:

1. *Free listing exercises.* The respondent is asked a question or set of questions that elicit a series of items (objects, events, issues) pertaining to a particular cultural domain. For example, the interviewer can ask, "What are all the different places where a person can get food for infants or young children?"

†* This case study is based on Pelto, G. H., & Armar-Klemesu, M. (2010). *Focused ethnographic study to assess the potential of a commercial complementary food*. Report prepared for the Global Alliance for Improved Nutrition, Geneva, Switzerland.

2. *Open-ended interviewing, with guiding questions.* In open-ended questions, the interviewer writes down what the respondent says in response to a question without using precoding. Questions can be broad or narrow—for example, "How do you prepare cereal for your baby?"
3. *Rating and ranking exercises.* Respondents are asked to rate and rank items such as foods and sources of health care. Methods include handing respondents' cards with pictures of objects to be ranked and asking them to arrange these from most to least important, as well as asking them to assign the card to a slot along a continuum. (A similar technique for the latter method is the familiar scale of perceived pain.)
4. *Mapping exercises.* Respondents create visual maps on which they indicate the locations of specific features of concern for the researcher, such as places to obtain commercial foods.

In this case study, there were two main types of key informant-respondents: (1) women who gave information from the perspective of people who take care of children and (2) people who gave information from the perspective of marketing infant and young child foods. Thirty primary care givers, 10 alternate caregivers, and 12 sellers of these foods were interviewed. The sellers were divided into street venders and keepers of small shops.

The results of the study provided evidence to answer the key question about the potential acceptability of a new fortified cereal for infants and young children. They revealed that fortified cereals are, indeed, used and accepted, and that a relatively high proportion of the food budget for households with infants is spent on these items. Thus the key question about a dietary niche for a fortified cereal was answered affirmatively.

Importantly, the study uncovered the reality that a food that must be cooked, however briefly, is *unlikely* to be chosen over instant foods that do not require cooking. Busy mothers will spend more money to purchase prepared cooked cereals from street vendor or buy small packets of instant cereal that can be mixed with water rather than cooking a cereal themselves. This was extremely important guidance, as it showed that the cereal planned for introduction would need to be modified so that it would not require cooking if it was to be a success.

Because families were already spending so much for these foods, there was clearly a niche for a lower-cost fortified cereal, but it must be instant and available in small packets to be a financially viable option for families. The findings of this study were valuable in planning a program to introduce the cereal, providing the guidance that healthcare practitioners needed to proceed with an appropriately modified product, and to avoid spending time and money on something that would not work.

Case Study: The Slim Disease—HIV/AIDS in Sub-Saharan Africa‡*

AIDS changed the way in which epidemiologic and behavioral research is conducted and health interventions designed and carried out. This case study illustrates virtually all the topics covered in this chapter.

Epidemiology

As of 2016, an estimated 36.7 million adults and children were living with HIV/AIDS worldwide (UNAIDS, 2017). Nearly half (17.8 million) were women, and 2.1 million were children. Of the 36.7 million persons with HIV/AIDS, 19.4 million reside in eastern and southern Africa, among whom 59% are women and girls. In 2016, 77,000 new HIV infections in eastern and southern Africa occurred in children. Per capita rates of infection rates in the United States decreased by 18% between 2008 and 2014, as a result of a combination of aggressive use of antiretroviral therapies and prevention strategies (HIV.gov, 2017a, 2017b; WHO, 2010).

Unlike in the Western world, where AIDS was originally associated with gay men and injection-drug users, in Africa the most common route of transmission is through heterosexual sex. Men infect their partners (often wives) as a result of their involvement with other partners. A pregnant, HIV-positive woman may transmit the virus to her fetus through the placenta or to her infant through breastfeeding.

Generally, AIDS patients in Africa suffer from intestinal infections, skin disease, tuberculosis, herpes zoster, and meningitis. In the industrialized countries, AIDS is associated with Kaposi's sarcoma (a skin cancer), meningitis, and pneumonia.

‡* This case study was developed by Isabel Martinez, MPH.

Why does the same disease spread so differently from one region of the world to another? History, politics, economics, and cultural and social environments influence the course of a disease in a society. In the case of Africa, traditional family, social, and environmental structures were disrupted by European colonization, which imposed changes on the existing culture. Even after countries became independent from Europe, their political, ecological, and economic structures remained disrupted and often unstable. Many of these factors contributed to an environment in which AIDS easily took hold (Akeroyd, 1997; Bond, Kreniske, Susser, & Vincent, 1997; Hunt, 1989; Jok, 2001). These factors and their association with the AIDS pandemic are described in the following subsections. In addition to illustrating the relationship between cultural norms, prevention and healthcare access, and disease, this case study demonstrates the profound relationship between the general sociocultural, political, physical, and economic environment and health.

Risk of AIDS Associated with Migratory Labor

The integral family structure of the African culture has been broken up by the migratory labor system in eastern, central, and southern Africa. This system was historically part of the region's industrial development and colonization by European powers. These large industries, which include mining, railroad work, plantation work, and primary production facilities (e.g., oil refineries), have absorbed a massive labor influx from rural areas. Men typically leave their homes and travel outside their communities to work sites, where they remain for long periods of time. This system has not only kept families apart, but also increased the numbers of sex partners for men—in turn, giving rise to a higher prevalence of sexually transmitted infections (STIs) and later AIDS. In many African cultures, regular sex is believed essential to health. Men in the migratory labor system have sex with prostitutes close to their work sites, become infected, and eventually return home and infect their wives, whose babies may in turn become infected (Hunt, 1989; Salopek, 2000).

War

In 2017, there were 14 major armed conflicts and 29 additional armed conflicts globally. Of these 43 conflicts, 21 occurred in Africa. A country at war typically faces a weakening of its political system, and this situation in Africa has intensified the impact of the AIDS epidemic. Several populations become more vulnerable to HIV/AIDS during wartime, including those affected by food emergencies and scarcity, displaced persons, and refugees. Women are especially at risk. They are six times more likely to contract HIV in refugee camps than women in populations that reside outside such camps. In addition, women are often victims of rape as a weapon of war by the enemy side. Armed forces and the commercial sex workers with whom soldiers interact are also affected by the epidemic (Akeroyd, 1997; Carballo & Siem, 1996; Jok, 2001, 2012; UNAIDS, 1999; United Nations, 1999a; Uppsala Conflict Data Program, n.d.; Wallensteen & Harbom, 2009).

Gender Roles and Cultural Traditions

The African woman's struggle with the AIDS pandemic has been depicted often in the literature (Akeroyd, 1997; Carballo & Siem, 1996; Hunt, 1989; Messersmith, 1991; Salopek, 2000; UNAIDS, 1999; Watkins, 2004). The risk to women from husbands or partners returning from work in other areas has already been discussed. Another risk—sex work or prostitution by women as a means of survival—is now almost a death sentence in Africa, considering the great risk of contracting HIV/AIDS through such employment. There are many reasons why some African women find the need to engage in sex work, although studies have linked most of these reasons to a political economy context. Sex in exchange for favors, material goods, or money is conducted in all socioeconomic levels, from female entrepreneurs in foreign trade having to use sexual ploys to ensure business to impoverished young women needing money to support themselves and their families (Swidler & Watkins, 2007). Even if women in sex work are knowledgeable about preventing HIV infection through use of condoms, their cost and availability, combined with the resistance of some males to use them, raise barriers for the safety of these women and play a part in further transmission of the disease (Akeroyd, 1997; Messersmith, 1991).

Having multiple sexual partners has increasingly been implicated in raising the risk for HIV for both men and women (Helleringer, Kohlerb, & Kalilani-Phiric, 2009). Called "concurrency," these practices are now a major focus of intervention efforts (Shelton, 2009).

Other cultural factors that place young women at greater risk for HIV infection include a superstition in some areas that having sex with a virgin will cure an HIV-infected man. Adolescent girls and young women are placed at higher risk of exposure to HIV due to the behavioral practices of engaging in sex with older men (Tulio de Oliveira et al., 2017). The practice of female circumcision also places young girls at risk. In both of these circumstances, the risk of contracting HIV through sex or infected surgical instruments increases for adolescents (Akeroyd, 1997; Salopek, 2000).

Additional Cultural Beliefs

Secrecy regarding HIV/AIDS is common within some sub-Saharan cultures. Denying that AIDS is affecting one's community or that one is infected increases the chances that the virus will be transmitted to other people because preventive actions are not taken (Akeroyd, 1997; Salopek, 2000; UNAIDS, 1999; United Nations, 1999d). Preventive actions go beyond preventing sexual transmission, to include concerns about transmission during treatment of ill individuals and during funeral practices.

In some parts of Africa, AIDS is referred to as the "slim disease" because of the wasting away that occurs as a result of the infections. Because of this belief, men prefer sex with plump women, believing that they are not infected. AIDS is called "white man's disease" in Gabon and "that other thing" in Zimbabwe. HIV and AIDS are a source of shame and denial in these African cultures. AIDS is also considered a punishment for overindulgence of the body. One *sangoma* (faith healer), who has helped revive an ancient Zulu custom of virginity testing of young girls, supported her belief in reviving this custom by saying, "We have adopted too many Western things without thinking, and we lost respect for our bodies. This has allowed things like AIDS to come torture us" (Akeroyd, 1997; Hunt, 1989; Salopek, 2000; UNAIDS, 1999).

Barriers to Prevention or Treatment of HIV/AIDS

Barriers to prevention of HIV/AIDS include lack of financial resources and allocation of funds to projects that might be less crucial than those related to health. For example, a foreign country funded a multimillion-dollar hospital in Zambia, even though the rural clinics where the majority of the population live are often not even stocked with aspirin (Bartholet, 2000; Salopek, 2000).

Changing people's health behavior and addressing cultural beliefs has also been a tough challenge when it comes to prevention efforts. Promoting safe sex and the use of contraception, as well as abstaining from some cultural rituals, can be perceived as changing traditional gender roles for both men and women, and may go against some religious values that are part of the core for some communities. The need to hide or look away from the problem of HIV/AIDS stems from the disgrace attached to the disease, which makes it difficult for people even to discuss it, much less be tested for this infection. The stigma of HIV/AIDS needs to be removed for prevention efforts to be more widely accepted by the African people (Akeroyd, 1997; Bartholet, 2000; Salopek, 2000; UNAIDS, 1999; United Nations, 1999c).

One project in Ghana used both the health belief model and social learning theory to examine the determinants of condom use to prevent HIV infection among youth. The authors of the study found that perceived barriers significantly interacted with perceived susceptibility and self-efficacy. Youth who perceived a high level of susceptibility to HIV infection and a low level of barriers to condom use were almost six times as likely to have used condoms at last intercourse. A high level of perceived self-efficacy and a low level of perceived barriers increased the likelihood of use three times (Adih & Alexander, 1999).

Prevention Efforts by Community and Governmental Agencies and Nongovernmental Organizations

In the 1990s, Uganda and Senegal reduced their HIV infection rates through aggressive public education and condom promotion campaigns, expanded treatment programs for other STIs, mobilization of nongovernmental organizations (NGOs), and reduction of stigma for people with HIV/AIDS. Health officials believe the education efforts surrounding AIDS have contributed to women delaying the onset of sexual intercourse and increased condom use among sex workers and men and women who have casual sex (UNAIDS, 1999; United Nations, 1999a).

The theory of self-efficacy has proved useful in addressing AIDS. For example, one study in South Africa found that knowledge of risk and its prevention was important, but not sufficient to change behavior. The authors stress the need to improve personal autonomy in decision making about sexual behavior and condom use for both men and women through skills development programs that promote self-efficacy (Reddy, Meyer-Weitz, van den Borne, & Kok, 1999).

The United Nations and its specialized agencies have created major programs to assist countries and communities in prevention efforts, including joining forces to accelerate the development of experimental vaccines. Academic institutions have also teamed up with local community and church organizations to create prevention projects and help organize the communities to reach more of the public. These efforts have assisted in empowering many volunteers, mostly women, to motivate others in their communities through education and increasing women's negotiation skills for safe sex or condom use (Msiza-Makhubu, 1997; United Nations, 1999d; WHO, 1997).

There is also a growing movement in which doctors in Africa work with traditional healers to do outreach and education on AIDS. As discussed earlier, traditional healers have better access to many populations. People seek their help because of tradition and lack of adequate health care (Associated Press, 2000; Green, 1994).

Antiretroviral Therapy

Donor agencies/organizations such as the Global Fund to Fight AIDS, TB, and Malaria; the U.S. President's Plan for AIDS Relief; the World Bank; the European Commission; WHO; and the Gates Foundation have aggressively provided testing for HIV/AIDS and antiretroviral (ARV) therapy during the past eight years, and infection rates have come down (UNAIDS, 2004a; WHO, 2004). Beginning in 2003, the United States implemented the President's Emergency Plan for AIDS Relief (PEPFAR/Emergency Plan). As of 2017, this program had provided antiretroviral treatment (ART) to more than 11 million HIV-infected people and supported HIV testing and counseling (HTC) for more than 56.7 million people (PEPFAR, 2017).

Innovative prevention programs such as the introduction of male circumcision in areas where it had not been practiced are also helping to reduce infection rates. Male circumcision has been found to help protect against infection, reducing transmission rates by as much as 60% (Bailey & Mehta, 2009; Bailey et al., 2007; Tobian et al., 2009; Westercamp & Bailey, 2007).

Antiretroviral Treatment Challenges

Diminished political and economic support for antiretroviral programs could lead to the interruption of treatment of HIV/AIDS patients, which in turn would provide the HIV virus with the potential to become drug resistant. Other challenges in Africa include a shortage of health professionals, many of whom have left their home countries for better opportunities in higher-income countries. In addition, a lack of treatment literacy poses a huge challenge for effective antiretroviral treatment (UNAIDS, 2004b).

The individual behaviors that place people at risk are part of the larger root causes of the problem in Africa, including colonialism, big industry's design of mass labor migration, poverty, gender inequalities, and war. The ideal prevention and intervention strategies would address health behavior changes as well as economic and community barriers to the provision of social services and treatment options (Akeroyd, 1997; Bond et al., 1997; Tylor, 1871; United Nations, 1999c, 1999d; WHO, 1997).

Ebola

The Ebola virus shares several characteristics with HIV/AIDs. First identified in 1976, Ebola is a newly emergent infection. Beginning as a zoonotic disease that infected fruit bats, nonhuman primates, and other mammals, it "jumped species" to infect humans as a result of ecological disruption—in this case, hunting wild animals for food or what is called "bush meat." An Ebola pandemic in 2013–2016 that was centered in the West African countries of Liberia, Sierra Leone, and Guinea infected some 17,145 individuals, among whom 6,070 died. The short interval between recognition of symptoms and death of 6–16 days, the appalling fatality rate of 25% to 90%, and the lack of effective antimicrobial agents and preventive vaccines produced panic that often impeded global health strategies to reduce infection and provide supportive care to the infected.

The two hardest-hit countries—Liberia and Sierra Leone—are impoverished. Liberia ranks 182th out of 187 countries in the United Nations Human Development Index; Sierra Leone comes in at 180 (United Nations Development Programme, 2017). Cultural patterns of care for the sick and dying, as well as careful tending of the dead in preparation for burial, were found to increase Ebola transmission, because contact with body fluids—even sweat—in the context of cleaning, feeding, or moving an infected person can lead to transmission (CDC, 2015).

When American and European global health personnel first arrived to help fight the epidemic, in many cases they were greeted with fear and even violence by villagers, who, upon seeing the Westerners in hazmat suits, resisted their approach and often hid their sick family members. A new psychosocial paradigm called "fear-related behaviors" in situations of mass threat is now recognized as an expected reaction to disasters and must be planned for with clear communication, cultural understanding, and the close collaboration of foreign public health workers with local professional and community leaders (Espinola et al., 2016). Viewed from this perspective, the panicked responses of the African villagers were not irrational (Richardson et al., 2016). Beginning with the slave trade and colonialism, Western incursions into Liberia and Sierra Leone have caused immense suffering. Further complicating the cultural misunderstandings is the fact that Liberia has 31 separate languages and Sierra Leone has 25 (Rodriguez, 1997).

With Ebola deaths mounting and the evidence that the health systems of the affected West African countries were under-resourced and overwhelmed, numerous well-meaning multinational, governmental, and nongovernmental agencies rushed to the scene to help.

Grieving African family members were not allowed to touch or kiss their dead. Rather than observing time-honored funereal practices, they were instead obliged to place their deceased loved ones in plastic bags, to be buried with chloride disinfectant. Those and other measures to prevent transmission represented enormous changes in cultural practices around death and dying, but they worked to reduce transmission (Agusto, Teboh-Ewungkem, & Gumel, 2015). Analyses of the Ebola pandemic, and the behavioral changes that led to decreases in transmission, demonstrated that serious attention to clear communication and understanding of cultural patterns are critical elements of responses to outbreaks and disasters.

Conclusion

This chapter has briefly explored cultural and behavioral issues that influence global health. Anthropology, sociology, and psychology have much greater depth in terms of both methods and theories than can be described in this chapter. A rich and extensive literature exists on health beliefs and behaviors, environmental and biological contexts, health systems, and programmatic successes and failures. It is essential to take these factors into account when contemplating global health work. In addition, a program must consider structural factors, such as setting, hours, child care, and ambience, as well as factors of content, such as culturally acceptable services, which includes providers who treat patients with respect and understanding.

Research and preventive services regarding health beliefs and behaviors must accept and integrate concepts different from those held by Western biomedicine, by middle- or upper-class healthcare providers, or by healthcare providers from an ethnic or cultural group that is different from their patients. This requirement demands the ability inherent in some of the anthropological methods and approaches discussed earlier—that is, the ability to "get into someone's head" and understand things from an insider perspective. There is nothing like the experience of spending time with people, in their own homes or community, and striving to reach that insider understanding.

Acknowledgments

Susan Scrimshaw would like to thank Carolyn Cline, Rose Grignon, and Lisa Brainard for their assistance in editing and preparing the bibliography; Pamela Ippoliti for her editorial assistance; Susan Levy and John Justino for providing key examples from the intervention literature; and Isabel Martinez and Janel Heinrich for their assistance with the literature search, for helpful comments on the chapter, and, in particular, for preparing and revising the case study on AIDS. Appreciation also goes to Carole Chrvala for sharing notes on the various intervention theories. In addition, we would like to thank and acknowledge Professor Gretel Pelto and Dr. Margaret Armar-Klemesu for permission to base the case study on the introduction of a fortified infant cereal on one of their recent projects.

Discussion Questions

1. Which prevention strategies for AIDS would you develop if you were the minister of health of a sub-Saharan African country? Which strategies would you use if you were a community leader? Would the strategies used for these two perspectives differ? If so, how? How would you address some of the cultural beliefs or traditions associated with HIV/AIDS mentioned in the case study?
2. If you were entering a community to introduce a health program, who would you talk to? What would you ask? Why?
3. Discuss the concepts of validity and reliability in research as they apply to the use of quantitative and qualitative methods. Next, discuss the same concepts as they apply to community participatory research.
4. What is the hot/cold illness belief system? Why is it important? How would you incorporate it into a maternal and child health program?
5. Many people believe that healers such as midwives and shamans are called to their profession by a greater spiritual power. What significance does this belief have for official health programs around the world? How should they address this belief?
6. If an indigenous practice seems peculiar to you, but does no apparent harm, what should you do?
7. How could you learn what people in a community really believe about health and illness?
8. Based on the theories of behavior change, create your own model by taking what you think is the best content from existing theories. Explain your reasoning.
9. How would you balance the need to interrupt Ebola transmission in a situation such as the recent West African outbreak with the importance of addressing the fears and panic of the population?

References

Abdullah, A., & Husten, C. (2004). Promotion of smoking cessation in developing countries: A framework for urgent public health interventions. *Thorax, 59*, 623–630.

Adams, R. N. (1955). A nutritional research program in Guatemala. In B. D. Paul (Ed.), *Health, culture, and community* (pp. 435–458). New York, NY: Russell Sage Foundation.

Adih, W. K., & Alexander, C. S. (1999). Determinants of condom use to prevent HIV infection among youth in Ghana. *Journal of Adolescent Health, 24*(1), 63–72.

Agusto, F. B., Teboh-Ewungkem, M. I., & Gumel, A. B. (2015). Mathematical assessment of the effect of traditional beliefs and customs on the transmission dynamics of the 2014 Ebola outbreaks. *BMC Medicine,13*, 96.

Ajzen, I., & Fishbein, M. (1972). Attitudes and normative beliefs as factors influencing behavioral intentions. *Journal of Personality and Social Psychology, 21*(1), 1–9.

Akeroyd, A. V. (1997). Sociocultural aspects of AIDS in Africa: Occupational and gender issues. In G. C. Bond, J. Kreniske, I. Susser, & J. Vincent (Eds.), *AIDS in Africa and the Caribbean* (pp. 11–32). Boulder, CO: Westview Press.

Alcalay, R., Ghee, A., & Scrimshaw, S. C. M. (1993, May–June). Designing prenatal care messages for low-income Mexican women. *Public Health Reports, 108*(3), 354–362.

Alland, A. (1970). *Adaptation in cultural evolution: An approach to medical anthropology.* New York, NY: Columbia University Press.

Allen, M., Auld, E., Logan, R., Montes, J. H., & Rosen, S. (2017). *Improving collaboration among health communication, health education, and health literacy. NAM Perspectives.* Discussion Paper. Washington, DC: National Academy of Medicine. Retrieved from https://nam.edu/improving-collaboration-among-health-communication-health-education-and-health-literacy

Anderson, L. M., Fielding, J. E., Fullilove, M. T., Scrimshaw, S. C., Carande-Kulis, V. G., & Task Force on Community Preventive Services. (2003, April). Methods for conducting systematic reviews of the evidence of effectiveness and economic efficiency of interventions to promote healthy social environments. *American Journal of Preventive Medicine, 24*(3S), 25–31.

Anderson, L. M., Scrimshaw, S. C., Fullilove, M. T., Fielding, J. E., & Task Force on Community Preventive Services. (2003, April). The community guide's model for linking the social environment to health. *American Journal of Preventive Medicine, 24*(3S), 12–20.

Aviv, R. (2017, April 3). The apathetic: Why are refugee children falling unconscious. *New Yorker, 68-77.*

Baer, H. (2008). The emergence of integrative medicine in Australia: The growing interest of biomedicine and nursing in complementary medicine in a southern developed society. *Medical Anthropology Quarterly, 22*(1), 52–66.

Bailey, R. C., & Mehta, S. (2009). Circumcision's place in the vicious cycle between HSV-2 and HIV. *Journal of Infectious Diseases, 199*, 923–925.

Bailey, R. C., Moses, S., Parker, C. B., Agot, K., Maclean, I., Krieger, J. N., . . . Ndinya-Achola, J. O. (2007). Male circumcision for HIV prevention in young men in Kisumu, Kenya: A randomised controlled trial. *Lancet, 369,* 643–656.

Bandura, A. (1977). *Social learning theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist, 44*, 1175–1184.

Bartholet, J. (2000, January 17). The plague years. *Newsweek*, *135*(3).

Beals, A. R. (1976). Strategies of resort to curers in south India. In C. Leslie (Ed.), *Asian medical systems: A comparative study* (pp. 184–200). Berkeley, CA: University of California Press.

Bernard, R. (2013). *Social research methods: Qualitative and quantitative approaches* (2nd ed.). London, UK: Sage.

Betancourt, J. R., Green, A. R., & Carrillo, J. E. (2002). *Cultural competence in health care: Emerging frameworks and practical approaches.* New York, NY: Commonwealth Fund.

Bhuyan, K. (2004). Health promotion through self-care and community participation: Elements of a proposed programme in the development countries. *BMC Public Health, 4*, 11. Retrieved from http://www.biomedcentral.com/1471-2458/4/11

Board on Neuroscience and Behavioral Health, Institute of Medicine. (2002), *Committee on Communication for Behavior Change in the 21st Century: Improving the health of diverse populations. Speaking of health: Assessing health communication strategies for diverse populations.* Washington, DC: National Academy Press.

Bond, G. C., Kreniske, J., Susser, I., & Vincent, J. (1997). The anthropology of AIDS in Africa and the Caribbean. In G. C. Bond, J. Kreniske, I. Susser, & J. Vincent (Eds.), *AIDS in Africa and the Caribbean* (pp. 3–9). Boulder, CO: Westview Press.

Booth, R., & Wiebel, W. (1992). Effectiveness of reducing needle related risk for HIV through indigenous outreach to injection drug users. *American Journal of Addictions, 1,* 227–287.

Buckley, T., & Gottlieb, A. (Eds.). (1988). *Blood magic.* Berkeley, CA: University of California Press.

Cameron, M. (2010). Feminization and marginalization? Women Ayurvedic doctors and modernizing health care in Nepal. *Medical Anthropology Quarterly, 24*(1), 42–63.

Carballo, M., & Siem, H. (1996). Migration, migration policy and AIDS. In M. Knipe & R. Rector (Eds.), *Crossing borders: Migration, ethnicity and AIDS* (pp. 31–48). London, UK: Taylor and Francis.

Centers for Disease Control and Prevention (CDC). (2015, July 22). Ebola: Transmission. Retrieved from https://www.cdc.gov/vhf/ebola/transmission/index.html

Chambers, R. (1992). Rapid but relaxed and particularly rural appraisal: Towards applications in health and nutrition. In N. S. Scrimshaw & G. R. Gleason (Eds.), *Rapid assessment procedures: Qualitative methodologies for planning and evaluation of health related programmes* (pp. 295–305). Boston, MA: International Nutrition Foundation for Developing Countries.

Chen, K.-H., & Murray, G. F. (1976). Truths and untruths in village Haiti: An experiment in Third World survey research. In J. F. Marshall & S. Polgar (Eds.), *Culture, natality, and family planning* (pp. 241–262). Chapel Hill, NC: Carolina Population Center.

Collinge, J., Whitfield, J., McKintosh, E., Beck, J., Mead, S., Thomas, D. J., & Alpers, M. P. (2006). Kuru in the 21st century: An acquired human prion disease with very long incubation periods. *Lancet, 367*(9528), 2068–2074.

Copeland, T. (2017, May). Self-managing HIV/AIDS: Cultural competence and health among women in Nairobi, Kenya. *Anthropology and Medicine*, 1–15. doi:10.1080/13648470.2017.1285002

Crisp, N. (2010). *Turning the world upside down: The search for global health in the 21st century.* London, UK: Hodder Education.

de Oliveira, T., Kharsany, A., Gräf, T., Cawood, C., Khanyile, D., Grobler, A., . . . Abdool Karim, S. (2017). Transmission networks and risk of HIV infection in KwaZulu-Natal, South Africa: A community-wide phylogenetic study. *Lancet HIV, 4*(1), e41–e50.

Dickson-Gomez, J. (2010). Structural factors influencing the patterns of drug selling and use and HIV risk in the San Salvador metropolitan area. *Medical Anthropology Quarterly, 24*(2), 157–181.

Espinola, M., Shultz, J. M., Espinel, Z., Althouse, B. M., Cooper, J. L., Baingana, F., . . . Rechkemmer, A. (2016). Fear-related behaviors in situations of mass threat. *Disaster Health, 3*(4), 102–111.

Fadiman, A. (1997). *A spirit catches you and you fall down: A Hmong child, her American doctors, and the collision of two cultures.* New York, NY: Farrar, Straus, and Giroux.

Foster, G. M. (1953). Relationships between Spanish and Spanish-American folk medicine. *Journal of American Folklore, 66,* 201–217.

Freed, S. A., & Freed, R. S. (1967). Spirit possession as illness in a north Indian village. In J. Middleton (Ed.), *Magic, witchcraft, and curing* (pp. 295–320). Garden City, NY: Natural History Press.

Gadjusek, D. C., Gibbs, C. J., & Alpers, M. (1967). Transmission and passage of experimental "kuru" to chimpanzees. *Science, 155,* 212–214.

Garro, L. (2000). Remembering what one knows and the construction of the past: A comparison of cultural consensus theory and cultural schema theory. *Ethos, 28,* 275–319.

Ghimire, M., Pradhan, Y. V., & Mahesh, M. K. (2010). Community-based interventions for diarrhoeal diseases and acute respiratory infections in Nepal. *Bulletin of the World Health Organization, 88,* 216–221.

Good, B. J., & Good, M. J. D. (1981). The meaning of symptoms: A cultural hermeneutic model for clinical practice. In L. Eisenberg & A. Kleinman (Eds.), *The relevance of social science for medicine* (pp. 165–196). Dordrecht, Netherlands: Reidel.

Green, E. (1994). *AIDS and STDs in Africa: Bridging the gap between traditional healing and modern medicine.* Boulder, CO: Westview Press.

Guarnaccia, P. J., Lewis-Fernandez, R., Martinez Pincay, I., Shrout, P., Guo, J., Torres, M., . . . Alegria, M. (2010, May 1). *Ataque de nervios* as a marker of social and psychiatric vulnerability: Results from the NLAAS. *International Journal of Social Psychiatry, 56*(3), 298–309.

Gwaltney, J. L. (1970). *The thrice shy.* New York, NY: Columbia University Press.

Haghi, M., Thurow, K., & Stoll, R. (2017). Wearable devices in medical Internet of Things: Scientific research and commercially available devices. *Healthcare Informatics Research*, *23*(1), 4–15. http://doi.org/10.4258/hir.2017.23.1.4

Helleringer, S., Kohlerb, H., & Kalilani-Phiric, L. (2009). The association of HIV serodiscordance and partnership concurrency in Likoma Island (Malawi). *AIDS, 23,* 1285–1290.

HIV.gov. (2017a, November 20). The global HIV/AIDS epidemic. Retrieved from https://www.hiv.gov/federal-response/pepfar-global-aids/global-hiv-aids-overview

HIV.gov. (2017b, February 14). New HIV infections drop 18 percent in six years. Retrieved from https://www.hiv.gov/blog/new-hiv-infections-drop-18-percent-in-six-years

Hjelm, K., Nyberg, P., Isacsson, A., & Apelqvist, J. (1999). Beliefs about health and illness essential for self-care practice: A comparison of migrant Yugoslavian and Swedish diabetic females. *Journal of Advanced Nursing, 30*(5), 1147–1159.

Hughes, C. (1990). Ethnopsychiatry. In T. M. Johnson & C. E. Sargent (Eds.), *Medical anthropology: Contemporary theory and method.* (p. 131). New York, NY: Praeger.

Hunt, C. W. (1989). Migration labor and sexually transmitted diseases: AIDS in Africa. *Journal of Health in Social Science Behavior, 30*, 353–373.

Jok, J. M. (2001). *War and slavery in Sudan.* Ethnography of Political Violence Series. Philadelphia, PA: University of Pennsylvania Press.

Jok, J. M. (2012). Negotiating security: Gender, violence, and the rule of law in post-war South Sudan. In H. Stein & A. Fadlalla Hassan (Eds.), *Gendered insecurities, health and development in Africa* (pp. 154–169). New York, NY: Routledge.

Kamat, V. R. (2008). Dying under the bird's shadow: Narrative representations of *degedege* and child survival among the Zaramo of Tanzania. *Medical Anthropology Quarterly, 22*(1), 67–93.

Karpman, H. L. (2016, March 6). Use of smartphone applications for patients' health and fitness: Internal medicine alert. *Relias.* Retrieved from https://www.ahcmedia.com/articles/137485-use-of-smartphone-applications-for-patients-health-and-fitness

Kendall, C., Foote, D., & Martorell, R. (1983). Anthropology, communications, and health: The mass media and health practices program in Honduras. *Human Organization, 42,* 353–360.

Kleinman, A. (1980). *Patients and healers in the context of culture.* Berkeley, CA: University of California Press.

Kleinman, A. (1986). *Social origins of distress and disease.* New Haven, CT: Yale University Press.

Kleinman, A. (1988). *The illness narratives.* New York, NY: Basic Books.

Lane, S., Lurie, P., Bowser, B., Khan, J., & Chen, D. (1999). The coming of age of needle exchange. In Harrison, L. (Ed.), *Harm reduction* (pp. 47–68). London, Sage.

Lane, S. D., & Rubinstein, R. A. (1996). Judging the other: Responding to traditional female genital surgeries. *Hastings Center Report, 26*(3), 31–40.

Lane, S. D., Rubinstein, R. A. Keefe, R. H., Satterly, L. B., Huntington, S., Ramachandran, T., & Amaus Student Researchers. (2017). Action anthropology in a free clinic. *Human Organization, 76*(4), 336–347.

Lindenbaum, S. (1971). Sorcery and structure in fore society. *Oceania, 41*, 277–287.

Logan, M. H. (1972). Humoral folk medicine: A potential aid in controlling pellagra in Mexico. *Ethnomedizin, 4,* 397–410.

Long, A., Scrimshaw, S. C. M., & Hernandez, N. (1992). Transcultural epilepsy services. In N. S. Scrimshaw & G. R. Gleason (Eds.), *Rapid assessment procedures: Qualitative methodologies for planning and evaluation of health related programmes* (pp. 205–213). Boston, MA: International Nutrition Foundation for Developing Countries.

Martiniuk, A., O'Connor, K., & King, W. (2003). A cluster randomized trial of a sex education programme in Belize, Central America. *International Journal of Epidemiology, 32*, 131–136.

Matsudo, V., Matsudo, S., Andrade, D., Araujo, T., Andrade, E., de Oliveira, L., & Braggion, G. (2002). Promotion of physical activity in a developing country: The Agita São Paulo experience. *Public Health Nutrition, 5*(1A), 253–261.

Mattingly, C., & Garro, L. (Eds.). (2000). *Narrative and the cultural construction of illness and healing.* Berkeley, CA: University of California Press.

McSweeney, J. C., Allan, J. D., & Mayo, K. (1997). Exploring the use of explanatory models in nursing research and practice. *Image Journal of Nursing Scholarship, 29*(3), 243–248.

Messersmith, L. J. (1991). *The women of good times and Baba's Place: The multi-dimensionality of the lives of commercial sex workers in Bamako, Mali.* PhD dissertation, University of California at Los Angeles.

Msiza-Makhubu, S. B. (1997). *Peer education and support for AIDS prevention among women in South Africa.* PhD dissertation, University of Illinois at Chicago.

Murray, G. F. (1976). Women in perdition: Ritual fertility control in Haiti. In J. F. Marshall & Polgar, P. (Eds.), *Culture, natality, and family planning* (pp. 59–78). Chapel Hill, NC: Carolina Population Center.

National Center for Complementary and Integrative Health (NCCIH). (2014, December 17). NIH complementary and integrative health agency gets new name. Retrieved from https://nccih.nih.gov/news/press/12172014

National Center for Complementary and Integrative Health (NCCIH). (2017, September 24). Use of complementary health approaches in the U.S. Retrieved from https://nccih.nih.gov/research/statistics/NHIS/2012/key-findings

Nguyen-Truong, C. K., Tang, J., & Hsiao, C. Y. (2017) Community Interactive Research Workshop Series: Community members engaged as team teachers to conduct research. *Progress in Community Health Partnerships,11*(2), 215–221. doi:10.1353/cpr.2017.0026

Nichter, M. (2008). *Global health: Why cultural perceptions, social representations and biopolitics matter.* Tucson, AZ: University of Arizona Press.

Nielsen-Bohlman, L., Panzer, A. M., & Kindig, D. (Eds.); Committee on Health Literacy, Board on Neuroscience and Behavioral Health. (2004). *Health literacy: A prescription to end confusion.* Washington, DC: National Academies Press.

Norr, K., Norr, J., McElmurray, B., Tlou, S., & Moeti, M. (2004). Impact of peer group education on HIV prevention among women in Botswana. *Health Care for Women International, 25*, 210–226.

O'Connor, B. (1995). *Healing traditions.* Philadelphia, PA: University of Pennsylvania Press.

Office of Minority Health, U.S. Department of Health and Human Services. (2016). National Standards for Culturally and Linguistically Appropriate Services (CLAS) in health and health care. Retrieved from https://www.thinkculturalhealth.hhs.gov/assets/pdfs/EnhancedNationalCLASStandards.pdf

Parsons, T. (1948). Illness and the role of the physician. In C. Kluckholm & H. Murray (Eds.), *Personality in nature, society, and culture* (pp. 609–617). New York, NY: Alfred A. Knopf.

Paul, B. D. (Ed.). (1955). *Health, culture, and community.* (p. 1). New York, NY: Russell Sage Foundation.

Pelto, P. J., & Pelto, G. H. (1978). *Anthropological research: The structure of inquiry.* New York, NY: Cambridge University Press.

Pennise, M., Inscho, R, Herpin, K, Owens, J., Jr., Bedard, B. A., Weimer, A. C., . . . Younge, M. (2015). Using smartphone apps in STD interviews to find sexual partners. *Public Health Reports, 130*(3), 245–252.

Perry, H. Shanklin, D., & Schroeder, D. (2003). Impact of a community-based comprehensive primary healthcare programme on infant and child mortality in Bolivia. *Journal of Health, Population, and Nutrition, 21*(4), 383–395.

Polgar, S., & Marshall, J. F. (1976). The search for culturally acceptable fertility regulating methods. In J. F. Marshall & S. Polgar (Eds.), *Culture, natality, and family planning* (pp. 204–218). Chapel Hill, NC: Carolina Population Center.

President's Emergency Plan for AIDS Relief (PEPFAR). (2017, July 22). Home page. Retrieved from https://www.pepfar.gov/

Prochaska, J., DiClemente, C., & Norcross, J. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47,* 1102–1104.

Ratzan, S. C., & Parker, R. M. (2000). Introduction. In C. R. Selden, M. Zorn, S. C. Ratzan, & R. M. Parker (Eds.), *National Library of Medicine current bibliographies in medicine: Health literacy* (pp. v–vii). NLM Pub. No. CBM 2000-1. Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services.

Reddy, P., Meyer-Weitz, A., van den Borne, B., & Kok, G. (1999). STD-related knowledge, beliefs and attitudes of Xhosa-speaking patients attending STD primary health-care clinics in South Africa. *International Journal of Sexually Transmitted Diseases and AIDS, 10*(6), 392–400.

Reichel-Dolmatoff, G., & Reichel-Dolmatoff, A. (1961). *The people of Aritama.* London, UK: Routledge and Kegan Paul.

Rekart, M. (2002). Sex in the city: Sexual behavior, societal change, and STDs in Saigon. *Sexually Transmitted Infections, 78* (suppl I), i47–i54.

Richardson, E. T., Barrie, M. B., Kelly, J. D., Dibba, Y., Koedoyoma, S., & Farmer, P. E. (2016). Biosocial approaches to the 2013–2016 Ebola pandemic. *Health and Human Rights Journal, 1*(18), 1–13.

Rodriguez, J. P. (1997). Oppression of Liberia. In *The historical encyclopedia of world slavery* (p. 410). Santa Barbara, CA: ABC-CLIO.

Rogers, E. M. (1973). *Communication strategies for family planning.* New York, NY: Free Press.

Rogers, E. M. (1983). *Diffusion of innovations* (3rd ed.). New York, NY: Free Press.

Rogers, E. M., & Shoemaker, F. F. (1972). *Communication of innovations* (2nd ed.). New York, NY: Free Press.

Rosenstock, I., Strecher, V., & Becker, M. (1974). Social learning theory and the health belief model. *Health Education Monograph, 2*, 328–386.

Rubel, A. J., & Haas, M. R. (1990). Ethnomedicine. In T. M. Johnson & C. D. Sargent (Eds.), *Medical anthropology, contemporary theory and method* (pp. 115–131). New York, NY: Praeger.

Rubel, A. J., O'Nell, C. W., & Collado-Ardon, R. (1984). *Susto: A folk illness.* Berkeley, CA: University of California Press.

Rubinstein, R. A., Scrimshaw, S. C., & Morrissey, S. (2000). Classification and process in sociomedical understanding towards a multilevel view of sociomedical methodology. In G. L. Albrecht, R. Fitzpatrick, & S. C. Scrimshaw (Eds.), *The handbook of social studies in health & medicine* (pp. 36–49). London UK: Sage.

Salem, R. M., Bernstein, J., Sullivan, T. M., & Lande, R. (2008, January). *Communication for better health.* Population Reports, Series J, No. 56, 28 pages. Baltimore, MD: INFO Project, Johns Hopkins Bloomberg School of Public Health.

Salopek, P. (2000, January 10). We die lying to ourselves: Part 2. *Chicago Tribune*. Retrieved from http://www.chicagotribune.com/news/chi-060826salopek2-story-story.html

Schumacher, S. A., Ockene, J. K., & Riekert, K. A. (2009). *The handbook of health behavior change.* New York, NY: Springer.

Scrimshaw, S. C. (1974). *Culture, environment, and family size: A study of urban in-migrants in Guayaquil, Ecuador.* PhD dissertation, Columbia University, New York.

Scrimshaw, S. C. M. (1992). Adaptation of anthropological methodologies to rapid assessment of nutrition and primary health care. In N. S. Scrimshaw & G. R. Gleason (Eds.), *Rapid assessment procedures: Qualitative methodologies for planning and evaluation of health related programmes* (pp. 25–49). Boston, MA: International Nutrition Foundation for Developing Countries.

Scrimshaw, S. C. M., Carballo, M., Carael, M., Ramos, L., & Parker, R. G. (1992). *HIV/AIDS rapid assessment procedures: Rapid anthropological approaches for studying AIDS related beliefs, attitudes and behaviours.* Monograph. Tokyo, Japan: United Nations University.

Scrimshaw, S. C. M., Carballo, M., Ramos, L., & Blair, B. A. (1991). The AIDS rapid anthropological assessment procedures: A tool for health education planning and evaluation. *Health Education Quarterly, 18*(1), 111–123.

Scrimshaw, S. C., & Hurtado, E. (1987). *Rapid assessment procedures for nutrition and primary health care: Anthropological approaches to improving program effectiveness (RAP).* Tokyo, Japan: United Nations University.

Scrimshaw, S. C., & Hurtado, E. (1988). Anthropological involvement in the Central American diarrheal disease control project. *Social Science and Medicine, 27*(1), 97–105.

Scrimshaw, S. C., & Pasquariella, B. G. (1970). Obstacles to sterilization in one community. *Family Planning Perspectives, 2*(4), 40–42.

Scrimshaw, S. C. M., Zambrana, R. E., Alcalay, R., Estupinan, S., Luna, R., & Rivera, H. (1990). *Prenatal health behaviors in Mexico and mass communication/education strategies for improvement.* Washington, DC: International Center for Research on Women.

Shelton, J. (2009, August 1). Why multiple sexual partners? *Lancet, 374*, 367–368.

Simons, R. C. (2001, November 1). Introduction to culture-bound syndromes. *Psychiatric Times, 18*(11). Retrieved from http://web.mnstate.edu/robertsb/306/Intro%20to%20Culture%20Bound%20Syndromes.pdf

Simons, R. C., & Hughes, C. C. (Eds.). (1985). *The culture-bound syndromes: Folk illnesses of psychiatric and anthropological interest.* Dordrecht, Netherlands: D. Reidel.

Sutton, S., McVey, D., & Glanz, A. (1999). A comparative test of the theory of reasoned action and the theory of planned behavior in the prediction of condom use intentions in a national sample of English young people. *Health Psychology, 18*(1), 72–81.

Swidler, A., & Watkins, S. C. (2007). AIDS and transactional sex in rural Malawi. *Studies in Family Planning, 38*(3), 147–162.

Tervalon, M., & Murray-García, J. (1998). Cultural humility versus cultural competence: A critical distinction in defining physician training outcomes in multicultural education. *Journal of Health Care for the Poor and Underserved, 9*(2), 117–125.

Thao, X. (1986). Hmong perception of illness and traditional ways of healing. In G. L. Hendricks, B. T. Downing, & A. Deinard (Eds.), *The Hmong in transition* (pp. 365–378). New York, NY: Center for Migration Studies of New York & Southeast Asian Refugee Studies Project of the University of Minnesota.

Thevos, A. K., Quick, R. E., & Yanduli, V. (2000). Motivational interviewing enhances the adoption of water disinfection practices in Zambia. *Health Promotion International, 15*(3), 207–214.

Tobian, A., Serwadda, D., Quinn, T., Kigozi, G., Gravitt, P., Laeyendecker, O., . . . Gray, R. (2009, March 26). Male circumcision for the prevention of HSV2 and HPV infections and syphilis. *New England Journal of Medicine, 360*(13), 1298–1309.

Topley, M. (1976). Chinese traditional etiology and methods of cure in Hong Kong. In C. Leslie (Ed.), *Asian medical systems: A comparative study* (pp. 243–265). Berkeley, CA: University of California Press.

Tylor, E. B. (1871). *Primitive culture.* London, UK: J. Murray.

UNAIDS. (1999). *AIDS epidemic update: December 1999.* Geneva, Switzerland: World Health Organization.

UNAIDS. (2004a). *2004 report on the global HIV/AIDS epidemic: 4th global report.* Geneva, Switzerland: Author. Retrieved from http://files.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2004/GAR2004_en.pdf

UNAIDS. (2004b). *AIDS epidemic update: 2004.* Geneva, Switzerland: Author. Retrieved from https://reliefweb.int/report/world/aids-epidemic-update-2004

UNAIDS. (2017). Fact sheet: Latest statistics on the status of the AIDS epidemic. Retrieved from http://www.unaids.org/en/resources/fact-sheet

United Nations. (1999a). *Acting early to prevent AIDS: The case of Senegal.* Geneva, Switzerland: Author.

United Nations. (1999b, June). *Sexual behavioral change for HIV: Where have the theories taken us?* Geneva, Switzerland: Author. (English original, text by Rachel King).

United Nations, Wellcome Trust Centre for the Epidemiology of Infectious Disease. (1999d). *Trends in HIV incidence and prevalence: Natural course of the epidemic or results of behavioral change?* Geneva, Switzerland: Author.

United Nations Development Programme. (2017). Human development reports: Human development data (1990–2015). Retrieved from http://hdr.undp.org/en/data

Uppsala Conflict Data Program. (n.d.). UCDP database, Uppsala University. Retrieved from www.ucdp.uu.se

Wallensteen, P., & Harbom, L. (2009). Armed conflict, 1946–2008. Department of Peace and Conflict Research. *Harbom Journal of Peace Research, 46*(4), 577–587.

Watkins, S. C. (2004, December). Navigating the AIDS epidemic in rural Malawi. *Population and Development Review, 30*(4), 673–705.

Weir, S. S., Pailman, C., Mahalela, X., Coetzee, N., Meidany, F., & Boerma, J. T. (2003). From people to places: Focusing AIDS prevention efforts where it matters most. *AIDS, 17*(6), 895–903.

Weller, S. C. (1983). New data on intracultural variability: The hot–cold concept of medicine and illness. *Human Organization, 42,* 249–257.

Wellin, E. (1955). Water boiling in a Peruvian town. In B. D. Paul (Ed.), *Health, culture, and community* (pp. 71–103). New York, NY: Russell Sage Foundation.

Westercamp, N., & Bailey, R. C. (2007). Acceptability of male circumcision for prevention of HIV/AIDS in sub-Saharan Africa: A review. *AIDS and Behavior, 11*(3), 341–355.

Wiebel, W. (1993). *The indigenous leader outreach model: Intervention manual.* DHHS Publication 93-3581.Washington, DC: National Institute on Drug Abuse, National Institutes of Health, U.S. Department of Health and Human Services.

Wiebel, W., Jimenez, A., Johnson, W., Oulette, L., Jovanovic, B., Lampinen, T., . . . Utne O'Brien, M. (1996). Risk behavior and HIV seroincidence among out of treatment injection drug users: A four-year prospective study. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology, 12*, 282–289.

World Health Organization (WHO). (1997). *Women and HIV/AIDS prevention training manual: Peer education program,*

a global approach. Collaborating Centre for International Nursing Development in Primary Health Care Research. Geneva, Switzerland: Author.

World Health Organization (WHO). (2004). *Investing in a comprehensive health sector response to HIV/AIDS: Scaling up treatment and accelerating prevention. WHO HIV/AIDS plan.* Geneva, Switzerland: Author. Retrieved from http://www.who.int/entity/3by5/en/HIV_AIDSplan.pdf

World Health Organization (WHO). (2010, June). Programmes. Retrieved from http://www.who.int/hiv/en/

Young, C. J. (1981). *Medical choice in a Mexican village.* New Brunswick, NJ: Rutgers University Press.

Zamen, F., & Underwood, C. (2003). *The gender guide for health communication programs.* Baltimore, MD: Population Communications Services, Center for Communications Programs, Johns Hopkins.

CHAPTER 3

Global Health, Human Rights, and Ethics

Tonya Littlejohn, Chris Beyrer, and Nancy Kass

The recent emergence and reemergence of infectious disease epidemics, the persistent burden of non-communicable disease, the enduring impact of environmental change, and the intractable effects of social disadvantage and inequity have all contributed to the rising prominence of health challenges at a global level. The promise and process of globalization, with its increased interdependence of economic, political, and social domains, accompanied by the integration of goods, services, values, and people, translates into new and interconnected global health threats, worsening health disparities, and heightened global health insecurity.

It is in this context that the human rights and ethics implications of public health challenges and interventions must be pursued and interrogated. Human rights and ethics, encompassing the right to health as well as some degree of the right to non-interference, health equity, and the determinants of health, are central to global health. This centrality is recognized by the World Health Assembly and the World Health Organization (WHO), and is enshrined in international conventions and treaties with relevance to global health, including the Universal Declaration of Human Rights; the International Convention on Social, Cultural and Economic Rights; the Convention on the Eradication of All Forms of Discrimination Against Women; and the most widely ratified of all of these foundational agreements, the Convention on the Rights of the Child. All of these conventions agree that human rights paradigms and ethics principles are inextricably linked to health and offer a normative framework to understand and address complex health issues experienced at a global level. Analysis of emerging and persistent global health challenges through the human rights and ethics lenses can be a more effective and nuanced approach for framing and responding to global health challenges and allows for focus on both the proximal causes and the social determinants of ill health (Mann, 1997).

Defining Global Health in the Context of Human Rights and Ethics

Though many constructs of global health exist, this field is generally understood to refer to a phenomenon whereby the determinants of health or health outcomes supersede the territorial boundaries of any given state. Global health recognizes that health is determined by conditions, issues, and concerns that typically transcend national boundaries (Stapleton, Schroder-Back, Laser, Meershoek, & Popa, 2014). It highlights issues that may be universally experienced, and accepts that while some burdens and challenges are shared, many

health burdens may be profoundly inequitable and disproportionate. Global issues are often recognized as beyond the capacity of individual countries alone to address through domestic institutions (Stapleton et al., 2014). The world recently witnessed this reality during the 2013–2016 Ebola outbreak in West Africa. The "global" in global health identifies the scope of problem, not the location (Koplan et al., 2009). The U.S. Institute of Medicine (1988) describes public health as "what we, as a society, do collectively to assure the conditions for people to be healthy." This definition has equal stature in relation to global public health and has particular application to both human rights and ethics.

Human rights instruments more broadly, the right to health specifically, and ethics provide entry points for understanding and outlining underlying systemic and structural social determinants of health that, if addressed, could advance health equity at a global level. Human rights and ethics discourse highlight the conditions necessary for people to be healthy, and the means and considerations necessary to enable these conditions. With human rights paradigms, these conditions also include fundamentals of good or poor governance as well as civil and political liberties, including the right to information and the right to participate in political decision making. The ethics framework insists that issues of equity, autonomy, and benevolence be included as determinants of health programs and paradigms. Underlying the social determinants of health, and considered essential elements necessary to live a healthy life, these sociopolitical and economic forces drive both local and global health inequality and disproportionate burdens of disease and shape the health, legal, institutional, and structural contexts in which they are embedded. They concern the conditions in which individuals are born, grow, live, work, and age and that influence their health status (WHO, 2017).

Just as health threats extend globally, so must the solutions. Increasingly, it has been recognized that there is both a shared global responsibility to realize the right to health and a moral impetus to address health challenges that are universally experienced. Thus, health, ethics, and human rights efforts are, by definition, synergistic. One undertaking cannot be viewed in isolation without a considered reference and integration of the others. While public health endeavors can at times impact on individual human rights and challenge ethical principles and decision making, such as may occur in a global public health crisis, the underlying complementarity of these disciplines has greater prominence than the inherent confrontation between them (Beyrer, 2004).

Setting the Context: Human Rights and Ethical Approaches to Global Health

Both human rights and ethics offer compelling discourses and frameworks to advance global health and ameliorate health disparities and differences that have a global force or reach. At the intersection of global health, human rights, and ethics, a number of consistent themes and challenges emerge. These will be explored throughout the chapter and are briefly outlined here.

First, fundamental human rights and the right to health are interrelated and inseparable. The realization of one right (such as the right to housing) can enable fulfillment and protection of the right to health; conversely, the deprivation of a right (such as the right to participation in civic life) may increase vulnerability to ill health or impair enjoyment of other human rights (Office of the United Nations Commissioner for Human Rights, 2008). Rights abrogation or violation can have measurable impacts on health, such that if an individual or population group experiences discrimination, or a deprivation of their rights, their ability to engage with and benefit from health interventions is diminished (Beyrer, 2004). Although traditionally described as being in conflict, the pursuit of individual human rights and enhanced global health at a population level are increasingly recognized as complementary and interdependent approaches to defining and advancing human well-being. The central premise of securing health for all by targeting marginalized, discriminated-against, and vulnerable populations parallels the sentiment of human rights. Promoting and protecting human rights is inextricably linked with promoting and protecting health.

Ethics, in relation to global health, is both the application of principles and norms for moral guidance and a process for identifying, analyzing, and resolving ethical issues inherent in the practice of public health. Ethics, which has traditionally been considered to exist within the purview of individual behavior and with a bioethical focus, may be recast in the context of public health ethics with broad application to global public health and governance. In turn, global health ethics may be considered a normative project with a common set of principles or values to deal with global health threats (Stapleton et al., 2014).

Second, to stimulate a more considered analysis of health, human rights, and ethics, it is imperative to understand the disproportionate burden imposed by disease and ill health and to appreciate how these

differences influence health outcomes across a range of populations. Health disparities and the social determinants of health are significantly driven by complex structural factors that can contribute to poor health outcomes and adversely affect groups of people who have systematically experienced both a greater burden of ill health and exposure to numerous structural-level social determinants (Jürgens, Csete, Amon, Baral, & Beyrer, 2010; Sollom et al., 2011).

Third (and a related concept), discrimination is often the basis of fundamental structural inequality, such that populations who are marginalized or experience discrimination are more vulnerable to ill health. According to the Office of United Nations High Commission of Human Rights (2008), discrimination refers to any distinction, exclusion, or restriction made on the basis of various grounds that has the effect or purpose of impairing or nullifying the recognition, enjoyment, or exercise of human rights and fundamental freedoms. The implication is that some diseases or health conditions that have a global distribution, such as tuberculosis or leprosy, may singularly affect marginalized populations or be more pronounced in these populations. Marginalization, stigma, and discrimination may be both the cause and the consequence of ill health. In acknowledging the importance of discrimination and marginalization, it must also be noted that some groups—such as women, children, and prisoners, for instance—face an additional risk or vulnerability due to particular circumstance or biological, structural, and social realities (Rubenstein et al., 2016). This vulnerability may be generalized, or it may be specific to certain diseases or interventions.

Fourth, human rights frameworks relate to obligations and duties at the level of the state and, therefore, are dependent on a state-based system of governance. As a signatory to the various human rights covenants, the state has an obligation to respect, protect, and fulfill those covenants in relation to the right to health. Such conventions outline the requirement of states to refrain from interfering in rights, to protect rights from demands of nonstate actors (such as armed groups and private corporations), and to take positive measures to facilitate and enable rights such as provision of immunization, training of health professionals, access to justice, access to sanitation and clean water, and other essential functions within a national health system.

How do ethical principles and human rights obligations and duties apply to states, and are they enforceable? And if the state is the primary custodian in enacting the right to health for populations within its jurisdiction, who should be made responsible and accountable for global health issues that fall outside state boundaries? Global health issues that transcend the boundary of the state require a coherent and coordinated response through public health activities such as epidemic preparedness, surveillance, health system research, treatment access, and immunization (**EXHIBIT 3-1**). This role and function in relation to global health issues is taken up by a multitude of actors and stakeholders, including nongovernmental organizations, private agencies, and businesses, each of which has its own mandates, motivations, and scope, and each of which also engages in the provision of health at a global level (see also Gostin & Mok, 2009). It should not be assumed that all actors are acting with the similar intent or understand the right to health in the same way. Determining which health challenges should be elevated to a global level of governance is an ongoing debate. Indeed, one of the ongoing challenges to global health is the need to strengthen claims to the right to health and support universality of the right to health (Gable & Meier, 2013).

Friedman and Gostin (2015) discuss the notion that global health—in this case, referring to public health, universal health coverage, and the social determinants of health—cannot be achieved without concurrent advancement of justice and global health equity. This development would lead to gains in both the aggregate level of health and the distribution of health across populations. The World Health Organization, under the leadership of Tedros Ghebreyesus of Ethiopia, has committed to universal health coverage (UHC) as the primary means of achieving the goal of global equity in healthcare access. However, others propose that improvements in health markers, such as prevalence, mortality, and morbidity, do not necessarily translate directly into improvements in health equity and justice (see, for example, Braveman & Gottlieb, 2014). Some would argue that even though war, crime, hunger, poverty, illiteracy, homelessness, and related human rights abuses interfere with the health of individuals and populations, it does not mean that eliminating these conditions is part of the mission of public health (Rothstein, 2002).

States have a shared responsibility for supporting conditions in which their populations can be healthy, including the implementation of preventive health strategies, health equity, and universal health coverage under a mandate of health for all (see also Friedman & Gostin, 2015). Incorporating the right to health, global health governance provides the institutional, financial, and legal mandates and mechanisms necessary to meet core global health challenges and risks that transcend boundaries in their origin and impact. Global health governance is an appeal to collective

EXHIBIT 3-1 Ebola, Ethics and Human Rights

The most recent Ebola outbreak, which occurred between 2013 and 2016, exposed in a very specific way the limitations of the global effort—namely, the inequalities in global power structures, the responsiveness of global health governance (see chapter on *Global Health Governance and Diplomacy* for more on strengthening global health governance after the 2013–2016 Ebola outbreak), and the limited epidemiologic information available early on—allowed the virus to flourish uncharted and undetected for up to three months. The sociopolitical realities of containing a highly pathogenic virus were laid bare. The fragility of the region, following years of conflict and instability, was evident in the weakened and dysfunctional public health governance and infrastructure, debilitated and impoverished institutions, and diverted resources, which contributed to a lack of capacity and preparedness. Local cultural practices, which included traditional burial rituals of hand-washing the body, and a highly mobile and interconnected population moving across arbitrary state boundaries allowed unhindered transmission of the virus (Richardson, 2016). It was the perfect milieu fomenting contagion, lethality, stigma, and neglect (Donnelley, 2014; see also Farmer, 2014). That so many health workers, community members, and graveyard workers were nevertheless willing to risk their lives to care for Ebola-infected patients, the many thousands of exposed persons, and those who died of Ebola complications is a tribute to the extraordinary dedication and humanity of these persons. This example reminds us all that the greatest human rights and ethical imperative remains to care for the sick, even when doing so may put us in harm's way.

The effects of the Ebola epidemic on global health structures, public health ethics, and human rights were far-reaching. First, the outbreak raised questions about who has responsibility to protect populations from public health emergencies (see also Calain & Poncin, 2015). The state must fulfill its duties and fundamental obligations to enable its population's enjoyment of the highest attainable standard of health. Revisions made to WHO's International Health Regulations (IHR) in the wake of the severe acute respiratory syndrome (SARS) outbreak required states to institute minimum core public health standards, including public health surveillance, laboratory capacity, and epidemic preparedness. However, the Ebola outbreak exposed the world's vulnerabilities due to the impoverished health systems in a small region of West Africa, decimating and overwhelming these already fragile systems that were unable, despite significant efforts, to meet their populations' most basic health needs (Gostin & Ayala, 2017). The moral impetus to protect and respond must therefore account for the deeply disproportionate impact that an outbreak has on the most marginalized and vulnerable in the event that the national government is incapacitated and unable to ensure health security for its own population.

The Ebola response was also a teaching moment regarding the weakness of global health governance, most compellingly seen in the failure of the World Health Organization to mount an effective response (see, for example, Sands, Mundaca-Shah, & Dzau, 2016). Without the intervention of nongovernmental groups, most notably Medecins Sans Frontieres, the loss of life would have been much greater. Eventually, the international community did mobilize resources, but the prominent role assumed by military medicine practitioners, notably from the United States, again exposed the weaknesses in civilian capacity to respond. While military engagement in large-scale disasters and epidemics is often essential, the delegation of humanitarian and public health responses to military and other security forces raises additional human rights and ethical challenges, including concerns from the host countries that public health responses led by military actors can be cover for other—more specifically, military or security—objectives. The U.S. Central Intelligence Agency's use of polio immunization as a cover for investigating the whereabouts of Osama Bin Laden in Pakistan is the most unfortunate example of this kind of abuse: It continues to bedevil polio eradication efforts in Pakistan and Afghanistan, and has led to the targeted killing of many immunization staff, mostly women ("Editorial," 2014).

effort, with engagement of state actors and nonstate actors alike, in pursuit of a collective outcome (see also Dodgson, Lee, & Drager, 2002).

The final point relates to priority setting. Priority setting raises ethical concerns of fairness, justice, and equity. Moreover, when it is applied to particular global health challenges, context-specific concerns inevitably emerge regarding how finite resources may be allocated in a fair and equitable manner. Priority setting will be informed by such considerations as infrastructural capacity, urgency, social value, governance, and financing. Ethics can guide approaches when a conflict arises between competing values or needs. In the case of a global health crisis, for example, it may be more readily apparent how goods should be prioritized and who should be the primary beneficiary. Resources may be reallocated or reprioritized as a means of containment to protect the affected and the most vulnerable, and this may prevent unaffected populations from maintaining their health needs. In contrast, the best decisions may be less obvious when setting priorities in the provision of non-emergent healthcare strategies, such as obesity or smoking.

Global Health and Human Rights

> All human beings are born free and equal in dignity and rights. (Article 1, Universal Declaration of Human Rights [UDHR])

Human rights pertain to the fundamental freedoms, inherent value, and dignity of all human beings. These rights are universal, inalienable, and indivisible. They confer both freedoms and entitlements. The improvement of one right facilitates advancement of others, and similarly the deprivation of one right adversely affects the others. Human rights are independent of nationality, place of residence, gender, religion, race, or any other status (United Nations, 1948).

The founding document of the modern human rights movement, the Universal Declaration of Human Rights (1948), Article 25 (1), provides the minimum standards necessary to ensure adequate health and lays out the duties and obligations of the state for upholding the rights of individuals. The UDHR is not legally binding, but states have bestowed it with great legitimacy through their actions, including its legal and political invocation at the national and international levels. Health is mentioned once in the UDHR , within the context of an overall right to an adequate standard of living:

> Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including food, clothing, housing and medical care and necessary social services, and the right to security in the event of unemployment, sickness, disability, widowhood, old age of other lack of livelihood in circumstances beyond his control. (United Nations, 1948)

Under international human rights, sovereign states have an obligation to recognize the rights of individuals who reside within their borders.

Subsequent treaties and covenants give further articulation of the rights related to health (**EXHIBIT 3-2**). Article 12 of the International Covenant on Economic, Social, and Cultural Rights (ICESCR) of 1966, in particular, establishes the right to health. It has a significant focus on what was then called international health and would now be called global health, in the three of its four specific provisions (a, b, and c below):

1. The States Parties to the Present Covenant recognize the right of everyone to the enjoyment of the highest attainable standard of physical and mental health.
2. The steps to be taken by the States Parties to the present Covenant to achieve the full realization of this right shall include those necessary for:
   a. The provision for the reduction of the stillbirth rate and of infant mortality and for the health development of the child
   b. The improvement of all aspects of environmental and industrial hygiene
   c. The prevention, treatment, and control of epidemic, endemic, occupational, and other diseases
   d. The creation of conditions which would assure access to all medical services and medical attention in event of sickness

Interpreting the Right to Health

In 2000, the committee on ICESR issued a guideline that provides more substantive details on Article 12, the right to health. General Comment 14 on the right to health notes that health is a fundamental human right indispensable for the exercise of other human rights (Office of the High Commissioner for Human Rights, 2000). General Comment 14 proposes

EXHIBIT 3-2 International Human Rights Instruments Recognizing the Right to Health

- The 1965 International Convention on the Elimination of All Forms of Racial Discrimination: Article 5 (e) (iv)
- The 1966 International Covenant on Economic, Social, and Cultural Rights: Article 12
- The 1979 Convention on the Elimination of All Forms of Discrimination Against Women: Articles 11(1)(f), 12, and 14(2)(b)
- The 1989 Convention on the Rights of the Child: Article 24
- The 1990 International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families: Articles 28, 43(e), and 45(c)
- The 2006 Convention on the Rights of Persons with Disabilities: Article 25

Reproduced from United Nations. Human Right to Health. 2016.

that the right to health does not equate to the right to be healthy, but rather is an inherent, inalienable right to the highest attainable right that must be progressively realized through available, accessible, and affordable health care, services, and conditions. Realization of the right to health is more than the absence of disease. Indeed, it implies consideration of the conditions necessary for the realization of the highest attainable standard of health. In addressing these issues, General Comment 14 suggests that achieving this goal must take into account the "individual's biological and socio-economic preconditions and a state's available resources" (Office of the High Commissioner for Human Rights, 2000). This provision gives focus to the relationship between the state and the individual by asserting that the state cannot, by itself, guarantee good health, just as it cannot protect against all aspects of ill health (Office of the High Commissioner for Human Rights, 2000).

Additionally, General Comment 14 states that the right to health is interpreted as an inclusive right. Specifically, it extends "not only to timely and appropriate health care but also to the underlying determinants of health, such as access to safe and potable water and adequate sanitation, an adequate supply of safe food, nutrition and housing, health occupational and environmental conditions, and access to health-related education and information, including on sexual and reproductive health" (Office of the High Commissioner for Human Rights, 2000).

The Challenge in Meeting the Right to Health

Under international human rights law, a "right" is put forward as a claim to those social arrangements—norms, institutions, laws, and enabling environment—that can best secure the enjoyment of this right (WHO, 2002). The legal integrity of human rights analysis, as much as its moral appeal, is what gives the human rights movement its authority and force. Under international treaty law, states assume obligations and duties to respect, protect, and fulfill human rights, including the right to health. Violation of this right implies that the obligation and duties have not been met or fulfilled (Cohen, Kass, & Beyrer, 2007).

The rights-based approach to health includes consideration of the underlying determinants to good health such as water, housing, education, and food. Thus, the right to health includes the minimum core services, such as maternal–child health and primary health care, that governments must provide, even in places of scarce resources. While other rights may encapsulate the essence of a right to health, the challenge remains as to how the right to health may be fulfilled and measured. It is for this reason that some proponents believe a more constrained interpretation of the right to health allows for a more meaningful and representative invocation of this right (Tasiouslas & Vayena, 2015).

As part of the application of human rights principles to global public health challenges, public health strategies, such as prevention and treatment programs, must not infringe upon the human rights of those whom they are intended to benefit (**EXHIBIT 3-3**). Similarly, rights-based approaches must ensure that all individuals, and particularly those who experience greater marginalization, stigma, and discrimination, enjoy equal access to services. All too often, however, discrimination and other human rights issues have been both a consequence of global health challenges and a societal cause of vulnerability to these conditions (Mann, 1997).

At least two misconceptions about human rights tend to complicate their application. The first is that often global health challenges are perceived as being part of the realm of "economic, social, and cultural rights" (such as the right to health care) as opposed to "civil and political rights," such as the right to free expression, association, and due process of law. In fact, many of the violations of human rights that most increase risk—imprisonment without due process, censorship of health information, and violence and discrimination against women as articulated in the International Covenant on Civil and Political Rights (United Nations, 1966) and most national constitutions—have direct public health implications (Beyrer, 2007b). The fact that these rights abuses worsen health outcomes underscores what has been called the "indivisibility" of human rights norms—the notion that civil and political rights and economic, social, and cultural rights are mutually reinforcing and derive from a single principle, the concept of fundamental human dignity.

A second, related misconception is that human rights impose undue constraints on state sovereignty and restrict the legislative branch of government in its efforts to give effect to the "will of the people" (usually the majority) and/or enact effective public health policy. In fact, human rights guarantees are almost always embodied in constitutions or international treaties that have been ratified by national legislatures, and are enforced by judges who are either elected or appointed by elected officials. Moreover, human rights guarantees can come with certain limitations, as long as these limitations can be shown to be necessary and proportionate to a legitimate policy objective. Human rights analysis involves identifying the rights infringement in question (usually by reference to a specific legal guarantee), and then balancing that infringement against

EXHIBIT 3-3 Zika, Human Rights, and Ethics

The most recent Zika virus outbreak, declared a Public Health Emergency of International Concern in February 2016, and its disproportionate concentration on women of childbearing age, is an example of systemic injustice and inequality in global health. Although many of the clinical manifestations and health sequelae of Zika infection are not all known, the recent outbreak highlights a global health problem with contemporaneous individual-, community-, and country-level dimensions and ineluctable social and ethical impacts.

Acquisition of Zika during the antenatal period has been associated with miscarriage, stillbirth, and neurologic malformations including congenital microcephaly. Infection may occur at any stage of pregnancy, raising a particular challenge for intervention and prevention. Evidence that the Zika virus may also be transmitted sexually has significant implications for women who live in a region with some of the most restrictive sexual and reproductive laws,[1] where approximately 56% of pregnancies are unplanned and gender-based violence is pervasive (Guttmacher Institute, 2017; Roa, 2016).

In the most affected region of Latin America, it is estimated that 23 million women have an unmet need for contraception and account for 75% of unintended pregnancies (Guttmacher Institute, 2017). As many as 95% of pregnancy terminations are performed in unsafe conditions, and a number of countries (including El Salvador, Chile, and Dominican Republic) criminalize abortion under any circumstance. In other countries, abortion may be allowable only in cases of rape or in the event of a significant risk to the mother's life but not in the event of fetal congenital anomalies, including those resulting from Zika infection. In these cases, some women have been legally compelled to carry a pregnancy of a possibly fatal fetal anomaly to full term.

Women living in marginalized and poor communities in Brazil and elsewhere faced the greatest disproportionate burden and deprivation of basic protections and rights and their ability to exercise reproductive autonomy (see also Rasanatham, MacCarthy, Diniz, Torreele, & Gruskin, 2017).[2] Discrimination, economic inequality, and poverty entrench both the risks and impacts of Zika (Phelan & Gostin, 2016).

In response to the outbreak, governments in the region recommended women of childbearing age postpone pregnancy, but few provisions or pathways to avoid an unintended pregnancy were made available, shifting the burden from the domain of the public to the individual. Restrictions on access to contraception, including emergency contraception, safe termination, and post-abortion care, predominated.

Questions have emerged in cases regarding whether the intrusion of the state in the life and decisions of a woman is justifiable and proportional. Is the global public health threat of Zika severe enough to limit sexual and reproductive rights and freedoms? Conversely, is the threat of Zika-associated fetal anomalies sufficient to change existing restrictive laws and policies? There is a moral imperative to create conditions that allow women to decide freely and responsibly on the number, spacing, and timing of their children. Similarly, there is an ethical obligation to ensure equitable access to contraceptive options and safe reproductive services, as well as a duty to minimize harms associated with a Zika-affected pregnancy. However, as the case of Zika succinctly illustrates, to achieve the most optimal outcome, employing ethical principles requires a balancing of interests and context.

1 International human rights documents pertaining to women's reproductive health rights include the International Covenant on Economic, Social, and Cultural Rights [right to health, Article 12; right to nondiscrimination based on sex, Articles 2(2) and 3; right to special protection for mothers and their children, Article 10] and the Convention on the Elimination of All Forms of Discrimination Against Women (women's right to nondiscrimination within law, Articles 2 and 3; right to health, Article 12; and right to reproductive self-determination, Article 16).

2 Laid out by the Committee on Economic, Social, and Cultural Rights, the right to sexual and reproductive health includes both freedoms and entitlements, including the freedom to make free and responsible decision and choices, free of violence, coercion, and discrimination, over matters concerning one's body and sexual and reproductive health (Diniz, 2016).

competing policy objectives. Most importantly, perhaps, human rights aim to give voice to minorities who may be marginalized or disenfranchised by the democratic process (Decker et al., 2015).

Rights-Based Approach to Global Health

> Social justice and the protection of human rights do not ensure good health—free people can make poor choices, and affluence carries its own burdens of morbidity and mortality. But social justice and limitations on basic rights and freedoms, on human dignity itself, can have direct and indirect effects on the health of individuals, communities, and populations. (Beyrer, 2007a)

Jonathan Mann, a public health physician, was an early proponent of the integration of medicine, public health, ethics, and human rights (Gostin, 2001). He believed that health and human rights were indistinguishable.

Human rights violations lead to adverse health events and, therefore, have measurable impacts on physical and mental well-being (Gostin, 2001). Similarly, framing health disparities as a violation of human rights facilitates a more useful interpretation and implementation of the human rights instruments. The progressive realization of the right to health at a global level gives emphasis to the shared responsibility and obligations of the global community.

Mann's legacy has grown and expanded since his tragic early death in 1998. The field of health and human rights, and of public health or global health and rights, has made critical contributions to our understanding of health inequalities, disparities, and the social and structural determinants of health that have proved so challenging to address through technological or biomedical advances alone.

Global Health and Public Health Ethics

Public health ethics uses ethics principles and norms to identify, analyze, and resolve moral challenges in public health. As described earlier, when public health ethics focuses its lens on global public health challenges, it might be examining issues that extend beyond the boundaries of the state, such as pandemics, or global issues that recur within territories' or countries' own borders.

Public health ethics has largely concentrated its attention around two main areas: (1) balancing individual liberties with the ethical duty to improve the welfare of the public and the needs of many and (2) the moral duty to improve social justice as a means of improving the public's health. As such, public health ethics is concerned with the pursuit of equity in health, which then guides the approach taken to remedy the inherently unjust health inequalities. Further, public health ethics seeks to determine how to best balance the needs of a population as a whole and the liberties and rights of the individuals who make up those populations. An oft-cited example is the case in which the state executes its agency through a public health intervention to restrict individual choice, privacy, or freedom of movement in an effort to prevent or contain disease as a means to protect others in the community or to limit suffering.

The greater the burden imposed by a public health intervention or program (e.g., limitations on individual liberties), the stronger the evidence must be to demonstrate that it will achieve its goals (Kass, 2001). Hence, a restriction of liberty must be justified and proportionate to the perceived or known threat. Public health goals of producing benefits, preventing harms, and generating utility may outweigh or override moral considerations of individual liberty and justice when certain justificatory conditions of effectiveness, necessity, least infringement, and proportionality are met (Childress, Faden, Gaare, & Gostin, 2002). In instances when the benefits and burden are uneven, the expected benefit must be even greater (Kass, 2001).

Public health ethics is somewhat distinct from medical ethics. The latter has notably deeper historical roots. Medical ethics provides a set of principles to help guide physicians through moral challenges in the practice of medicine, including those that may arise between physician and patient. Tom Beauchamp and James Childress (2013) have outlined four ethical principles to guide medicine:

1. *Principle of respect for autonomy*—that is, values and loyalties are determined freely and voluntarily without coercion (e.g., informed consent).
2. *Principle of nonmaleficence*, which implies minimization of harm or injury.
3. *Principle of beneficence*, which refers to doing good by acting with the best interest of the other in mind. It implies a duty to balance benefits and harms and ensuring that benefits outweigh harms.
4. *Principle of justice*, which concerns fairness and distribution of common goods fairly. Thus, when goods or resources are limited, a means for fair distribution must be determined such that "those persons who are equal should qualify for equal treatment."

These principles are intended as nonhierarchical, such that no one principle is superior to another and each principle is considered *prima facie*.

While many frameworks for medical ethics have been put forward, frameworks for the ethics of *public health* work are a more recent contribution. While the foundations of both public health and medical ethics are similar, they also differ in some critical ways. Most notably, while both require the provision of benefit and the minimization of harm, public health targets this requirement to communities as a whole, and measures benefit the entire society. Medical ethics, by contrast, is more likely to focus exclusively on the well-being of individuals. Furthermore, public health is afforded legal authority in many environments to ensure that the *public's* health is improved and/or not threatened and, at times, can use invasive measures, as needed,

ranging from medical isolation and quarantine to mandatory hospitalization, as seen in multiresistant tuberculosis and extensively drug-resistant tuberculosis (XDR-TB), to secure that goal. Public health ethics, then, provides a framework for balancing the important and core value of public or societal benefit with restraint to ensure that individual rights and values are neither ignored nor compromised whenever possible. Public health ethics also puts emphasis on justice and equity, both as predictors of good public health outcomes and as a moral good for their own sake.

Public health ethics shares many core values with human rights but, again, is different in some critical ways. Public health ethics, at its core, is concerned with ensuring that public health interventions provide benefits and minimize harms; that they respect individuals' dignity and rights to the greatest extent possible; and that they are implemented fairly and, to a great degree, serve to increase equity. These values are completely consistent with those of human rights covenants and paradigms. Unlike human rights, however, ethics is not founded in law. While ethical norms may contribute the rationale for many laws, ethics itself has no legal standing, and "moral rightness" is not legally enforceable.

Appealing to ethics in crafting and implementing public health interventions nonetheless is critical for three reasons. First, engaging in morally right action is important in and of itself. That interactions, including interactions related to public health, ought to be implemented consistent with the highest ethical standards is itself an important end. Civil societies, by definition, are bounded by shared moral norms and practices (as well as by more formal laws and regulations), and those who are in leadership positions should uphold them.

Second, professional ethics requires a commitment to engaging *in one's work* in ways that are consistent with high ethical standards. Most of the established health professions, including medicine and nursing, have longstanding codes of professional ethics; public health more recently adopted its own code of ethics for public health professionals. Such professional codes help to self-regulate a profession (**EXHIBIT 3-4**). Acting in accordance with moral norms and codes is important to the integrity and trustworthiness of the professionals who practice public health and doing so also serves to instill trust in the profession on the part of the public.

EXHIBIT 3-4 Global Health, Ethics, and Research

Health research in low- and middle-income countries (LMICs), in pursuit of global health outcomes, raises complex ethical challenges. Research in such countries is vital to advancing knowledge on preventable disease, infectious disease epidemics, and treatments that have a global reach. However, ethical guidelines and regulations may be vastly inconsistent across different country settings due to limitations of capacity and infrastructure. Maintaining sound ethical guidelines in these settings is considered critical to safeguard against potential exploitation of research participants.

Amon et al. (2012) highlight the challenges that arise when investigating human rights violations as determinants of, or structural barriers to, health, and in particular when engaging with marginalized, stigmatized, and criminalized populations, especially when they face a disproportionate burden of ill health as a result of their social status. Amon et al. relate the experience of agency in health-based research in more repressive countries where a research ethics committee functions as an agent of the state, protecting state interests rather than legitimately representing, or protecting, the interests of vulnerable groups and research participants. This kind of environment can prevent health researchers from investigating health outcomes associated with state actors, government laws and policies, and social and cultural norms.

Other issues relate to consent in emergency settings and access to data, interventions, and treatments for the target population when research is complete. An example of the latter arose during HIV perinatal trials in the late 1990s when more inferior interventions, compared to the standard best available treatment, were employed (IJsselmuiden, Kass, Sewankambo, & Lavery, 2010). In this case, debate arose over whether use of an inferior regimen, considered more logistically or economically feasible to distribute and contributing to overall public health gains in a resource-constrained setting, should have precedence over the gold-standard intervention even though it would reach fewer people (IJsselmuiden et al., 2010).

Further examination of this trend was undertaken in a report by the Commission on Health Research for Development, which describes the vast differential between global health needs and global health research. In what is referred to as the 10/90 gap, less than 10% of global resources targeting global health research are allocated to health in low-income countries, which account for more than 90% of the global burden of preventable deaths, included neglected diseases such as Chagas disease, leishmaniasis, and human African trypanosomiasis. Conflict and political or social instability can also have profound effects on health interventions and epidemiologic surveillance efforts through disruption of health data collection and information systems, compromising the quality and sampling of core data.

(continues)

EXHIBIT 3-4 Global Health, Ethics, and Research *(continued)*

The Nuremberg Code formalized the ethical standards and principles for the conduct of medical ethics, particularly in research involving human participants. Formulated in 1947 in Nuremberg, Germany, following revelation of the murderous and tortuous conduct of Nazi doctors in World War II concentration camps, it serves as a blueprint for the rights of subjects in medical research (Shuster, 1997). The Nuremberg Code delineates core principles critical to clinical research, which are also broadly applicable to nonresearch contexts. These principles include informed voluntary consent given freely, without coercion, duress, or force, as well as the right for withdrawal, thereby extending the ethic to "first, do no harm" and protecting the rights of the individual. Article 7 of the International Covenant on Civil and Political Rights (United Nations, 1966) articulates this requirement by stating: "No one shall be subjected to torture or to cruel, inhuman or degrading treatment or punishment. In particular, no one shall be subjected with his free consent to medical or scientific experimentation."

A second foundational document concerning ethics is the Declaration of Helsinki. Introduced in 1964 and revised in 1975 by the World Medical Association, it outlines standards and protocols of research, including the requirement that the proposed research be reviewed by an independent research ethics committee. In 1979, the U.S. Department of Health and Human Services' Belmont Report provided further moral guidance by reiterating principles of respect for persons, beneficence, and justice as part of conduct of research using human subjects. These significant contributions continue to govern approaches to medical ethics and health research in humans. Research ethics committees are considered a widely accepted and fundamental component of conducting health-based research involving systematic collection of data and analysis of data on humans.

Third, as a means to an end, good ethics makes good public health sense. Ethics asks public health to ensure programs will be beneficial before imposing them on the public. It requires that individuals and communities be treated with dignity, and it requires that *all* communities receive appropriate public health interventions, not just those with more privilege or influence. Upholding these moral norms is relevant not just for their own sake: They make obvious sense for public health. Such principles help to ensure that more people get health benefit and that the "targets" of public health programs will better trust that such programs are initiated to further their own interests rather than to further some arbitrary goal on behalf of the state's leaders. Promotion of health equity has been shown though countless studies to improve the public's health. And where public health actions *do* need to infringe on civil liberties, such as in enforcing immunization requirements for public schools, it is critically important that targeted populations perceive that these programs are fair, that they are based in sound science, and that the privileged (by resources or power) are not exempted from the programs (Beyrer, 2004).

Several approaches to considering ethics in public health proposals have been put forward. One six-step approach is offered here (Kass, 2001).

Step 1. What are the public health goals for the proposed intervention, policy, or program? These goals generally should be expressed in terms of public health improvement, such as the degree to which the program will reduce morbidity or mortality. For example, a human immunodeficiency virus (HIV) screening program should have as its ultimate goal that fewer incident cases of HIV will occur, rather than that a certain proportion of individuals will agree to be tested.

Also relevant when considering public health goals and benefits is to whom the benefit will accrue. Public health interventions often are targeted to one set of individuals so as to protect *other* citizens' health. For example, partner notification programs and directly observed therapy for tuberculosis are designed, primarily, to protect citizens from the health threats posed by others. Restricting someone's liberty to protect that person—generally framed as paternalism—poses different ethical burdens than restricting liberty to protect the interests of others.

Step 2. How effective is the intervention or proposed program at achieving its stated goals? Proposed interventions or programs are based on certain assumptions that lead us to believe the programs will achieve their stated goals. Step 2 asks us whether actual *data* exist to support these assumptions. In general, the greater the burdens posed by a program—for example, in terms of cost, constraints on liberty, or targeting particular, already vulnerable segments of the population—the stronger the evidence must be to demonstrate program effectiveness. Indeed, because many public health programs are imposed on people by governments and not sought out by those targeted by the programs, the burden of proof lies with governments or public health practitioners to prove that a program will achieve its goals. If there are no good data to demonstrate program effectiveness, the analysis can stop right here, and, ethically, the program should not be implemented. Conversely, the presence of good data alone does not justify the program; it

simply allows public health practitioners to move to the next stage of the analysis.

Step 3. What are the known or potential burdens of the program? If data suggest that a program is reasonably likely to achieve its stated goals, the potential burdens or harms that could result from the public health work must be identified. The majority of such harms will fall into four broad categories: (1) risks to privacy and confidentiality, especially in data collection activities; (2) risks to liberty and self-determination, given the power accorded public health to enact almost any measure necessary to contain disease; (3) risks to health, if public health interventions carry some risk to the individuals affected; and (4) risks to justice, if public health practitioners propose targeting public health interventions only to certain groups. Data collection may not simply be viewed as a violation of personal privacy to individuals; breach of confidentiality of such data—deliberate or incidental—can lead to significant and tangible harms. Personal health information, such as that related to HIV status or sexual orientation and/or gender identity, can be dangerous in certain settings, if obtained by authorities and/or social acquaintances; even seemingly benign data such as vital statistics can reveal patterns about ethnic groups or neighborhoods that could lead to stigma, discrimination, violence, and/or forced relocation of identifiable groups.

Health education generally is thought of as the ideal public health intervention because it is completely voluntary and seeks to empower individuals to make their own decisions regarding their health. Unfortunately, education may not be effective in all settings. When it fails to meet the state's expectations, policy makers may feel the need to resort to more restrictive measures.

Regulations and legislation, strictly speaking, are coercive, since they impose penalties for noncompliance. As such, they pose risks to liberty and self-governance. While many such measures (e.g., mandatory immunizations) have demonstrated efficacy, they nonetheless are the most intrusive approach to public health. Certain mandated interventions, such as immunization or mass deworming campaigns, are implemented only when the population benefit is considerably larger than the risk to individuals, yet ethics tensions remain when even very small risks are created for healthy individuals because of public health campaigns. Further, the law can impose threats to justice if regulations pose an undue burden on particular segments of society, and the law can be designed in ways to reduce inequalities as well.

Step 4. How can burdens be minimized? Is the least burdensome approach being implemented? Once burdens have been identified, ethics requires programs to be modified in ways that minimize burdens without greatly reducing efficacy. If disease surveillance is equally effective with unique identifiers as with names, if voluntary programs yield almost identical cooperation and effectiveness as mandatory ones, or if individuals can be informed in advance about why interventions are being introduced, then these "burden-reducing" approaches should be taken.

Step 5. Is the program implemented fairly? Consistent with the principle of distributive justice, there must be a fair distribution of benefits and burdens in public health programs. Public health benefits such as clean water cannot be limited to one community alone, and equitable treatment is similarly required when restrictive measures are proposed. Injustice is wrong not just for its own sake, but for the material harms that can follow. This does not mean that programs or resources must be allocated equally or identically to all communities; rather, allocations must be *fair*. That is, differences cannot be proposed arbitrarily or based on historical assumptions about who might be at risk or who is more responsible. Instead, targeting of programs to one community and not another must be justified with strong attention to data. Moreover, the social consequences must be considered when targeting of programs occurs, and balanced against the benefits to that community or others. Also central is the role that public health can play in righting existing injustices, especially given the strong link between social inequities and poor health outcomes (Starfield, 2005). Several conceptions of justice allow and even require unequal allocation of benefits to remedy existing inequities (Daniels, 1985; Rawls, 1971).

Step 6. How can the public health benefits and the accompanying burdens be balanced? Even to the extent that public health professionals aim to follow the previously described requirements, disagreements invariably will emerge over interpretation—which types of freedoms must prevail, which types of burdens are acceptable, and which types of targeting are unjustified? Procedural justice, then, requires fair procedures to determine which public health interventions, in the end, should go forward. This process will require communities to discuss what is gained from good public health, and why such benefits often must be organized collectively. Dissent around proposed programs or interventions deserves special attention if raised by an identified subgroup, such as an ethnic minority, particular age group, or residents of a particular region. In general, the greater the burden imposed by a program, the greater must be the expected public health benefit. Likewise, the more uneven the benefits and burdens—that is, when one group is burdened to protect the health of others—the former must be both the scientific justification and the expected benefit.

Case Studies in Global Health, Human Rights, and Ethics

Obesity and Taxation on Sugar-Sweetened Beverages

Obesity and its associated conditions represent an epidemic of global proportions that does not discriminate by gender or socioeconomic status. While the rate of obesity is increasing globally, the absolute rates of its occurrence are higher with low income and low education. Most of the world's population now lives in countries where there are more deaths attributable to being overweight than underweight (Basu, McKee, Galea, & Stuckler, 2013; see also Development Initiatives, 2017).

The etiology of obesity is complex, with this condition resulting from an interplay of genetic, biological, environmental, sociopolitical, and behavioral factors (Malik, Pan, Willet, & Hu, 2013). Yet the rapid change in obesity rates globally suggests that the primary factors are environmental and behavioral rather than due to genetics, as the human genome could not have changed so rapidly. Traditional obesity prevention strategies, which have targeted individual behavior change, have yielded limited success. In more recent years, there has been growing interest in taking more of a public health approach to obesity, by intervening in the food environments in which people make their food consumption choices. Interventions related to the food environment can include changes in the production, availability, regulation, cost, and marketing of food and beverages.

One component of diet significantly implicated in the overweight and obesity epidemic globally is the high intake of dietary free sugars and, more specifically, the consumption of sugar-sweetened beverages (SSBs) (Brownell et al., 2009; Colagiuri, 2017). In the United States, for example, SSBs are the single largest contributor to adults' and children's daily caloric intake (Bleich, Wang, Wang, & Gortmaker, 2009), and increased consumption of SSBs has been associated with increased prevalence of obesity, cardiovascular disease, and diabetes (Brownell et al., 2009; Malik et al., 2013; Schulze et al., 2004; Te Morenga, Mallard, & Mann, 2013). In the United States, consumption of energy-dense foods in the form of sugary beverages accounts for almost half of the total daily intake of added sugar (Malik et al., 2013; see also Colagiuri, 2017). It has been estimated that 8.5 million disability-adjusted life years (DALYs) are related to SSB intake. The proportional mortality due to SSBs is highest among younger adults, with this factor contributing to more than 1 in 10 of all diabetes- and obesity-related deaths in nearly every region of the world; in Mexico, 30% of deaths in those younger than age 45 years are attributed to SSB consumption (Singh et al., 2015). Restriction or elimination of sugary beverages has been associated with positive changes in body weight. Sugar-sweetened beverages are less satiating, have a high caloric load, and are nutritionally deplete.

There is also compelling recent evidence that the food industry colluded with scientists in the 1960s and 1970s, when evidence began to appear on the harmful effects of sugar consumption, to focus on the purported harmful effects of dietary fat, rather than sugar (O'Connor, 2016). In a situation strikingly reminiscent of the tobacco industry's tactics, the sugar, corn-producing, and soft drink industries sought to avoid regulation of their products even as evidence of their harms emerged (Kearns, Schmidt, & Glantz, 2016). For example, industry-supported scientists at Harvard University, in a 1965 paper in the *New England Journal of Medicine*, deliberately underplayed the role of sugar in coronary artery disease, instead focusing their attention on the harms associated with fat and cholesterol (Kearns et al., 2016).

Public health efforts targeting obesity have traditionally relied on health education about good nutrition (including, for example, requirements for food labeling and public awareness of the "food pyramid" or "food plate"); individual-oriented strategies are underscored by the fact that the diet industry is a $60 billion business in the United States (Kass, Hecht, Paul, & Birnbach, 2014). In proposing public policy related to the food environment—whether related to marketing, pricing, taxation, or other influences—one must consider the ethical ramifications of alternative approaches and then choose the strategy that will yield the desired results while simultaneously posing the fewest threats to other ethically important values (Kass et al., 2014).

An ethics analysis of alternative policy options must first start by determining what the public health *goal* of a proposed policy is—for example, to decrease population obesity, rather than to reduce SSB consumption, which might simply lead to substitution with an equally harmful beverage. Second, one must determine how much evidence there is that the proposed policy will actually result in the intended public health benefit. Third, public health programs should constrain liberties as little as possible, and minimize the risk of other important harms or burdens. The fourth requirement is a justice consideration, requiring

interventions to avoid disproportionately burdening a particular population without important justification and to aim to, at least in part, reduce disparities in the population. Fifth, the introduction of a policy must follow fair procedures and should be accompanied by accountability measures. Finally, the symbolic relevance of public institutions must be preserved.

One public health intervention that has garnered particular attention as a response to the burgeoning obesity epidemic is taxation of SSBs—an approach that has been implemented in several cities in the United States and also in Mexico. Taxation is a form of public policy that allows governments to exert mild influence over products or practices they wish to encourage or discourage. For example, some jurisdictions have chosen to allow cigarettes and alcohol to remain completely legal, but have imposed higher levels of tax than the typical sales tax to provide a disincentive for their purchase (Kass et al., 2014).

One of the main ethical critiques of the SSB tax has been a liberty-based one: This view states that taxation is government overreach, and that individuals should be able to consume what they want and make their own choices. A second critique is a justice-based one, rightly citing that sales taxes are regressive, meaning that the impact on lower-income individuals is proportionally higher if they choose to buy SSBs than it is on wealthier individuals.

An ethics analysis would go deeper than these arguments, asking whether progressive taxation is required for products where there is not a government (public health) interest in ensuring access, and then asking which types of liberties governments are required to preserve. Most central to the answer would be the fundamental human rights—rights to freedom of religion, press, whom to love, and speech are most essential for governments to preserve. Claiming that there is a "liberty interest" in being able to buy and consume unhealthy products is not entirely irrelevant—the ability to pursue pleasurable activities is certainly important—but when the liberty is so deeply distant from the fundamental rights and liberties that must be preserved, more room for compromise is absolutely allowable. Indeed, the ability to still allow free access to SSBs while simply providing a financial disincentive for their purchase, similar to that instituted in many regions for cigarettes, is increasingly being viewed as both an ethically acceptable strategy and one with public policy acceptability.

While these critiques are important, an ethics analysis must go beyond claiming that there is a liberty infringement in taxing SSBs. Although a central responsibility of government is to protect foundational liberties from unwarranted intervention, it does not necessarily follow that fundamental liberties are threatened when public policy discourages consumption of unhealthy products or prohibits government spending on them (Kass et al., 2014). Also, and less often discussed in the context of liberty and taxation, too little state intervention in improving population health can violate individuals' right, just as too much can (Wilson, 2016). Limitations on health can threaten individuals' ability to pursue their life course and independent priorities. While SSBs remain widely available on the market, the intent of a tax is to act as a disincentive, rather than to prohibit their use outright. The personal pleasure to be derived from consumption of SSBs is absolutely worthy of consideration, yet such pleasure does not rise to the level of a fundamental freedom (Kass et al., 2014).

Invoking the right to health, and the concomitant obligations laid out in human rights frameworks, brings the protective function of the state and its role in obesity prevention through taxation of SSBs into sharp focus. Under the human rights provisions, states must take steps to enable full realization of the right to achieve the highest attainable standard of health. The state violates individuals' rights if it fails to take cost-effective and proportionate measures to remove health threats from the environment. For example, the right to adequate food implies enabling food security through access to nutritionally dense foods, responsible food labeling and appropriate food regulation, and a particular emphasis on advertising and marketing of foods and beverages to children and to schools. High intake of SSBs among the food-insecure and low-income populations is an important consideration, especially given that nutrient-dense foods are often more difficult to access or cost prohibitive for these populations.

HIV/AIDS

The HIV/AIDS pandemic remains one of the greatest challenges to human health, with more than 70 million being infected with the virus, and more than 35 million deaths due to AIDS-related causes having occurred by the end of 2016 (Joint United Nations Programme on HIV/AIDS, 2017). Since its inception, the HIV pandemic has been different from the pandemics of influenza, smallpox, or polio; all of those diseases were greatly feared, yet none generated the kinds of social opprobrium against the infected individuals that the HIV pandemic has so regularly created (Cohen et al., 2007). The scale of the HIV pandemic—and the stigma, discrimination, and violence that surrounded

its sudden emergence—catalyzed a public health response that expanded human rights in principle and practice and ignited ethical debates.

While the prospect of a cure and a preventive vaccine remain elusive, comprehensive and combination treatment/preventive programs have proved to be critical interventions to mitigate the impact of HIV infection and reduce HIV incidence in many settings and populations. Despite these advances, a large majority of individuals living with HIV or at risk of HIV still do not have access to prevention, treatment, and care and are unaware of their HIV status (Joint United Nations Programme on HIV/AIDS, 2017; WHO, 2014). The vast majority of people unable to access treatment reside in low- or middle-income countries and communities. In key affected populations,[3] identified as people who have both higher likelihood of HIV infection and greater risk of being excluded form essential HIV services, HIV incidence continues to rise, even as incidence stabilizes or declines in the general population (WHO, 2014). Outside of sub-Saharan Africa, 80% of all new HIV infections among adults occur among people from key populations and their immediate sexual partners (Joint United Nations Programme on HIV/AIDS, 2017).

Key populations, collectively at the global level, are disproportionately affected as a result of the additional social and structural factors that increase their vulnerability. Stigma is recognized as a powerful social determinant of health and a key driver of health disparities in vulnerable groups (Poteat et al., 2015). In many instances, this vulnerability increases when individuals are unable to realize their rights and face stigma, exclusion, harassment, and violence as a result of both their HIV status and their membership in a key population. When stigma and discrimination in healthcare settings are both overt and hidden, they may lead to delayed HIV testing, concealment of positive serostatus, and poor uptake of HIV services (Fay et al., 2011).

In 2017, 78 countries had repressive laws criminalizing homosexuality or propaganda of homosexuality; in 13 of these countries (e.g., Sudan, Iran, Nigeria, Somalia), homosexual acts are punishable by death (Carroll, 2016). In an overwhelming number of countries that criminalize homosexuality, HIV prevalence is at the highest levels, although the absence of such laws in other countries does not necessarily mean they have lower rates of HIV. Criminalization sanctions and reinforces existing prejudices and legitimizes violence; in practical terms, it translates into increased barriers to equitable access to essential services and treatment. Enforcement of antidiscrimination and protection laws invokes core human rights standards and is a key component toward advancing the health of the most vulnerable. Protecting the human rights of the most vulnerable populations necessarily includes efforts to decriminalize sexual behaviors, recognize diverse gender expression and identity, and provide inclusive and sensitized services.

Against this background, this section will explore in detail the human rights and ethical dimensions of the global health challenge that is HIV. The focus will be on two public health approaches that have defined the response: (1) HIV prevention and (2) equity of and access to HIV treatment. The use of human rights and ethics tools applied to the HIV context highlights the integrative and complementary nature of these perspectives, as well as important differences between them.

Prevention

From early on in the pandemic, tension between individual human rights and the role of the state as duty bearer drove the response. Prominence was given to individual human rights, with a focus on risk, behavior change, and agency through campaigns on access to care, information, HIV testing and counseling, and use of condoms. However, population-based policies, such as universal access to condoms and needle syringe programs, that prioritized access to prevention tools were often selectively implemented in different countries, leaving large networks within many populations at significant risk and lacking the basics of HIV prevention services.

The emergence of newer preventive paradigms, including treatment as prevention, has reframed prevention efforts to have an increased focus on the integration of individual rights into collective public health policies. Strategies such as preventing mother-to-child transmission, harm reduction, pre-exposure prophylaxis (PrEP), post-exposure prophylaxis (PEP), prevention of violence and decriminalization, and provision of opiate substitution medication for people who inject opiates have had a profound impact by stemming the acquisition and transmission of HIV. To further explore HIV prevention and the application of human rights and ethics frameworks, the section that follows highlights two examples: Harm reduction and treatment as prevention.

3 Key affected populations are defined as people who have been identified as belonging to population groups most at risk of HIV. They include men who have sex with men (MSM), transgender people, people who inject drugs, and sex workers. Vulnerability extends to groups of people who have an increased susceptibility to HIV infection as a result of their specific circumstance or context—for example, adolescents, people in closed settings, orphans, street children, and migrant and mobile workers.

People who inject drugs are at risk of HIV and other blood-borne viral infections, including hepatitis B and C, through sharing of contaminated, nonsterile syringes and injecting equipment. Globally, 123 countries—that is, 78% of the total 158 countries that report data on injecting drug use—attribute cases of HIV to injecting drug use (WHO, 2014). In 2012, it was estimated that approximately 12.7 million (range: 8.9–22.4 million) people on a worldwide basis had recently injected drugs; of those, 1.7 million people (13.1%) were living with HIV. Data from 49 countries show that the risk of HIV infection was, on average, 22 times greater among people who inject drugs than among the general population (WHO, 2014).

Injecting-drug users endure institutionalized vulnerabilities as a result of arbitrary deprivation of rights in the form of mandatory testing, inadequately targeted services, social marginalization, and punitive measures to contain drug dependence through arrest and imprisonment. Indeed, the most elemental concerns of the human rights agenda are also the determinants of the health outcomes for injecting-drug users, who face the prospects of incarceration, violence, stigmatization, isolation, and discrimination (Wolfe & Cohen, 2010). As Wolfe and Cohen (2010) note, core principles of human rights—including liberty and security of the person, autonomy, privacy, and freedom from cruel, inhuman, or degrading treatment—are vital components of effective health programs for this population group.

In the few countries that do not criminalize drug use, political and legal barriers prevent access to prevention without discrimination, and to justice such as freedom from arbitrary arrest and detention (WHO, 2014; Wolfe & Cohen, 2010). Examples can be cited of compulsory treatment centers, ostensibly used to rehabilitate or detoxify, that hold people without charge, right of appeal, or evaluation by a health professional, addiction specialist, or psychiatric specialist. In other cases, individual liberty has been usurped by law enforcement as a means of containment, further impeding universal access and public health efforts. Interned in prisons, injecting-drug users may be exposed to infectious diseases, violence, overcrowding, and high-risk behaviors. Access to key treatments, such as antiretroviral therapy, may also be limited due to ingrained stigma and perceptions of instability that may jeopardize adherence. A more recent and pressing example is the "war on drugs" in the Philippines, which highlights the difficulty of globalizing ethical, moral, and evidence-based public health interventions in highly criminalized settings where extrajudicial killings are being actively encouraged and enjoy broad popularity in the populace. This generates a broader discussion of how human rights and ethics may be applied in these settings.

Treatment as Prevention

Results of a landmark clinical trial in 2011 demonstrated that early treatment of the HIV-infected partner in sero-discordant couples with antiretroviral therapy provides durable and reliable protection against sexual transmission to the uninfected partner. This intervention was found to 96% effective in decreasing the risk of HIV acquisition (Cohen et al., 2011). This trial, named HIV Prevention Trials Network (HPTN) 052, supported the use of medical therapy as a public health prevention strategy, thereby changing the landscape for HIV prevention.

These findings have been further supported by the landmark "Opposites Attract" study of HIV sero-discordant male same-sex couples, led by Andrew Grulich and colleagues, which showed that men living with HIV and virally suppressed had no documented HIV transmission to uninfected partners despite low use of condoms (Bavinton et al., 2014). The study findings suggest that when treatment as prevention is at its most optimal, an HIV-positive individual with an undetectable viral load has a negligible risk of transmission.

This situation changes issues of disclosure and raises questions about whether there remains a moral requirement to disclose HIV status if all reasonable attempts to reduce risk are taken. Similarly, if the viral load levels are undetectable to the extent that the risk of transmission is practically negligible, is there an obligation to disclose (Haire & Kaldor, 2015)? Indeed, the question may be whether the obligation to prevent viral transmission lies solely with the individual who is infected, or whether all parties have a collective responsibility to protect and prevent this disease (Sugarman, 2013). Is it morally wrong to situate the health of an infected person as secondary to the public health benefits of a suppressed viral load, particularly when the treatment may carry side effects or toxicities (Haire & Kaldor, 2015)?

Additional issues arise for key population groups. Despite the evidence, in a number of instances, trials and rollouts of this strategy are being hampered in countries where repressive laws exist and criminalization of homosexuality is pervasive and, in some cases, where there has been an expansion of the HIV epidemic in many MSM populations (Beyrer et al., 2016). Treatment as prevention has been demonstrated to have limited utility in other key population groups,

such as people who inject drugs, particularly in circumstances where other interventions such as oral substitution therapy are unavailable.

Treatment: Equity and Access

The emergence of treatment options was a critical turning point in the HIV response, with dramatic reductions in morbidity and mortality following in their wake. The potential benefits of earlier initiation of therapy far outweigh the potential risks of increased exposure to drug toxicity and emergence of viral resistance in the setting of suboptimal adherence.

Universal access to medicine has the promise to deliver much-needed treatment to those most at need and to extend life expectancy. Nevertheless, identifying those who are infected shortly after acquisition and initiating treatment remain a challenge to the arresting the epidemic. Treatment is reliant on individuals knowing their status and having access to noncoercive testing.

In 2016, WHO released consolidated guidelines recommending that all people living with HIV be provided with antiretroviral therapy, removing limitations on eligibility for HIV-positive individuals. Expanded access to treatment resonates with the 90-90-90 targets—that 90% of people living with HIV have knowledge of their HIV status, 90% of people who know their status have access to and receive ART, and 90% of people who receive ART have suppressed viral loads. Despite this concordance, even the most ambitious targets for expanding access to antiretroviral therapy recognize that many people in need of treatment will not receive it. Utilitarian arguments as applied to access to medicines promote an ethics of resignation and the observance that resource scarcity is accepted as inevitable, and the pressure to identify and address inequality is diminished by the dissemination of those scarce resources within a population (Smith, 2016). Populations and groups who have been systematically excluded and disenfranchised from treatment programs, such as migrants and ethnic minorities, and stigmatized groups including sex workers, injection-drug users, prisoners, and sexual and gender minorities, continue to face difficulty in obtaining treatment access. The issue of equity in treatment access is complicated by the fact that in some countries, the overwhelming majority of people in need of treatment come from politically marginalized or vulnerable groups. Perhaps the clearest example of this is injection-drug users, who are consistently under-represented in national treatment programs despite accounting for the overwhelming majority of people in need of treatment in many countries. The reason for such under-representation in treatment access may be an underlying pattern of human rights abuse that renders certain populations less able to obtain basic health care.

Under human rights law, the guiding principle for equitable access to antiretroviral treatment is that of nondiscrimination and equality under the law. This principle accepts that governments have difficult choices to make and that not every "social good" can be made universally available. Undertaking a commitment to do the greatest good for the greatest number of people may be at odds with treatment paradigms. Expanding access to medicines that are known to be less efficacious or carry a greater toxicity profile, but that are inexpensive and, therefore, can reach a greater number of people, including those who may otherwise not have access to treatment, may be more beneficial or preferable than provision of the gold-standard treatment in all settings, with the understanding that this resource may be finite and available to fewer people (Persad & Emanuel, 2016). The former option challenges the idea that refusal to offer treatment, if deemed substandard or potentially more harmful, overemphasizes nonmaleficence to the extent that it is a dereliction of duty (Persad & Emanuel, 2016). The latter position reinforces the notion that equal access to the best available treatment for all people should be the primary goal (Smith, 2016).

In international covenants, governments are generally prohibited from intentionally or unintentionally denying social benefits to individuals on the basis of, among other things, race, sex, national or ethnic origin, religion, and political viewpoint. Sometimes such denial will be justified, as when pregnant women are given preference for treatment to prevent HIV transmission from parent to child. Nevertheless, such decisions should never be based on unfounded or stereotypical assumptions about marginalized groups—such as the assumption that injection-drug users should not qualify for treatment programs because they are "noncompliant" or incapable of adhering to a treatment regimen. In the case of antiretroviral treatment, their eligibility should be based on clinically relevant criteria and, beyond that, on criteria that are justified in the circumstances and that do not offend human dignity (**EXHIBIT 3-5**).

The procedures by which these criteria for treatment access are set should likewise be subject to both human rights and ethics standards. International human rights law recognizes the "right to participate," which should include a positive obligation on governments to solicit the views of affected populations in formulating public policy (International Covenant on Civil and Political Rights, 1966). More urgently,

EXHIBIT 3-5 Test and Treat

While it has been estimated that 53% of all people diagnosed with HIV are on antiretroviral therapy (Joint United Nations Programme on HIV/AIDS, 2017), approximately 30% of people living with HIV at a global level remain undiagnosed and unaware of their seropositive status (WHO, 2017). In resource-poor settings, roughly half (19.5 million) of those individuals eligible for treatment are receiving treatment (Joint United Nations Programme on HIV/AIDS, 2017). The evidence consistently shows that stigma and discrimination in healthcare settings are both overt and hidden, leading to delayed HIV testing, concealment of positive serostatus, and poor uptake of HIV services.

"Test and treat" is a strategy that universalizes voluntary counseling and testing and offers immediate treatment to individuals who test positive, irrespective of clinical stage or CD4 count. It is premised on the notion that if individuals are made aware of their status, they will access treatment and alter high-risk behaviors.

Test and treat is a proven intervention in settings where the HIV epidemic is not widespread or generalized. Even so, some challenges remain. Despite this strategy's potential to recruit newly diagnosed persons into treatment, to preserve autonomy, and to normalize and destigmatize HIV, barriers remain in regard to linkage to care once the diagnosis is known as well as persistent issues related to viral resistance and adherence.

authoritarian governments that place restrictions on civil society, including crackdowns on AIDS activism and censorship of the press, risk adopting HIV/AIDS policies that do not reflect the needs of their population. These "first generation" human rights guarantees (freedom of speech, freedom of the press, freedom of association) are essential to treatment rollouts to the extent that they foster the political participation of the widest range of stakeholders possible. Procedural justice, as mandated by ethics in creating significant public health policy, has such rights as its cornerstone.

Conclusions

Public health, human rights, and ethics have universal and deeply shared values. It is the mission of public health to improve health, reduce morbidity, and reduce mortality wherever it does its work. Human rights and ethics, too, have basic sets of principles and rules in common that are intended to guide and/or dictate behavior in a variety of situations so as to ensure that human rights and ethics norms are not compromised in the pursuit of good public health outcomes.

While the principles and values of human rights and public health ethics generally are shared, a significant difference between them remains in terms of targets and redress. Ethicists generally have little legal power to challenge what may be viewed as unethical practices or programs, even where the targets of such ethical critiques or advocacy are governments or publicly sanctioned policies. Human rights activists, in contrast, explicitly target governments or the policies they endorse; international tribunals, war crimes trials, and the like are among the most potent tools wielded by human rights advocates. Thus, whereas human rights organizations explicitly try to challenge rights violations through existing legal systems, ethical frameworks try to shape shared societal norms for morally appropriate behavior—norms that, in turn, may be reflected in the law. In addition, human rights proponents often channel their arguments through media or advocacy, whereas in ethics this is less likely to be the primary means of communicating.

In certain contexts, the goals of public health, ethics, and human rights can be very well aligned. Public health, as a branch of government, has extraordinary power to further the public's health in ways that are beneficial, careful, and fair. But because such significant power, wherever it is given, can so easily be abused, both ethics and human rights have created their own sets of checks and balances. Ethics and human rights provide the moral and legal "brakes" to redirect public health to more constructive tactics, and to highlight to public health professionals, through advocacy, argumentation, and accountability, what are and are not justifiable uses of state power and intervention in the name of furthering public health. Ethics and human rights also advocate proactively for just structures for public health—for creating the conditions under which individuals and communities can thrive.

It has been the thesis of this chapter that the human rights context in which public health work is conducted has an extraordinary impact on which public health tools ultimately must be selected to have ethically sound public health responses. The relationship of citizens to their governments has a tremendous impact on public health status in different locales; indeed, the degree to which governments believe that they have a responsibility to care for the health of their public varies strikingly from country

to country. Similarly, when public health seeks to intervene, the human rights or political context into which it enters will influence to a great degree both the potential public health benefits of a given intervention and the ways in which a given intervention is deemed ethically acceptable, ethically unacceptable, or ethically required.

Discussion Questions

1. How might the sale of high-sugar soft drinks in public schools be framed as a human rights issue?
2. What role do states have in regulating foods known to be damaging to health but popular with citizens?
3. In country X, HIV rates are disproportionately high among sexual- and gender-minority adolescents and adults. Same-sex behavior between consenting adults is criminalized. How might decriminalization of these behaviors affect HIV programs? How might this be addressed from a human rights perspective?
4. Adults have a right to purchase and consume the foods and beverages they like and want. Yet aggressive marketing of unhealthy foods can disproportionally affect the health and well-being of the poor, the marginalized, and those with less access to quality health care. How does a public health ethics framework help address the balance of freedom and responsibility? Choice and exploitation?

References

Amon, J. J., Baral, S., Beyrer, C., & Kass N. (2012). Human rights research and ethics review: Protecting individuals or protecting the state? *PLoS Medicine, 9*(10), e1001325.

Basu, S., McKee, M., Galea, G., & Stuckler, D. (2013). Relationship of soft drink consumption to global overweight, obesity, and diabetes: A cross-national analysis of 75 countries. *American Journal of Public Health, 103*(11), 2071–2077.

Bavinton, B. R., Jin, F., Prestage, G., Zablotska, I., Koelsch, K. K., Phanuphak, N., . . . Grulich, A. E.; Opposites Attract Study Group. (2014). The Opposites Attract study of viral load, HIV treatment and HIV transmission in serodiscordant homosexual male couples: Design and methods. *BMC Public Health, 14*, 917.

Beauchamp, T. L., & Childress, J. F. (2013). *Principles of biomedical ethics* (7th ed.). Oxford: Oxford University Press.

Beyrer, C. (2004). Public health, human rights, and the beneficence of states. *Human Rights Reviews, 5*(1), 28–33.

Beyrer, C. (2007a). Human rights and the health of populations. In C. Beyrer & H. F. Pizer (Eds.), *Public health and human rights* (pp. 1–10). Baltimore, MD: Johns Hopkins University Press.

Beyrer, C. (2007b). Responding to epidemic disease threats in Burma and lessons for China: Why good governance matters. In S. R. Quah (Ed.), *Crisis preparedness: Asia and the global governance of epidemics* (pp. 47–62). Stanford, CA: Stanford University Walter H. Shorenstein Asia-Pacific Research Center, and Brookings Institution.

Beyrer, C., Baral, S. D., Collins, C., Richardson, E. T., Sullivan, P. S., Sanchez, J., . . . Mayer, K. H. (2016). The global response to HIV in men who have sex with men. *Lancet, 388*(10040), 198–206.

Bleich, S. N., Wang, Y. C., Wang, Y., & Gortmaker, S. L. (2009). Increasing consumption of sugar-sweetened beverages among US adults: 1988–1994 to 1999–2004. *American Journal of Clinical Nutrition, 89*(1), 372–381.

Braveman, P., & Gottlieb, L. (2014). The social determinant of health: It's time to consider the causes of the causes. *Public Health Reports, 129*(suppl), 19–31.

Brownell, K., Farley, T., Willett, W., Popkin, B., Chaloupka, F. J., Thompson, J. W., & Ludwig, D. S. (2009). The public health and economic benefits of taxing sugar-sweetened beverages. *New England Journal of Medicine, 361*(16), 1599–1605.

Calain, P., & Poncin, M. (2015). Reaching out to Ebola victims: Coercion, persuasion or an appeal for self-sacrifice. *Social Science and Medicine, 147,* 126–133.

Carroll, A. (2016). *State sponsored homophobia 2016: A world survey of sexual orientation laws: criminalization, protection and recognition.* Geneva, Switzerland: International Lesbian, Gay, Bisexual, Trans and Intersex Association.

Childress, J. F., Faden, R. R., Gaare, R. D., & Gostin, L. O. (2002). Public health ethics: Mapping the terrain. *Journal of Law, Medicine and Ethics, 30*(2), 170–178.

Cohen, J., Kass, N., & Beyrer, C. (2007). Human rights and public health ethics: Responding to the global HIV/AIDS pandemic. In C. Beyrer & H. F. Pizer (Eds.), *Public health and human rights* (pp. 362–390). Baltimore, MD: Johns Hopkins University Press.

Cohen, M. S., Chen, Y. Q., McCauley, M., Gamble, T., Hosseinipour, M. C., Kumarasamy, N., . . . Fleming, T.; HPTN 052 Study Team. (2011). Prevention of HIV-1 infection with early antiretroviral therapy. *New England Journal of Medicine*, *365*, 493–505.

Colagiuri, S. (2017). The obesity epidemic and sugar sweetened beverages: A taxing time. *Medical Journal of Australia, 206*(3), 109–111.

Daniels, N. (1985). *Just health care.* Cambridge, UK: Cambridge University Press.

Decker, M. R., Crago, A. L., Chu, S. K., Sherman, S. G., Seshu, M. S., Buthelezi, K., . . . Beyrer, C. (2015). Human rights violations against sex workers: Burden and effect on HIV. *Lancet, 385*(9963), 186–199.

Development Initiatives. (2017). *Global nutrition report 2017: Nourishing the SDGs.* Bristol, UK: Author.

Diniz, D. (2016). The protection of women's fundamental rights violated by the Zika virus epidemic. *American Journal of Public Health, 106*(8), e9.

Dodgson, R., Lee, K., & Drager, N. (2002). Global health governance: A conceptual review. World Health Organization and the London School of Hygiene and Tropical Medicine. Retrieved from http://libdoc.who.int/publications/2002/a85727_eng.pdf

Donnelley, P. (2014). Ebola and human rights: Insights from the experts. *Health and Human Rights Journal*. Retrieved from

https://www.hhrjournal.org/2014/11/ebola-and-human-rights-insight-from-experts/

Editorial. Polio eradication: The CIA and their unintended victims. (2014). *Lancet, 383*(9932), 1862.

Farmer, P. (2014). Diary. *London Review of Books, 36*(20), 38–39.

Fay, H., Baral, S. D., Trapence, G., Motimedi, F., Umar, E., Iipinge, S., . . . Beyrer, C. (2011). Stigma, health care access, and HIV knowledge among men who have sex with men in Malawi, Namibia, and Botswana. *AIDS Behavior, 15*(6), 1088–1097.

Friedman, E., & Gostin, L. (2015). Imagining global health with justice: In defense of the right to health. *Health Care Analysis, 23*(4), 308–329.

Gable, L., & Meier, B. M. (2013). Global health rights: Employing human rights to develop and implement the framework convention on global health. *Health and Human Rights, 15*(1), 17–31.

Gostin, L. O. (2001). Public health, ethics, and human rights: A tribute to the late Jonathan Mann. *Journal of Law & Medical Ethics, 29*(2), 121–130.

Gostin, L., & Ayala, A. (2017). Global health security in an era of explosive pandemic potential. *Journal of National Security, Law and Policy, 9*, 1–24.

Gostin, L. O., & Mok, E. (2009). Grand challenges in global health governance. *British Medical Bulletin, 90,* 7–18.

Guttmacher Institute. (2017). *Abortion in Latin America and the Caribbean: Fact sheet.* New York, NY: Author.

Haire, B., & Kaldor, J. (2015). HIV transmission law in the age of treatment-as-prevention. *Journal of Medical Ethics, 41,* 982–986.

IJsselmuiden, C. B., Kass, N. E., Sewankambo, N. K., & Lavery, J. V. (2010). Evolving values in ethics and global health research, *Global Public Health, 5*(2), 154–163.

Institute of Medicine. (1988). *The future of the public's health in the 21st century.* Washington, DC: National Academies Press

International Covenant on Civil and Political Rights (INCCPR). (1966). G.A. res. 2200A (XXI), 21 U.N. GAOR Supp (No.16) at 52, U.N. Doc. A/6316 article 19.

International Covenant on Economic, Social and Cultural Rights (ICESCR). (1966). G.A. Res. 2200 (XXI). 21 U.N. GAOR Supp (No 16) at 49, U.N. Doc. A/6316, article 12.

Joint United Nations Programme on HIV/AIDS. (2017). *UNAIDS data 2017.* Geneva, Switzerland: UNAIDS.

Jürgens, R., Csete, J., Amon, J. J., Baral, S., & Beyrer, C. (2010). People who use drugs, HIV, and human rights. *Lancet, 376*(9739): 475–485.

Kass, N. (2001). An ethics framework for public health. *American Journal of Public Health*, *91*(11), 1776–1782.

Kass, N., Hecht, K., Paul, A., & Birnbach, K. (2014). Ethics and obesity prevention: Ethical considerations in 3 approaches to reducing consumption of sugar-sweetened beverages. *American Journal of Public Health, 105*(5), 787–795.

Kearns, C. E., Schmidt, L. A., & Glantz, S. A. (2016). Sugar industry and coronary heart disease research. *JAMA Internal Medicine, 176*(11), 1680–1685.

Koplan, J. P., Bond, T. C., Merson, M. H., Reddy, K. S., Rodriguez, M. H., & Sewankambo, N. K. (2009). Towards a common definition of global health. *Lancet, 373*(9679), 1993–1995.

Malik, V. S., Pan, A., Willet, W. C., & Hu, F. B. (2013). Sugar-sweetened beverages and weight gain in children and adults: A systematic review and meta-analysis. *American Journal of Clinical Nutrition, 98*(4), 1084–1102.

Mann, J. M. (1997). Medicine and public health, ethics and human rights. *Hastings Center Report, 27*(3).

O'Connor, A. (2016, September 12). How the sugar industry shifted blame to fat. *New York Times*. Retrieved from https://www.nytimes.com/2016/09/13/well/eat/how-the-sugar-industry-shifted-blame-to-fat.html

Office of the High Commissioner for Human Rights. (2000). Committee on Economic, Social and Cultural Rights (CESCR), General Comment No. 14, U.N. Doc. E/C.12/2000/4.

Office of the United Nations Commissioner for Human Rights. (2008). *The right to health.* Fact Sheet No 31. Geneva, Switzerland: United Nations.

Persad, G. C., & Emanuel, E. J. (2016). The ethics of expanding access to cheaper, less effective treatments. *Lancet, 388,* 932–934.

Phelan, A., & Gostin, L. O. (2016). The human rights dimensions of Zika. Retrieved from https://academicimpact.un.org/content/human-rights-dimensions-zika

Poteat, T. C., Logie, C. H., Adams, D., Mothopeng, T., Lebona, J., Letsie, P., & Baral, S. (2015). Stigma, sexual health and human rights among women who have sex with women in Lesotho. *Reproductive Health Matters, 23*(46), 107–116.

Rasanathan, J. J. K., MacCarthy, S., Diniz, D., Torreele, E., & Gruskin, S. (2017). Engaging human rights in the response to the evolving Zika virus epidemic. *American Journal of Public Health, 107*(4), 525–531.

Rawls, J. (1971). *A theory of justice.* Cambridge, MA: Belknap Press of Harvard University Press.

Richardson, E. T., Bailor Barrie, M., Kelly, D., Dibba, Y., Koedoyoma, S., & Farmer, P. E. (2016). Biosocial approaches to the 2013–2016 Ebola pandemic. *Health and Human Rights Journal, 18*(1), 115–128.

Roa, M. (2016). Zika virus outbreak: Reproductive health and rights in Latin America. *Lancet, 387*(10021), 843.

Rothstein, M. (2002). Rethinking the meaning of public health. *Journal of Law, Medicine and Ethics, 30*(2), 144–149.

Rubenstein, L. S., Amon, J. J., McLemore, M., Eba, P., Dolan, K., Lines, R., & Beyrer, C. (2016). HIV, prisoners, and human rights. *Lancet, S0140-6736*(16), 30663.

Sands, P., Mundaca-Shah, C., & Dzau, V. J. (2016). The neglected dimension of global security: A framework for countering infectious disease crises. *New England Journal of Medicine, 374*(13), 1281–1287.

Schulze, M. B., Manson, J. E., Ludwig, D. S., Colditz, G. A., Stampfer, M. J., Willett, W. C., & Hu, F. B. (2004). Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. *Journal of the American Medical Association, 292*(8), 927–934.

Shuster, E. (1997). Fifty years later: The significance of the Nuremberg Code. *New England Journal of Medicine, 337*, 1436–1440.

Singh, G. M., Micha, R., Khatibzadeh, S., Lim, S., Ezzati, M., & Mozaffarian, D.; on behalf of the Global Burden of Diseases Nutrition and Chronic Disease Expert Group (NutriCoDE). (2015). Estimated global, regional, and national disease burdens related to sugar-sweetened beverage consumption in 2010. *Circulation,132*(8), 639–666.

Smith, J. (2016). Dilemmas in access to medicines: A humanitarian perspective. *Lancet, 389*, 1007–1008.

Sollom, R., Richards, A. K., Parmar, P., Mullany, L. C., Lian, S. B., Iacopino, V., & Beyrer, C. (2011). Health and human rights in Chin State, Western Burma: A population-based assessment using multi-staged household cluster sampling. *PLoS Medicine, 8*(2), e1001007.

Stapleton, G., Schroder-Back, P., Laser, U., Meershoek, A., & Popa, D. (2014). Global health ethics: An introduction to prominent

theories and relevant topics. *Global Health Action, 7*(23569), 1–7.

Starfield, B. (2005). Equity, social determinants, and children's rights: Coming to grips with the challenges. *Ambulatory Pediatrics, 5*(3), 134–137.

Sugarman, J. (2013). HIV cure research: Expanding ethical considerations. *Annals of Internal Medicine, 159*(7), 490–491.

Tasioulas, J., & Vayena, E. (2015). Getting human rights right in global health policy. *Lancet, 385*, e42–e44.

Te Morenga, L., Mallard, S., & Mann, J. (2012). Dietary sugars and body weight: Systematic review and meta-analyses of randomised controlled trials and cohort studies. *British Medical Journal, 346,* e7492.

United Nations. (1948). Universal Declaration of Human Rights. General Assembly resolution 217.

United Nations. (1966). International Covenant of Economic, Social, and Cultural Rights. General Assembly resolution 2200A.

Wilson, J. (2016). The right to public health. *Journal of Medical Ethics, 42*, 367–375.

Wolfe, D., & Cohen, J. (2010). Human rights and HIV prevention, treatment, and care for people who inject drugs: Key principles and research needs. *Journal of Acquired Immune Deficiency Syndrome, 55*(suppl), S56.

World Health Organization (WHO). (2002). *Questions and answers on health and human rights.* Health and Human Rights Publication Series Issue No 1. Geneva, Switzerland: Author.

World Health Organization (WHO). (2014). *Consolidated guidelines on HIV prevention, diagnosis, treatment and care for key populations.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2017, November). *HIV /AIDS fact sheet.* Geneva, Switzerland: Author. Retrieved from http://www.who.int/mediacentre/factsheets/fs360/en/

CHAPTER 4

Understanding and Acting on Social Determinants of Health and Health Equity

Rene Loewenson and Sarah Simpson

Introduction

A large and growing body of evidence shows that health improves when action is taken on the underlying causes of ill health, and particularly those causes that lead to social differences in ill health. "These underlying causes are complex, often reflecting systematic social, political, historical, economic and environmental factors that accumulate across people's lifetimes and are transferred across generations" (Sadana et al., 2014, p. 8). Such underlying social, economic, political, legal, and material factors that affect health are collectively called *social determinants of health* (SDH). Beyond biological processes in the human body and the disease-causing germs that are determinants of health, health outcomes have social origins, such as in the way people live and work and their economic and political systems, known as the social determinants of health. Identifying them enables us to understand and intervene to improve health, and has been a longstanding goal of public health measures. For example, such interventions may include measures to counter the marketing of tobacco to reduce smoking, or creation of cycling paths and green spaces in urban areas to promote physical activity.

These SDH have a general impact on health outcomes. With the global understanding that everyone has the right to the highest attainable standard of health (United Nations, 1976), there is also increasing international attention on social differences in health that are avoidable and unfair (termed "health inequities"). Beyond the broad understanding of the social factors that cause disease, the persistent and sometimes growing differences in health within and between social groups and countries raises questions about how SDH are associated with avoidable differences in the opportunities for a healthy life as well as the differential benefits from interventions for improved health. When this analytic lens is applied, those factors may be termed the *social determinants of health equity* (SDHE). This chapter explores the understanding and application of SDH to improve health, and, as a matter of increasing importance globally, to improve health equity.

DEFINITIONS

Equity is the absence of unfair, avoidable, or remediable differences among groups of people, whether those groups are defined socially, economically, demographically, or geographically (Regional Network on Equity in Health in East and Southern Africa [EQUINET], 2012).

Health inequities are systematic differences in health that can be avoided or remedied and that are therefore viewed as unfair or unjust (Commission on the Social Determinants of Health [CSDH], 2008). This is particularly so for health given the global context, in which the highest attainable standard of health is one of the fundamental rights of every human being (World Health Organization [WHO], 2006).

- *Health inequality* refers simply to differences in health between different individuals, without a normative judgment about those inequalities.

In a chapter in an earlier edition of this text, Kelly and Doohan (2014) explored in some detail how the understanding of social variations in ill health and its prevention developed over time. They described the causal pathways from SDH to individual disease outcomes and population-level patterns of disease.

This chapter does not seek to repeat their well-presented explanations of the causal pathways from SDH to health outcomes. Rather, it focuses on the *application* of the concepts, with examples from a range of countries globally, particularly from low- and middle-income countries (LMICs), especially those in east and southern Africa. The conceptual frameworks, discussed later, show how SDH can be organized into different and deepening levels, each of which incorporates different elements. We do not aim to discuss all the diverse SDH in this chapter, but instead present more detail on some of these elements, given their role in improving health and health equity.

Thus, this chapter describes published work on the following topics:

- The development of conceptual frameworks for understanding SDH and SDHE, including how these frameworks relate to frameworks for human rights and gender equality
- Intersectoral action for health (IAH) and health in all policies (HiAP) as approaches used to apply the conceptual understanding of different types and levels of SDH, from immediate material determinants to deeper socioeconomic and policy determinants, to improve health and health equity
- The implications of an SDH perspective for health systems and services, understanding the health system itself as a social determinant of health and health equity
- The roles of and interventions to address social exclusion, social agency, and power as cross-cutting SDH, affecting other sociopolitical, socioeconomic, and material determinants
- The increasingly global dimensions of SDH in a globalizing world, their role in national-level SDH, and the different levels of responses to them
- How to evaluate action on SDH aimed at improving health and health equity

Conceptual Frameworks for Understanding Social Determinants of Health and Health Equity

Analysis of the relationship between disease and the social and material environment dates back to ancient medico-philosophical systems. With advancing understanding of the causes of ill health and their origin in human activity, it has become clearer that socially determined health risks can be prevented, including as a matter of social justice (Kelly & Doohan, 2014). In recent decades, as public health has shifted its focus from individual risk–health relationships to more complex, multifactorial causal networks, a range of increasingly comprehensive conceptual frameworks have emerged, drawing on evidence to support analysis of the SDH and their role in health equity. This section explores how these conceptual frameworks have developed, with increasing attention over time to the relationship between SDH and health equity.

Turrell et al. (1999) identified SDH at three discrete, yet closely interrelated stages or levels—namely, *upstream*, *midstream*, and *downstream*. The upstream (or macro-level) factors include international influences, government policies, and the fundamental social, physical, economic, and environmental determinants of health. The midstream (or intermediate-level) factors include psychosocial factors, health-related behaviors, and the role of the healthcare system. Some social factors, such as

culture, beliefs, values, and norms, are seen to influence decision making, actions, and behavior at both upstream and midstream levels. The downstream (or micro-level) factors include physiological and biological functioning. This model has been used to identify interventions targeted at entry points in all three levels, either singly or in combination. While acting upstream (such as with tax policies) may have wider population effects, demonstrating its influence is often more complex.

Dahlgren and Whitehead (2007) developed a similarly multilayered and widely used "rainbow" model of determinants, shown in **FIGURE 4-1**. This model has at its core the individual biological determinants that are not considered to be SDH—age, sex, and constitutional—as characteristics affecting health that are largely fixed. From here, the framework adds layers of determinants that are socially determined from individual lifestyle factors; farther upstream are those factors and services that impact on health at the population level and that are theoretically modifiable by policy. This model shows not only the different levels, but also the preponderance of factors that may be considered as SDH. These factors may be health promoting (e.g., provision of adequate housing) or protective, by eliminating risk of disease (e.g., pollution control). In this model, healthcare services coexist with other determinants to impact on health. Dahlgren and Whitehead (2007) propose that a comprehensive health strategy should address and link the different levels of downstream and upstream determinants of health shown in Figure 4-1.

The Dahlgren and Whitehead model suggests the determinants that may be included in the definition of what is socially determined and the relationship between these different levels of SDH on health outcomes. Nevertheless, it does not show how these different SDH relate to health equity outcomes. The authors do, however, argue that the determinants of *inequities* in health may be different from the determinants of health. For example, poor working conditions may account for a higher share of the *difference* in the burden of disease between affluent and low-income groups than in the *overall* burden of disease. Thus, they argue that actions on SDH may not automatically address equity, and that specific attention needs to be paid to the *distributional* impact of those actions for them to achieve this goal. For example, urban

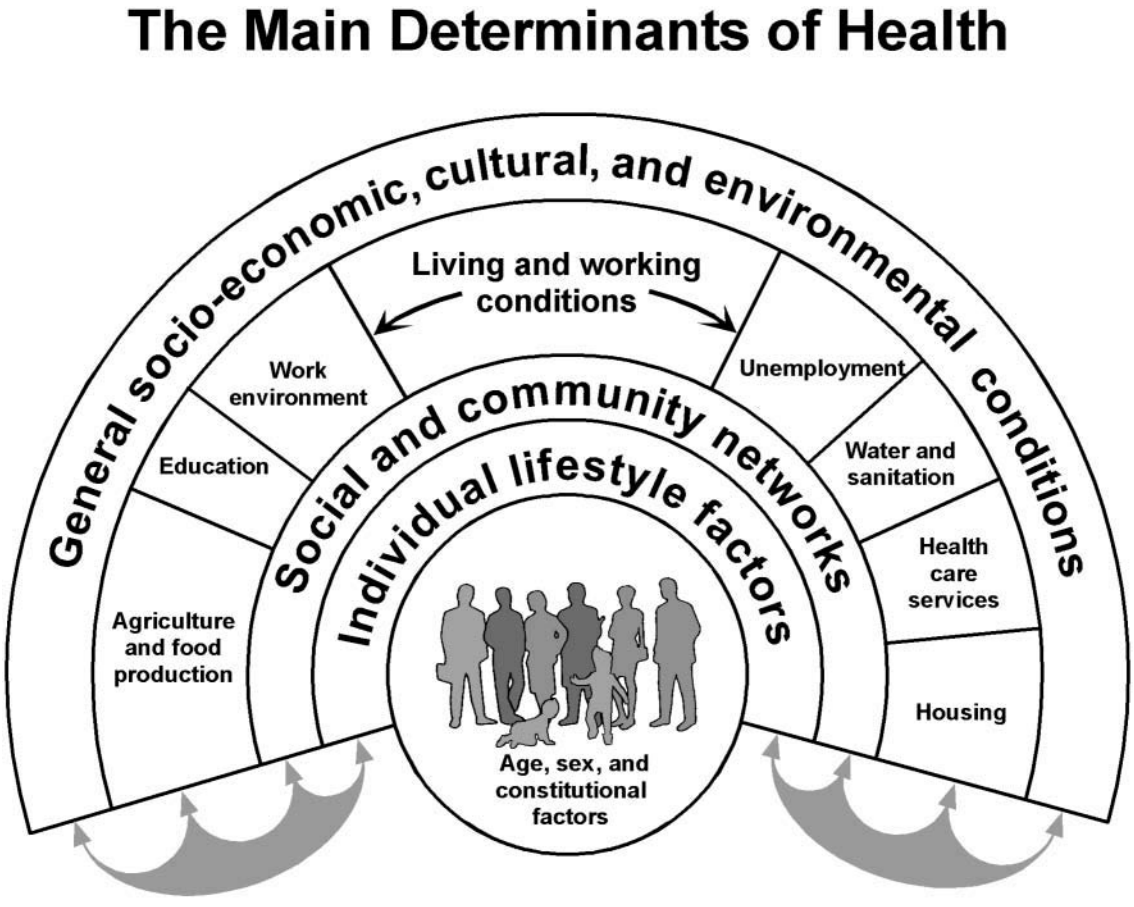


FIGURE 4-1 The "rainbow" model of determinants of health, including social determinants of health.

Reproduced from Dahlgren G., Whitehead M. (2007). *Levelling up (part 2): A discussion paper on European strategies for tackling social inequities in health*. Copenhagen, Denmark: WHO Regional Office for Europe. http://apps.who.int/iris/bitstream/10665/107791/1/E89384.pdf (accessed 22 June 2017).

developments to improve the quality of housing and green spaces are actions on SDH that may improve health, but may also raise housing costs and push local residents into more marginal urban zones, forcing them out from areas where they have been long-time residents. This trend will affect the distribution of health benefits, unless specific measures are in place to protect their tenure or well-being.

The *social gradient* in health describes the pattern that is formed by comparing health outcome measures with some measure of social position, such as occupation, education, or income (Kelly & Doohan, 2014). A population-level public health intervention that affects the whole population in the same way may shift the gradient upward by the same amount for all, without affecting the relative differences between the different social groups.

If the goal of a policy or an intervention is to improve health equity, the rate of improvement should be faster for those groups whose starting point is worse initially, making the health gradient less steep (Kelly & Doohan, 2014). If the gradient simply shifts upward at the same rate for all groups, such that all groups improve at the same rate in their absolute outcomes, then the relative differences between the groups remain unchanged—and in that sense, health inequity has not changed. Population health interventions that are applied universally to address specific SDH without taking social differentials into account may, indeed, worsen inequalities in outcomes. For example, as health technologies are introduced or health screening scales up, uptake may initially be more rapid in wealthier households, which widens (rather than narrows) the social gradient, unless specific measures are implemented to promote uptake in poorer groups (**EXHIBIT 4-1**). A more detailed discussion of absolute and relative measures of social status and their relationship to absolute and relative measures of health outcomes can be found in the chapter by Kelly and Doohan (2014).

EXHIBIT 4-1 Closing the Equity Gap: A Case Study of the Health Gradient

This example illustrates the way the health gradient does, and does not, respond to new public health interventions and the "stubborn" influence of SDH. In the state of Ceará, in a poor area of northeastern Brazil, the 1980 infant mortality rate was greater than 100 infant deaths per 1,000 live births and malnutrition was common. Based on the findings of a 1986 statewide survey of child health and nutrition, new health policies were implemented, including growth monitoring, oral rehydration, breastfeeding promotion, immunization, and vitamin A supplementation—interventions collectively known as GOBI (growth monitoring, oral rehydration, breastfeeding, and immunization). As lack of access to health services was a major problem, community health workers and traditional birth attendants were widely introduced and responsibility for health services was decentralized to rural municipalities, where health indicators were poorest. A social mobilization campaign for child health used media and small radio stations to broadcast educational messages. This work was supported by four consecutive state governors, all of whom gave high priority to improving child health, and the outcomes were reviewed through surveys in 1990 and 1994.

The surveys showed improved population coverage of the four GOBI interventions by 1994. The use of oral rehydration increased to more than 50% in children with diarrhea; nearly all children had a growth chart, and half had been weighed within the previous three months; immunization coverage rose above 90%; and the median breastfeeding duration increased from 4.0 months to 6.9 months. These improvements were noted in all income groups. The prevalence of low-weight-for-age and low-height-for-age children in the population younger than age 5 fell from 13% to 9% and from 27% to 18%, respectively, while diarrheal disease incidence in the previous two weeks was nearly halved.

While the child health interventions were applied to all families, including the poorest, the social differentials in disease and infant mortality rate between rich and poor remained largely unchanged between 1987 and 1994. In that period, family income inequalities persisted and remained largely unchanged. Diarrheal disease incidence remained approximately 60% higher among poor children, income inequalities remained largely unchanged, and the inequity ratio between rich and poor persisted.

One explanation for these findings given by the authors of the Ceará study was that wealthy families had made greater and earlier use of both public-sector and private-sector services to protect their children's health. They noted that even when public health programs are targeted at the poorest households, it is difficult to close the inequity ratio group if the richer households have not yet themselves achieved high levels of coverage with the specific services. The conclusions suggested that, even when public health programs are targeted at the poorest members of a society, the wealthiest are likely to continue to benefit from the introduction of new health technologies, and that further investments are needed to make existing and new interventions more widely accessible to the poorest populations so as to change the health gradient and close the inequity gap (Bonnefoy et al., 2007; Victora et al., 2000, cited in Kelly & Doohan, 2014).

Dahlgren and Whitehead (2007) note that understanding how SDH are linked to improving health equity calls for investigation of what is causing the social gradients in exposure to different health hazards and in access to health-promoting or health-protective factors, such as those described in Exhibit 4-1. Analysis of these SDH associated with distributional outcomes is used to identify the approaches and additional resources that may be needed to address social gradients in health, such as through "leveling-up" strategies. Once these SDH that have positive or negative influence on social gradients are identified, specific strategies can be designed to integrate them within health policies and programs. Dahlgren and Whitehead argue further that policies and programs that influence health should be assessed for their distributional impacts across different socioeconomic groups.

The SDH that have a distributional impact may arise within any of the bands of SDH in the rainbow model shown in Figure 4-1. For example, social position is an important determinant of inequities in health, as are social networks and health behaviors. Groups that are socially and economically better off typically have more power and opportunities to live a healthy life than groups that are less privileged (Australian Institute of Health and Welfare [AIHW], 2012). Legal and institutional arrangements, political and market forces can consolidate these inequalities. The relationships are not purely negative. In contrast, and as exemplified in Exhibit 4-8, shown later, education may be a positive factor in closing inequalities, as it can enhance the power and opportunity of less privileged groups to avoid unhealthy living and working conditions. The freedom and power that people have to influence their own life and society, including through greater political and economic democracy and through control of commercial markets (such as marketing of alcohol), is also a positive factor in tackling social inequities in health (EQUINET, 2012; EQUINET SC, 2007).

Given the increasing concern voiced regarding remediable inequalities in health, Solar and Irwin (2010) elaborated the conceptual framework that was used by the WHO's Commission on Social Determinants of Health (CSDH). This framework, shown in **FIGURE 4-2**, makes an explicit connection to health equity by seeking to profile how SDH relate to and address unfair, avoidable, or remediable differences in health among population groups, as an appeal to ethical norms and a matter of social justice.

The CSDH framework repeats many elements of the earlier frameworks for SDH, but applies an equity lens. It locates SDH within a hierarchy of structural and intermediary determinants that gives specific focus to the political, social, and economic contexts and institutions and the cultural and societal values (all SDH) that *generate, configure, and maintain social hierarchies and that stratify people* by income, education, occupation, gender, race/ethnicity, and other factors (also

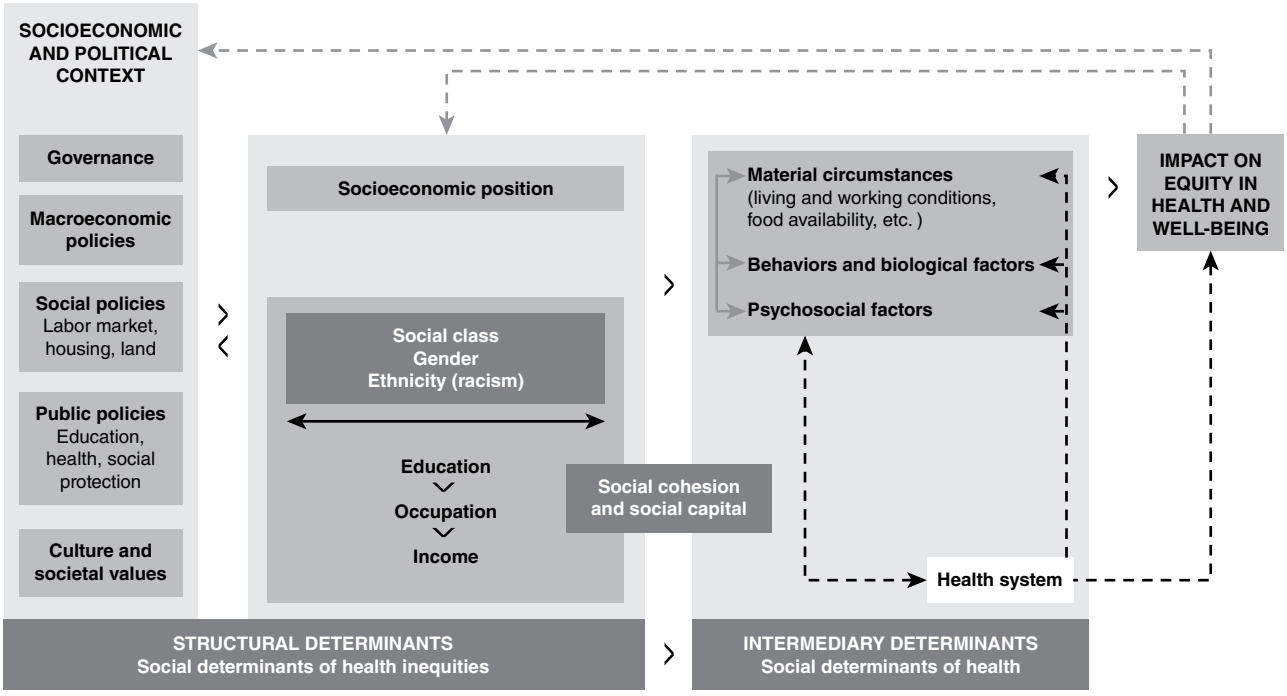


FIGURE 4-2 The Commission on Social Determinants of Health (CSDH) conceptual framework.

Reproduced from Solar, O. and Irwin, A. (2010). *A conceptual framework for action on the social determinants of health. Social Determinants of Health.* WHO Discussion Paper 2 (Policy and Practice), Geneva, Switzerland: WHO. http://www.who.int/sdhconference/resources/ConceptualframeworkforactiononSDH_eng.pdf (accessed 22 June 2017).

SDH). It includes the SDH—including state policies and programs—that *redistribute resources* and nurture relationships and systems that *close social differentials*. These structural determinants and the socioeconomic positions to which they lead, shown in the left-hand columns in Figure 4-2, thus have a central role as SDHE. They jointly shape intermediary SDH, such as material, behavioral, and psychosocial determinants. These factors largely act as SDH but may have a role in equity where they influence *differentials in exposure and vulnerability* to conditions that affect health or the social consequences of ill health, as discussed for the health system later in this chapter. The health outcomes of these conditions can themselves "feed back" on people's social position, such as by compromising employment opportunities or reducing income.

While previous models have noted the role of the health system in dealing with social determinants of health, this framework positions the health system as one of the *intermediary determinants*. It recognizes that health systems both reflect existing patterns of social inequality and provide a site from which to contest them (a topic discussed in more depth in a later section). The CSDH framework also recognizes the effect of globalization and global-level drivers on SDH at the national and local levels (also discussed later).

Policies and interventions that seek to improve equity may be targeted at SDH at both the structural and the intermediary levels: from the "micro" level of individual interactions (such as individual dietary practices), to the "meso" level of community conditions (such as neighborhood environments), to the broadest "macro" level of universal public policies (such as tax and investment policies), including in the global environment and in relation to the role and impact of private actors and corporations. Applying the framework entails assessing the distributional impact of actions and action across multiple areas of SDH. As shown in **FIGURE 4-3**, the CSDH framework identifies these distributional impacts at the level of impacts on social stratification, on differentials in exposure, on differentials in vulnerability to risks, and on the unequal consequences of illness.

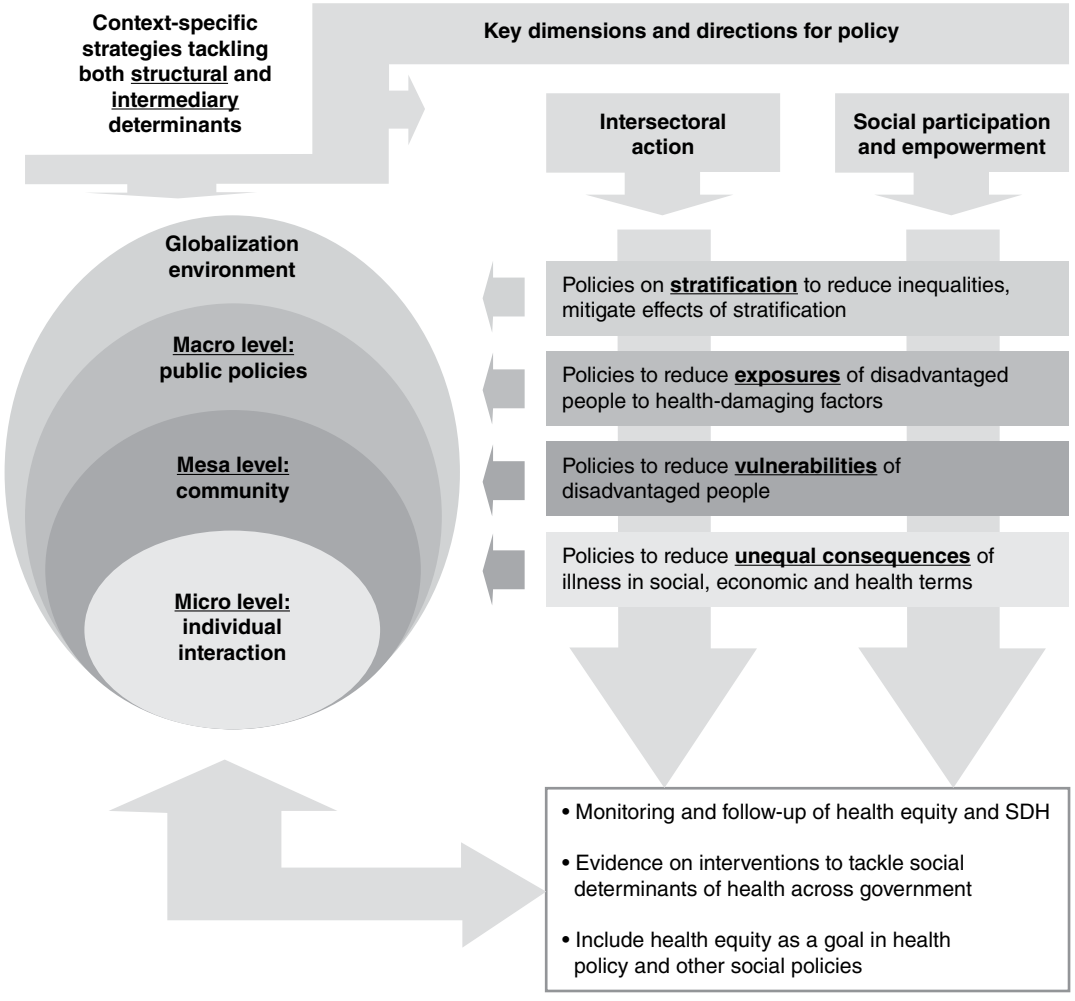


FIGURE 4-3 The CSDH framework for tackling inequalities in health.

Reproduced from Solar, O. and Irwin, A. (2010). *A conceptual framework for action on the social determinants of health. Social Determinants of Health*. WHO Discussion Paper 2 (Policy and Practice), Geneva, Switzerland: WHO. http://www.who.int/sdhconference/resources/ConceptualframeworkforactiononSDH_eng.pdf (accessed 22 June 2017).

Across each of these dimensions, identifying and taking action on SDH to reduce health inequities may have the following benefits:

- Remedy health disadvantage among specific populations—as, for example, in SDH interventions for the Roma community, outlined in Exhibit 4-7
- Close health gaps between two groups, usually the worse-off and better-off groups, as intended in the work of community health workers in Guatemala to improve service uptake in poorer groups, described in Exhibit 4-7
- Address the social health gradient across the whole population—as, for example, is a policy intention of the various features of universal health systems described in later in this chapter

These approaches are not mutually exclusive, but rather may bring together action on different SDH at different levels. As discussed later in this chapter and in Exhibit 4-1, strategies that tackle one or more of these outcomes may be complementary, such as when additional measures for single mothers are embedded within general social protection schemes. Such action on SDH calls for input from multiple sectors, as explored later in this chapter.

Understanding the SDH that affect how different social groups access or experience universal interventions can point to specific additional measures that need to be integrated to close the gap or gradient. Such analysis, as a form of "equity proofing," can, for example, be integrated in health impact assessments (HIA) to check for the likely equity implications and consequences—both intended and unanticipated—of a particular action, policy, or intervention, and to recommend proactive measures to improve health equity (Simpson, Mahoney, Harris, Aldrich, & Williams, 2005).

In all of these models, applying an equity lens to SDH to identify and address these social determinants of health equity calls for robust evidence on the distribution of health and its determinants, and for monitoring and evaluation systems that contribute such evidence to policy decision making. Evidence on SDH often extends beyond familiar health indicators, and assessing changes in health gradients and across social features, areas, and time can be complex.

Solar and Irwin (2010) observe that there is a dearth of evidence, including data on impacts, to support policy action on SDH and inequities in health. Cochrane reviews of studies of tobacco control, for example, rarely assess the impact of policies or programs in relation to socioeconomic dimensions (Oxman, Lavis, Lewin, & Fretham, 2009). As noted earlier, the interventions are often multifaceted, combining different levels of SDH and sectors, relative to the often narrower biomedical/clinical health interventions. The former are context dependent, are delivered within a constantly changing health and social policy environment, and involve diverse disciplines (Oxman et al., 2009). While there is growing recognition of and research on these SDHE, further development of methods that take the complexity of such interventions into account is still needed, and funding for this research is still limited relative to that allocated for the biomedical sciences (CSDH, 2008). The CSDH (2008, p. 186) has noted that "action on SDH is best served through developing a rich and diverse evidence base." This chapter thus includes evidence from multiple disciplines and methodological traditions, ranging from analysis of repeated household surveys, to qualitative studies, policy analysis, studies of power relations, practice reviews, and system-level evaluations.

The conceptual frameworks discussed in this section also suggest that social processes and power differentials play an important role in the social hierarchies that stratify people and in acting on SDH that affect health equity, as exemplified in the case of indigenous people in **EXHIBIT 4-2.**

EXHIBIT 4-2 Understanding Systematic Discrimination Against Indigenous People as an SDH

Indigenous people in Australia (as in other countries) are not merely "disadvantaged citizens." The poverty and inequality in health that they experience reflect the association with systemic discrimination over centuries. Thus, addressing these issues is not simply a matter of provision of material or service inputs, but rather requires addressing ways to increase their control over their physical environment, their dignity, and their community self-esteem, and the matter of justice (Calma, 2007). Such systemic processes of discrimination and disempowerment cannot be simply understood as historical events reflected in poorer current health outcomes for indigenous people, but must be recognized as persistent in their social consequences and within current sociopolitical contexts, with continued negative impacts on health outcomes (Axelsson, Kukutai, & Kippen, 2016). The CSDH (2008) and the indigenous health research community have called for such effects of colonialism to be more explicitly incorporated into epidemiologic analysis and monitoring systems.

Health interventions and systems are thus not simply technical or biomedical in nature, but rather affect social power and status and reflect procedural justice, such as through decision-making processes that are transparent to, accessible to, and inclusive of the input of those affected by them. The next sections explore further how intervening on SDH to promote equity calls for institutions and processes that provide opportunities and spaces for disadvantaged and marginalized groups to engage in decisions that affect their health, that are responsive to need, and that recognize and deliver on health rights, as noted in the 2008 CSDH report: "Changing the social determinants of health and health equity is a long term agenda requiring sustained support and investment . . . recognising that . . . at the centre of this action should be the empowerment of people, communities and countries that do not have their fair share" (p. 23).

Integrating concepts of power into the analysis of SDH and into the responses to the resulting inequities in health both clarifies and complicates frameworks. Solar and Irwin (2010, p. 21) argue that while power is "arguably the single most important organizing concept in social and political theory," it is also contested and subject to diverse and often contradictory interpretations. It draws public health professionals into the sometimes less familiar terrains of social justice and rights–based approaches, of sociopolitical and systems theory, and of activism, discussed further in the chapter.

Relationships Between SDH, Gender Equality, and Human Rights

The conceptual frameworks highlighted earlier indicate that SDH intersect with the distinct but linked concepts of gender and human rights, particularly when applying an equity lens. While sex is a biological determinant, gender refers to the socially constructed roles, rights, responsibilities, and limitations assigned to women and men, boys and girls—which often privilege male power or characteristics (WHO, 2011). These aspects of gender are socially constructed and amenable to change, as an SDH that leads to *differentials in exposure and vulnerability* to conditions that affect health and thus health equity (Sen & Ostlin, 2011).

Gender norms, roles, and relations can affect (health) risk and vulnerability, health-seeking behavior, and health outcomes for men and women of different ages and social groups, including through stereotypes, discrimination, and the gender-based division of labor (Sen & Ostlin, 2011; WHO, 2011). Gender norms and relations are a persistent basis of the social hierarchies and stratification mentioned earlier, intersecting with social class, ethnicity, education, occupation, and income; influencing socioeconomic position and the distribution of other SDH; and being influenced by the wider socioeconomic and political context, culture, and societal norms and values (Sen & Ostlin, 2011). The Definitions box summarizes the related concepts of *gender equality* and *gender equity*. Within universal policies such as "health for all," the "all" are not the same. Gender-related differences that lead to inequities in health arise from the different health needs and challenges that men and women face across their life course and the ways in which they intersect with other SDH. Gender-related differences exist in a range of SDH, including living and environmental conditions, employment and income opportunities, and control over decisions about and uptake of health services (WHO, 2011). Gendered norms, behaviors, and socially constructed roles intersect with other SDH to generate differential social and health outcomes.

Many of these differentials derive from women's status in society, and their control over a range of areas affecting health, including over their own bodies, their reproductive health, and their working conditions and income (for example, see **EXHIBIT 4-3**). Women are at greater risk of physical violence and sexual abuse and face deficits in protection in law or its enforcement (Sen & Ostlin, 2011). In addition, female, lesbian, gay, bisexual, transgender, and intersex (LGBTI) people risk poorer health outcomes than males or people who are not LGBTI because of gender power relations (Sen & Ostlin, 2011).

DEFINITIONS

- *Gender equality* refers to women and men having equal conditions and opportunities to realize their rights and potential to be healthy, to contribute to health development, and to benefit from the results. Gender inequality puts the health of women and girls at risk globally. Improving gender equality in health enables the improvement in the health of women.
- *Gender equity* refers to fairness and considers women's and men's different needs to achieve gender equality. It implies the different treatment needed to ensure equality of opportunity. Both gender equality and gender equity are needed to achieve health equity (Sen & Ostlin, 2011; WHO, 2011).

EXHIBIT 4-3 Child Marriage and Its Gendered Health Effects

Child marriage is a union (official or not) of two persons, at least one of whom is younger than 18 years of age. It is driven by gendered values and beliefs about girls' roles and contributions, affects girls and boys differently, and is more prevalent among girls. In Niger, for example, 77% of women age 20 to 49 were married before age 18, compared to 5% of men in the same age group. Collecting data on such unions poses a challenge, because child marriages are often not registered. However, worldwide, more than 700 million women in 2014 were married before age 18, and about 250 million before age 15, with the highest rates found in South Asia and sub-Saharan Africa. There are educational, regional, urban–rural, and wealth differences in this practice: In Malawi, nearly two-thirds of women with no formal education were child brides compared to 5% of women with secondary or higher level schooling; in Amhara, northern Ethiopia, the rate of child marriage is 75%, compared to 26% in Addis Ababa; and girls in the poorest quintile are 2.5 times more likely to marry in childhood than those in the wealthiest quintile.

Girls who are married before age 18 experience greater negative impacts on physical and mental health and well-being than do boys in this situation, primarily due to complications from early pregnancy and childbearing, increased risk of and vulnerability to human immunodeficiency virus (HIV) and other sexually transmitted infections, and gender-based violence and discrimination. Complications during pregnancy and childbirth are the second leading cause of death for 15- to 19-year-old girls globally. Adolescent mothers are often socially isolated from family, friends, and other sources of support, such as health services; have limited opportunities for education and employment; and may be less empowered to break the cycle of poverty due to their lower levels of education. Early maternal age affects infant mortality and growth, compounding intergenerational inequality. Furthermore, unregistered child marriages may mean that the child's birth is also unregistered, potentially affecting the child's access to health, social, and educational services (UNICEF, 2014, 2016; WHO, 2014; WHO Regional Office for Europe [WHO Europe], 2016).

In terms of reducing health inequities,

> [because] of the numbers of people involved and the magnitude of the problems, taking action to improve gender equity in health and to address women's rights to health is one of the most direct and potent ways to reduce health inequities and ensure effective use of health resources. Deepening and consistently implementing human rights instruments can be a powerful mechanism to motivate and mobilize governments, people and especially women themselves. (Sen & Ostlin, 2011, p. 74)

International human rights instruments provide a framework for universality and a legal obligation for states to ensure policies that support conditions and opportunities for health for all (Braveman & Gruskin, 2003). Human rights–based approaches thus raise the relevance of acting on SDH for improved health, so as to meet the 1948 Universal Declaration of Human Rights Article 25 provision: "Everyone has the right to a standard of living adequate for the health and well-being . . . including food, clothing, housing and medical care and necessary social services" (United Nations, 1948).

A human rights–based approach also draws attention to the SDH that affect health equity, to ensure the "right to the enjoyment of the highest attainable standard of health . . . without distinction of race, religion, political belief, economic or social condition" for all, as provided in Article 12 of the 1966 International Covenant on Economic, Social and Cultural Rights (ICESCR) (United Nations, 1976). The ICESR General Comment 14, in addition to identifying universal access to specific SDH for the whole population, identifies the core state obligations in relation to health as ensuring *nondiscriminatory access* to health facilities, goods and services, *especially for marginalized groups* (United Nations, 2000). The right to health provides for claims by people as rights bearers to key SDH, including health care, and its achievement depends on the realization of these entitlements and freedoms (Gesellschaft für Internationale Zusammenarbeit [GIZ] & WHO, 2011a). While some rights violations, such as access to safe water, are measurable and have clear health consequences, there is some debate on what the right to "the highest attainable standard of health" means in practice. It has been argued that one dimension of this is closing avoidable inequalities in health within and across countries (EQUINET, 2012).

Rights-based approaches in health imply not only intervention on intermediary SDH, but also processes that include participation, information sharing, and accountability in setting priorities, planning and design, and implementing and monitoring policies and programs that are set in more structural SDH. They imply that states and others charged with duties

The United Nations Human Rights-Based Approach (UNHRBA, 2003) to development has as its purpose to assist UN agencies in mainstreaming human rights into their activities and programs within their area of work with the expected outcome of realization of one or several human rights (e.g., the right to health). Human rights principles should guide programming in all sectors, and include capacity building of rights holders to make their claims and of duty bearers to meet their obligations (UNHRBA, 2003).

have the resources and capacities to meet their obligations; and that social groups—particularly vulnerable groups—have the information and capacity to claim their rights, with access to redress mechanisms such as human rights commissions, policy reviews, audits, and courts to pursue violations.

Nevertheless, the resource constraints facing states, particularly in low-income countries, also generally imply a principle of progressive realization of the right to health (and its determinants)—that is, to move as quickly and effectively as possible toward their achievement. ICESCR General Comment 14 refers to core obligations that require immediate action versus obligations that can be progressively realized, supported by measures to show constant progress (United Nations, 2000). States need to demonstrate that they are taking deliberate, concrete, and targeted steps towards realizing these health and SDH rights, to put all appropriate and available means in place for this progress, and to monitor and report on their implementation against agreed benchmarks (GIZ & WHO, 2011b).

There is an interaction between gender and human rights frameworks in addressing SDH. Human rights principles of equality, participation, and nondiscrimination applied in addressing SDH address gender inequalities and gender equality not only as an SDH, but also as a human right. Addressing gender inequalities is, further, essential to realizing human rights, including the right to health (WHO, 2011). Gender mainstreaming as a strategy thus applies a human rights–based approach, in analyzing the impact of gender norms, roles, and relations on rights to health, and in implementing measures to address rights violations and institutionalize gender equality, in a manner that integrates the participation of affected women (WHO, 2011).

Hence, despite their differences in focus, SDH, gender equality, and human rights frameworks share some common features: All are explicitly guided by an ethical (normative) foundation that values and promotes procedural justice; all identify health outcomes as the result of wider social and material conditions (and seek to generate evidence and analysis on this relationship); and all recognize the role of social processes and power differentials in creating health inequities. Moreover, all three see the state and participation of affected populations as central in the response to health inequity and action on the SDH, through actions across numerous sectors, including the health sector.

Addressing SDH Through Intersectoral Action and Health in All Policies

The diversity of SDH noted in the frameworks introduced earlier in this chapter implies that actions on SDH call on many sectors. The concept of *intersectoral action for health* was formally introduced at the 1978 International Conference on Primary Health Care (PHC) in Alma-Ata. The PHC approach explicitly identified the need for both comprehensive health services and joint action with other sectors to tackle the SDH and root causes of poor health (Public Health Agency of Canada [PHAC], 2007), with subsequent global processes further stressing the importance of dialogue and joint action with other sectors (CSDH, 2008)

Many SDH lie outside the remit of the health sector (CSDH, 2008). For example, a global review of evidence on the determinants of diabetes identified that improving health outcomes for this condition involves the following measures: (1) limiting the availability of unhealthy food and environments; (2) tackling the obesogenic environment through improving urban infrastructure to promote physical activity; and (3) reducing exposure and addressing increased vulnerability among certain groups by improving

DEFINITION

Intersectoral action refers to processes in which "the objectives, strategies, activities, and resources of each sector are considered in terms of their implications and impact on objectives, strategies, activities, and resources of other sectors." It is advanced as a means for overcoming policy fragmentation; as a way to plan, implement, and monitor service delivery; and as a means to address "upstream" the determinants of health (Loewenson, 2013a, p. 8).

screening and access to and uptake of health care (Whiting, Unwin, & Roglic, 2010). The health sector thus needs to cooperate with other sectors to improve health outcomes, as demonstrated in **FIGURE 4-4** in regard to the actions needed to tackle the health burdens from exposure to air pollution.

Such intersectoral action to address SDH can be located as a shared goal across government, with "Health in All Policies" (HiAP) being one such approach. A HiAP approach is driven from within government as a whole, linking specific programs and measures to wider government policy agendas

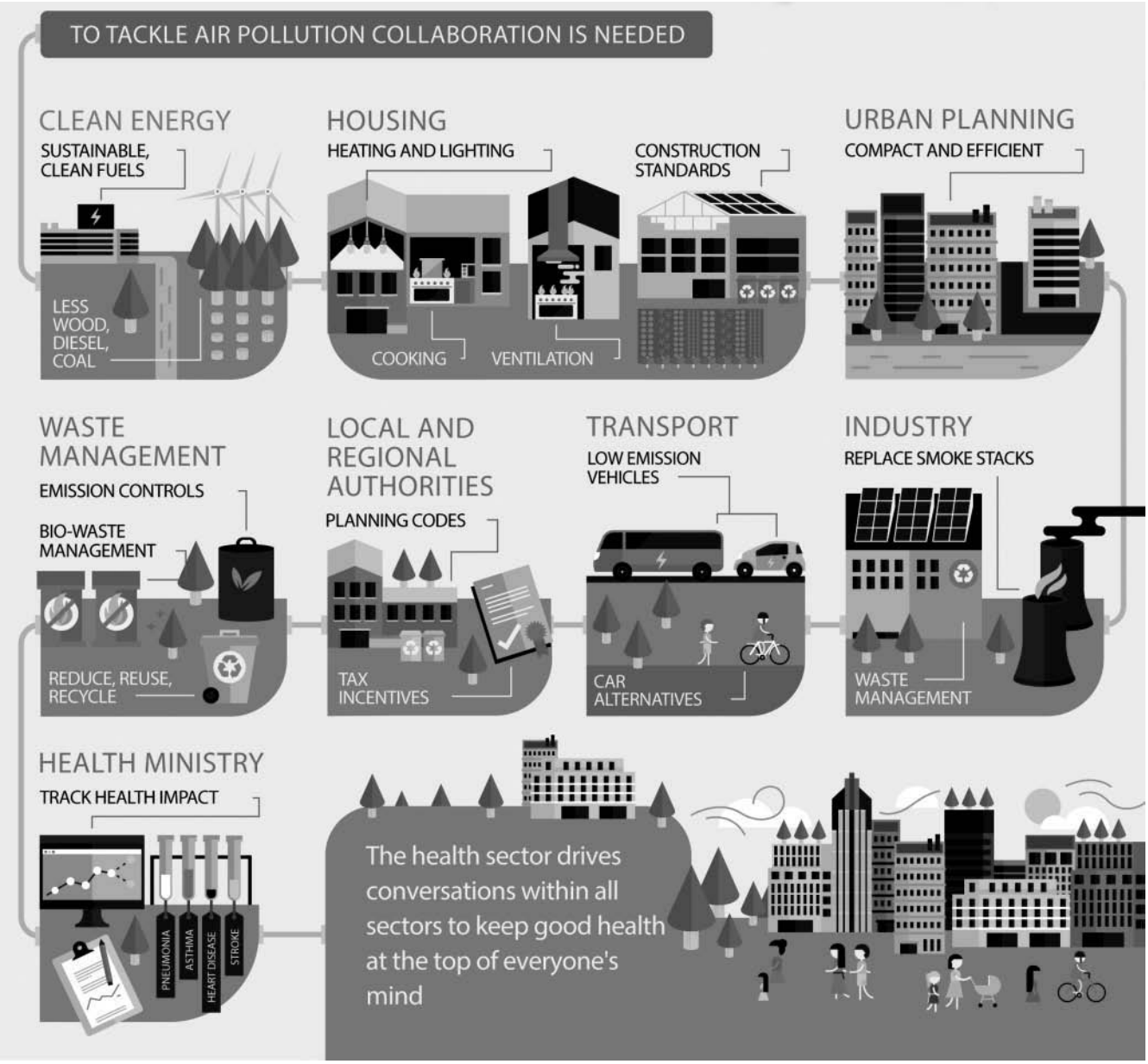


FIGURE 4-4 Intersectoral collaboration on air pollution.

Reproduced from WHO. (2015). *What is health in all policies?*, Infographic on air pollution. http://who.int/social_determinants/publications/health-policies-manual/HiAP_Infographic.pdf?ua=1 (accessed 19 May 2017).

DEFINITION

Health in All Policies is an approach to public policies across sectors that systematically takes into account the health and health systems implications of decisions, seeks synergies, and avoids harmful health impacts, with the goal of improving population health and health equity. A HiAP approach is founded on health-related rights and obligations.

(Ollila, Baum, & Peña, 2013, p. 6)

(Shankardass, Solar, & O'Campo, 2012). While formally articulated as policy concepts from the 1970s, HiAP and intersectoral action for health have their roots in the early history of public health and "healthy public policy." HiAP builds on experiences of intersectoral action and health promotion, and has been given a greater international profile since 2000 (Friel, Harris, Simpson, Bhushan, & Baer, 2015; PHAC, 2007).

EXHIBIT 4-4 provides an example of HiAP in South Australia, located within the overall state strategic plan. Applying HiAP calls for information systems and strategic review involving all relevant sectors to gather evidence and evaluate how joint actions are affecting the distribution of SDH as well as the social gaps and gradients in health, so as to adopt appropriate responses (Loewenson, 2013a; WHO, 2013a).

As noted in Exhibit 4-4, intersectoral action calls for measures to build collaboration and coordination across sectors, often departing from the administrative, budgetary, and functional siloes that exist within states. It involves whole-of-government approaches, which lever action across all levels and sectors of government, and whole-of-society approaches for public outreach, including through parliaments/legislatures, civil society, and the private sector (Delaney et al., 2016; WHO Europe, 2013). A review of experiences in more than 15 countries identified key measures for the successful implementation of intersectoral action, shown in **EXHIBIT 4-5** (PHAC, 2007; Rasanathan, 2011).

The HiAP approach and the implementation steps in Exhibit 4-5 can be applied to embed health equity as a goal across all relevant national policy, law, and

EXHIBIT 4-4 The South Australian Approach to Health in All Policies

HiAP in the South Australian government, as adopted in 2008, was developed as a central process of government, rather than one run by and for the health sector. It was integrated in the implementation of the South Australian Strategic Plan, which aimed to improve prosperity and well-being, foster creativity, build communities, and sustainably expand opportunity. The approach sought to better utilize social, economic, and environmental levers to influence population health, and in turn to contribute to achievement of the overarching government vision for South Australia as articulated in the strategic plan. HiAP is thus deliberately and strategically positioned as a central process of government, rather than a process run by the health sector to achieve solely its objectives. A range of agencies use HiAP in partnership with South Australian Health as a mechanism for achieving their own goals, aligned to the objectives and related targets of the strategic plan. The implementation of HiAP is supported by central government, in partnership with the health sector.

The 2011 South Australian Public Health Act provides for the systematic integration of HiAP approaches in government functions and for mechanisms for embedding health considerations in government decision-making processes. The act mandates use of HiAP in public health planning by local municipalities and in health impact assessments and any other measures that they use, such as a health lens analysis. Health lens analysis builds on traditional methods for health impact assessment and incorporates additional methods used by other sectors, such as economic modeling. Such analyses have, for example, focused on water sustainability, regional migrant settlement, digital technology access and use, healthy weight, and active transport. A small HiAP unit was established within South Australian Health to facilitate health lens analysis and sustain the partnership with other government agencies in applying HiAP to their targets (Bucket, Williams, & Wildgoose, 2011; WHO, 2013a).

EXHIBIT 4-5 Measures for Steps Necessary for Successful Implementation of Intersectoral Action for Health

1. Create a policy framework and an approach to health that are conducive to intersectoral action.
2. Emphasize shared values, interests, and objectives among all partners and potential partners.
3. Ensure political support, building on positive factors in the policy environment.
4. Engage key partners at the very beginning.
5. Ensure appropriate horizontal linking across sectors and vertical linking of levels within sectors.
6. Invest in the alliance-building process by working toward consensus at the planning stage.
7. Focus on concrete objectives and visible results.
8. Ensure that leadership, accountability, and rewards are shared among partners.
9. Build a stable team of people who work well together, with appropriate support systems.
10. Develop practical models, tools, and mechanisms to support implementation of intersectoral action.
11. Ensure public participation through education and awareness raising of SDH and intersectoral action.

Reproduced from WHO. World Conference on Social Determinants of Health. Closing the Gap: Policy into Practice on Social Determinants of Health. Retrieved from: http://www.who.int/sdhconference/Discussion-paper-EN.pdf

programs, when specific focus is given to those SDH that have a role in closing the gap or gradient. For example, in 2007, a Norwegian parliament whole-of-government approach aimed to reduce social inequalities in health in 2007–2017 by "leveling up" across the social gradient. The strategy linked efforts to reduce social inequalities in health to government policy initiatives to improve employment, welfare, and inclusion, and to support early intervention for lifelong learning (Norwegian Ministry of Health and Care Services, 2007). Another example of such a "whole-of-government" approach in Mozambique is described later in Exhibit 4-12.

While there may be technical evidence or consensus on the relevance of specific SDH to health outcomes, not all countries have the high-level leadership support, policy space, or institutional systems to support such whole-of-government approaches to cooperation across sectors. Where this is the case, intersectoral action may be built by the health sector itself collaborating with one or more sectors to support the shared goals of both sectors. For example, education measures that support secondary school completion rates for girls, such as those described later in Exhibit 4-9, contribute both to improved education outcomes and to improved adolescent health, with longer-term health, socioeconomic, and well-being gains for girls—and their future children—extending into adulthood (Loewenson, 2013b).

These approaches help to build the confidence in, relationships for, and systems to support intersectoral action for health. Tackling health inequities, and particularly those that are affected by structural determinants, often demands the deeper and more sustained levels of collaboration through the whole-of-government approaches described earlier, as found in a review of experiences of 19 countries in Africa, Southeast Asia, and the Western Pacific (WHO, 2013a). In one such experience, described in **EXHIBIT 4-6**, Vietnam sought to improve road safety and to alleviate traffic congestion, both of which had the potential to have positive impacts on health and health equity.

Advancing intersectoral action calls for governance arrangements and systems to design and implement coherent multisectoral and cross-sectoral policies on SDH. This is often a challenge, particularly when health professionals are perceived to be overlooking other sectors' goals and challenges and drawing resources *from* other sectors *to* a health sector agenda, rather than as efforts to mutually improve each sector's policies (Rasanathan, 2011). The goals of other sectors should thus be used to orient analysis and explore areas of mutual interest across sectors. WHO's SDH Sectoral Briefing Series (covering housing, education, transport, energy, and social protection) provides further information

EXHIBIT 4-6 Multisectoral Collaboration Through the National Helmet Law in Viet Nam

The motorcycle is the primary mode of transportation in Viet Nam and contributes to more than half of all road traffic fatalities (Passmore, Tu, Luong, Chinh, & Nam, 2010). Passengers, particularly children, are vulnerable to injury due to widespread lack of helmet use. Despite helmet legislation since 1995, with amendments in 2000, 2001, and 2003, the National Traffic Safety Committee charged with implementation of this law faced challenges due to legislative loopholes. A 2007 law sought to close these loopholes by requiring all riders and passengers to wear helmets on all roads without exception and by increasing penalties tenfold. Although 50,000 helmets were distributed to low-income families nationwide, more loopholes were identified after introduction. For example, no reference was made to the correct use of helmet wearing, and the existing legislation meant neither children younger than 16 years of age nor adults carrying them could be fined. These legislative loopholes were subsequently addressed and attention given to barriers to effective helmet use, including improved availability and quality of helmets for the climate, raising public awareness, and changing beliefs about the impact of helmet use. Multisectoral collaboration was built across government agencies and with key nongovernmental organizations and the private sector. This resulted in three national mass-media campaigns and distribution of free helmets to school-age children sponsored by private companies. Monitoring showed helmet-wearing increased from less than 30% to more than 95%, with this change estimated to have saved more than 1,500 lives and prevented almost 2,500 serious injuries.

This example illustrates an approach to policy change backed by strong political support, integrating legislative, research, and media work and strategic alliances. There was cross-sectoral support because the actions were seen to yield benefits for a range of sectors, including road safety, alleviation of traffic congestion, improved health, and reduced costs to the state and families from serious injury and potential loss of income. Focusing distribution of helmets on low-income families aimed to support equity by facilitating compliance with the law among those who were least able to afford a helmet, but most reliant on motorcycle travel (Nguyen, Passmore, Cuong, & Nguyen, 2012; Passmore, Tu, Luong, Chinh, & Nam, 2010; Passmore, Nguyen, Nguyen, & Olivé, 2010; WHO, 2013a).

to support such collaboration with sectors outside health, such as in supporting transport authorities' banning of lead and sulfur in gasoline as not only a health-promoting measure, but also as a contributor to improved air quality and a sustainable environment (WHO, 2017a).

Implications for Health Systems and Services

As further elaborated in *The Design of Health Systems* chapter, effective health services are a determinant of population health, contributing to reducing mortality and increasing life expectancy, particularly through preventive services and health promotion (Loewenson & Whitehead, 2012; WHO Regional Office for Europe [WHO Europe], 2011). They encompass various functions, including stewardship, resource mobilization, financing, and service delivery, and involve a network of public and private organizations, institutions, and resources (WHO Europe, 2011).

Health systems not only provide individual- and population-level services that position them as an intermediary SDH, but, as outlined earlier in this chapter and in Figure 1.1, they can also influence the policies and coordinate with the actions of other sectors to address SDH, thereby providing a site from which to contest social inequalities (Solar & Irwin, 2010; WHO Europe, 2011). Health services and institutions can address differences in exposure and vulnerability, take a leadership or facilitating role in intersectoral action, and mediate or mitigate the financial, social, and physical consequences of illness in people's lives.

This role of the health system in addressing both disease and its underlying SDH is consistent with the application of the PHC approach as an organizational strategy and an underlying philosophy. Within the healthcare system, PHC requires the provision of comprehensive, integrated, and appropriate health services, emphasizing prevention, promotion, and the role of primary care and the role of the health sector in intersectoral action for health (Gilson, Doherty, Loewenson, & Francis, 2008). While past experience of the application of PHC has produced mixed outcomes, evidence of its potential to improve health equity—compiled, for example, by CSDH—has sustained and increased calls for its implementation (Gilson et al., 2008).

There are a range of ways that health systems can achieve positive outcomes and influence the distribution of health and well-being, summarized in **FIGURE 4-5** overleaf. In the figure, the solid arrows show drivers of health inequity, whereas the dotted arrows show how the health system can mitigate these effects or promote health equity. Health systems can promote health equity when they tackle the physical and social environments that affect differential exposure and vulnerability to ill health, including through intersectoral action. They can reduce social gaps and gradients in health by influencing how health services perform, how different social groups experience the services they receive, how widely their uptake or contact translates into effective coverage and care, and whether health funding protects against impoverishment when people fall ill (Gilson et al., 2011).

In a rights-based approach, health systems can provide space for people to exercise their rights, to have an informed say in decisions and actions on health, and to hold the system accountable for its performance. Health systems directly influence other SDH, such as local employment and economies, gender equality, and organizational policies, by their own behavior and organization, in the way they help patients to claim welfare benefits; provide rehabilitation to enable people to keep jobs; manage personnel; procure inputs locally; and champion or facilitate political support to introduce or sustain actions that address SDH and promote health equity (Loewenson & Whitehead, 2012).

While the manner and extent to which these measures are implemented depends on the specific country and sociopolitical contexts, some principles and areas of learning apply more broadly in terms of how the health system implements its role in SDH and health equity.

Investing in Primary Health Care-Oriented Services at the Local Level

Addressing health equity and SDH implies that health systems will pay attention to and direct resources toward health promotion and prevention of disease

DEFINITION

Health systems encompass all the activities whose primary purpose is to promote, restore, or maintain health. They comprise the public and private organizations, institutions, and resources that aim to prevent disease, promote health, and provide health care, and they shape wider societal norms and values (Gilson, Doherty, & Loewenson, 2011).

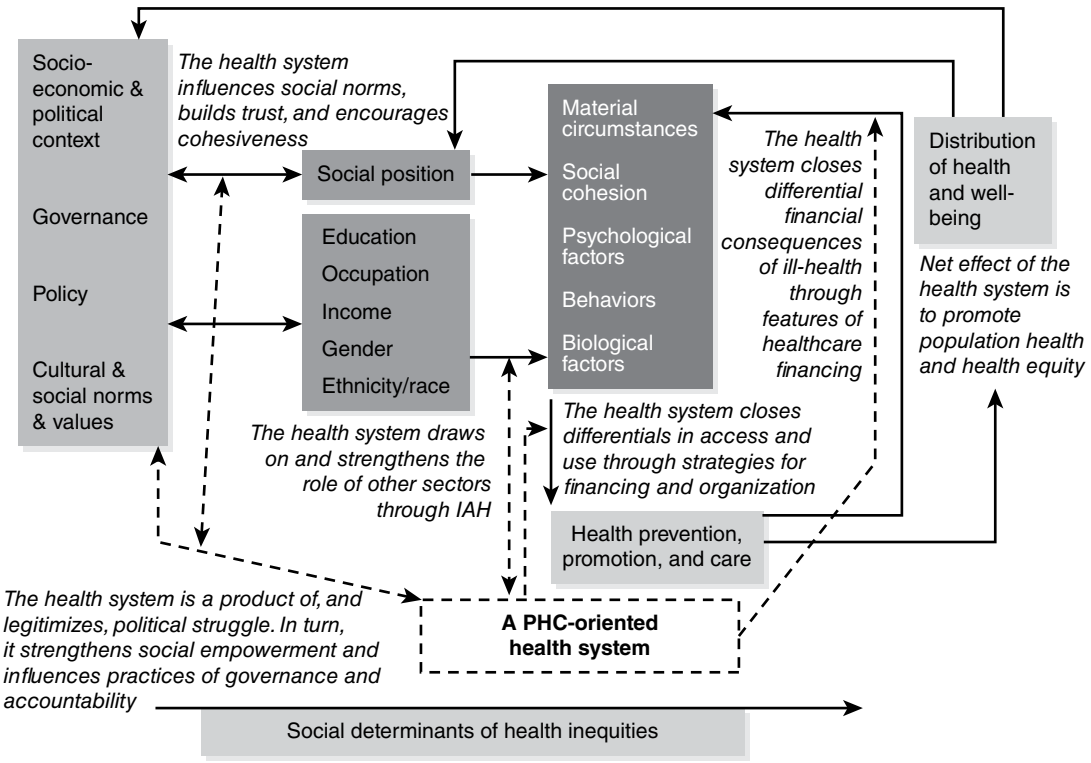


FIGURE 4-5 The health system as a social determinant of health.

Reproduced from Gilson, L., Doherty, J., and Loewenson, R. (2011). 'Challenging inequity through health systems,' in J. H. Lee and R. Sadana (Eds.). *Improving equity in health by addressing social determinants*. Geneva, Switzerland: World Health Organization (WHO). http://apps.who.int/iris/bitstream/10665/44793/1/9789241503037_eng.pdf (accessed 22 June 2017).

at both the individual and population levels, rather than focusing exclusively on individual curative or emergency care. Doing so requires investment in community- and primary care–level services and activities, to open up opportunities for engagement with civil society, communities, and other local services or interventions needed by marginalized groups (Gilson et al., 2008). It is at this level and in this kind of local interaction that many of the SDH and access barriers affecting service uptake can be most directly recognized and addressed, including cost, social, gender, and cultural and other barriers and discrimination; lack of information and knowledge; lack of voice or power; and unresponsive or poorly oriented service providers. This is especially the case if services are to reach and address the health needs of the most marginalized groups (Gilson et al., 2011).

As noted earlier, for health systems to address SDH successfully, they also need to provide public health leadership for and work with other sectors to address the wider social and economic determinants of health, so that people do not repeatedly need treatment after returning to the same conditions that made them ill. This is not simply a matter for those working in community health or health promotion. Rather, it is affected by the continuity and comprehensiveness of the healthcare system as a whole: how well it is able to link preventive, treatment, and rehabilitation services; cover immediate- and longer-term mental, physical, and social health needs; and coordinate pathways from primary care to referral to secondary and specialist services around individuals and populations (Gilson et al., 2008, 2011; Loewenson & Whitehead, 2012; Training and Research Support Centre [TARSC], 2014). It calls for leadership, communication, and facilitation capacities; for the decision-making space and resources to make and manage links with the public and other sectors; and for the political and social awareness and support for health personnel when they act on public health needs (Gilson et al., 2008; Loewenson, 2016; WHO, 2008).

Organizing Equitable and Universal Provision and Financing

In a universal system, everyone in a country should be able to access the same range of services on the basis of their need and pay for these services on the basis of their income. Universal health systems are generally publicly funded largely through general taxation or mandatory insurance and provide care for free or at very low cost at the point of delivery. To support equity, wealthier (and relatively healthy) people cross-subsidize the use of health care by poorer people, who are also more

likely to be ill. This transfer of resources from wealthier to poorer groups in redistributive health systems can also assist to close gaps in income and living standards between poorer and wealthier groups (Mackintosh, 2007). Universality is a key goal of the UN Sustainable Development Goals, expressed as achieving universal health coverage (UHC)—a system in which all people and communities can use the promotive, preventive, curative, rehabilitative, and palliative health services they need, of sufficient quality to be effective, while also ensuring that the use of these services does not expose them to financial hardship.

How countries achieve this goal varies across settings and may change over time. Many countries have funded health from general tax revenues or national mandatory insurance, in large pooled funds that allow for both cross-subsidies and the range of promotion, prevention, and care approaches necessary for addressing the social determinants of health equity. Currently, new prepayment approaches are being explored to secure adequate public funding, including for health promotion, taking into account not only how progressive they are, but also how they will impact on the goods, services, and behaviors that affect health. Such innovative financing sources include taxes on financial transfers, air travel, high-sugar products, tobacco, and alcohol, which may also be earmarked for specific purposes. In situations where formal and informal private payments for health services have risen, especially when arising as a result of inadequate or falling public funding, there are challenges for how to organize private payments into the forms of pooled prepayment that support universal systems and health equity and that invest in areas of public health that have less immediately perceived personal benefit (WHO, 2005). Organizing PHC-oriented, equitable, and universal health systems is not only a matter of how money is raised. It is also affected by whether resources and commodities are allocated according to need and invested in addressing social and service barriers to coverage, and whether there are mechanisms, such as health technology assessment, that bias public spending toward services with proven cost-benefit (Loewenson & Whitehead, 2012). Moreover, it depends on whether there is a motivated, competent health workforce deployed to areas of health need, trained and supported to implement actions on SDH and health equity.

Addressing Barriers to Equity *Within* Universal Policies

As noted earlier, policies for universal coverage do not necessarily lead to equity of access or impact on the SDH that affect uptake of care. Additional measures may be needed. Some measures for such "leveling up" have already been mentioned, such as ensuring that services are free at point of care, either universally, or at least for specific disadvantaged or high-need groups; providing infrastructure and staff in under-served areas; and ensuring a strong community and primary care services system that provides accessible entry points for health promotion, early detection, and care and that coordinates use of other levels of the health-care system. There may be demand for specific subsidies or tax exemptions for commodities consumed or used more by disadvantaged groups, such as basic foods. However, not all of the interventions are technical or on the supply side. Many address determinants that discourage uptake of services, such as by working with indigenous health systems and local civil society organizations (CSOs); providing joint entry points or working as multidisciplinary teams across health and other sectors, together with local political and community leaders and volunteers; creating entry points for health actions in settings (e.g., schools, workplaces, communities, marketplaces, areas) that people frequent; and sharing information and building a more participatory culture to involve community members in decisions and actions on their priority health needs (Gilson et al., 2011; Loewenson & Simpson, 2014). **EXHIBIT 4-7** provides some examples of programs that address the SDH affecting equity in health systems.

Actions such as those described in Exhibit 4-7 can strengthen, or protect, equity-promoting features of health systems and are often connected with wider struggles for social justice and transformation (Rede Brasil de Direitos Humanos, n.d.). For example, wider struggles for democracy in southern Africa set the political basis for post-independence comprehensive PHC approaches (EQUINET SC, 2007), while struggles for participatory democracy in Brazil shaped the entitlements and governance arrangements established in the Brazilian universal health system (SUS) in 1998 (Cornwall & Shankland, 2008). A study of primary care in 31 European countries found that sustained social support for government's responsibility for welfare provision was important to enable the longer-term reforms that strengthened the comprehensiveness and equity of primary care services (Kringos, Boerma, Zee, Van der Zee, & Groenewegen, 2013). In the United Kingdom, a redistributive publicly funded National Health Service has been sustained over more than 60 years by social and cross-party support (Pennington & Whitehead, 2014). In in other countries, rights-based approaches have been used to clarify state duties and public entitlements and responsibilities (Loewenson & Simpson, 2014).

EXHIBIT 4-7 Addressing the SDH That Affect Equity in Health Systems

In Spain, the *Health Promotion among Navarre Ethnic Minorities* program aims to reduce health inequities by improving the health of the Roma community in one area; the Roma are one of Europe's most disadvantaged social groups. Efforts are being made to design and implement prepayment schemes and to raise awareness with the Roma community about their health entitlements. Issues raised by the Roma community are integrated into training for health workers. The health information system carries out health equity surveillance, including for Roma populations, and the evidence is used to review and share practices across sectors and across countries. People from within the Roma community are trained as mediators and act as peer educators and as a liaison between the community and the central health, social, and education services. These mediators play a key role in documenting the health history of families in the area covered and drawing up a health plan in cooperation with the appropriate service providers. They highlight areas in which they need more information or education, as identified through the Roma associations—for example, on outbreaks of communicable diseases or issues related to lifestyle, life transitions, chronic diseases, and prevention. Staff from the relevant agencies meet once a year to incorporate additional items into the annual training program (Perez Jarauta, Goya Arive, & Merino Merino, 2010).

In England, *health champions*, local people trained by a U.K. nongovernmental organization, work with others in their communities to improve health. Health champions work with primary care personnel on health actions in the community; visit schools and community centers and bring local community members to health meetings; link local networks, knowledge, and experience with the practice knowledge and resources; use culturally relevant methods to gather local people to work with general practice staff; and explain to other patients how to make best use of the facilities and services provided. Health champions are reported to have improved input to local commissioning decisions, to have set up social and support groups for young mothers and for people with chronic conditions, to have improved individual and community literacy on prevention initiatives such as immunizations, and to have supported the use of appointment guides and other practice tools, especially for those persons for whom English is a second language. Their work is reported to have led to service and quality improvements (Pennington & Whitehead, 2014).

In rural Guatemala, *promotoras* (community health workers) have supported health service access and uptake. Low literacy rates, high rates of poverty, and gender inequalities impact on the uptake of child health services by indigenous Mayan women in the highlands of Guatemala. Indigenous leaders within the community—usually women who understand the local dialect and culture—have been trained as *promotoras*. The training is culturally adapted and recognizes local experience and environments. It includes space for the community to teach the teachers, such as on knowledge about the use of folk medicine to treat common ailments, while the *promotoras* are given information and skills to improve and manage basic health issues. An evaluation found these community health workers have had a positive impact on health and well-being, and the authors suggested that this experience may be relevant to meeting the health needs of immigrant populations in the United States.

Another paper describes how a needs assessment by the Organization for the Development of Indigenous Maya found diabetes to be a priority concern in indigenous communities. In 2012, the Organization for the Development of Indigenous Maya worked with community health workers trained as diabetes health promoters linked to primary care. Each worker had a caseload of 15–20 patients with whom they held a weekly diabetes club meeting, providing the attendees with information on self-management and group exercises. The community health workers also conducted weekly home visits and preconsults in the clinic to monitor each patient's progress and discuss specific challenges, create goals, and tailor exercise and nutrition plans. A significant decrease was found in mean blood glucose levels between baseline and follow-up in the people involved (Amerson, 2013; Micikas et al., 2015).

In recent decades, neoliberal globalization-related economic policies have dominated health system development in many countries. Given a more individualistic and consumerist society combined with conflicting political forces and values, these policies have challenged, fragmented, and weakened the universality and equity of health systems (Gilson et al., 2011; Mackintosh & Koivusalo, 2005). Thus, acting on SDH increasingly demands action to build public, professional, and political support, while also raising the role of social agency and power as a social determinant of health equity, as discussed further in the next section.

Social Exclusion, Social Agency, and Power as a SDH

As noted earlier in this chapter, people attain different positions in the social hierarchy, often characterized

by their social class, educational achievement, occupational status, and income level, or based on gender. Socioeconomic position can derive from the following sources:

- Resource-based measures, referring to material and social resources and assets, including income, wealth, and educational credentials, as well as poverty and deprivation
- Prestige-based measures, including the level and quality of access to and consumption of goods, services, and knowledge; occupation; and education (WHO, 2016)

Beyond resources and prestige, stratification arises due to discrimination on social factors such as race, gender, color, sex, language, religion, political or other opinion, national or social origin, or property (WHO, 2016). As social beings, humans need to feel valued and appreciated, and to derive meaning from and exercise a degree of control over their work and life conditions (AIHW, 2012). These drivers of social position are linked. Material deprivation generates social exclusion and stigma. Social power affects the ability people have to influence and make choices over health inputs and to use these inputs to improve their well-being (EQUINET SC, 2007). These factors and their role in stratification and power are summarized in **FIGURE 4-6**.

Social exclusion—and, conversely, social cohesion—has grown in profile as a conceptual lens through which to view problems of inequality, poverty, and disadvantage. Social exclusion has the following characteristics:

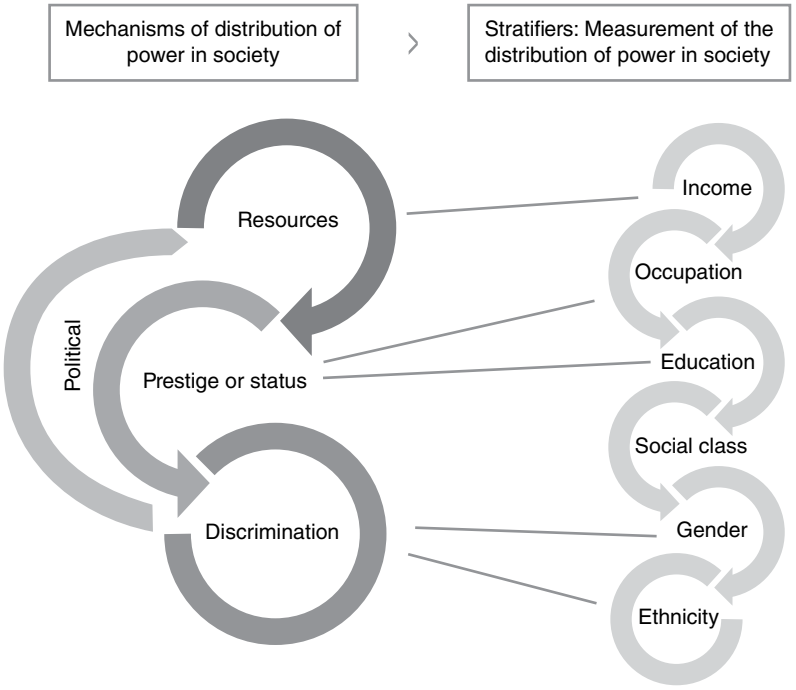


FIGURE 4-6 Mechanisms of distribution of power and their stratifiers.

Reproduced from WHO. (2016). *Innov 8: The Innov8 approach for reviewing national health programmes to leave no one behind*, Geneva, Switzerland: WHO http://apps.who.int/iris/bitstream/10665/250442/1/9789241511391-eng.pdf?ua=1 (accessed 22 June 2017).

DEFINITION

Social cohesion refers to the mechanisms and perceptions that exist in a society regarding social integration across various differentials and for confronting discrimination. It affects the sense of belonging within society, together with features of trust, participation, and reciprocity (WHO, 2016). This concept is applied differently in different regions, contexts, and communities. The European Commission's (2016) concept of "active inclusion" means that every citizen, including the most disadvantaged, fully participates in society. It means having adequate opportunity for work, support for income and employment, and access to quality services that enable active participation in society, including through investment in individuals' capacities and opportunities for participation. In contrast, in post-independence South Africa, social cohesion has been identified with nation, peace-building, and diversity in a democratic dispensation, as a response to past racism and inequality (Palmary, 2015).

- Multidimensional, encompassing social, political, cultural, and economic dimensions, operating at different social levels. and potentially reflecting violation of social rights to participation and inclusion
- Dynamic, impacting populations in different ways and to differing degrees at different social levels over time
- Relational, focusing on exclusion as the rupture of relationships between a group of people and the wider society (Popay et al., 2011)

On the one hand, "exclusion" may be seen as a less stigmatizing label than "poor," and to make clearer links to concepts of social rights and justice. On the other hand, a focus on social exclusion can place too much emphasis on the social inadequacies of specific groups of people, rather than on the environments and processes that generate poverty and inequality (Popay et al., 2011).

The CSDH framework shown in Figure 4-1 positions social cohesion as an SDH that is structural, generating social differentials, and intermediary, as a condition. Popay et al. (2011) argue that applying an equity lens to the analysis of and response to SDH is more effective when social exclusion is understood in terms of processes that embed unequal power relationships and produce social inequalities, rather than as a static "state" in which particular groups are labeled as "excluded."

This perspective places inclusion and exclusion on a continuum that is a result of the distribution of social, economic, political, and cultural resources, capabilities, rights, and opportunities (Popay et al., 2011). This chapter points to various ways that such exclusionary processes manifest and affect health outcomes, such as in the transmission and response to the HIV epidemic in east and southern Africa (**EXHIBIT 4-8**), in slower and reduced access to services, in the impact of gender norms, or through precarious environments or displacement of people by large extractive industry projects.

Responding to processes and conditions that generate social exclusion is not a straightforward matter, and the responses are themselves affected by social determinants. In Colombia, for instance, people displaced by war were required to register with civil or military authorities as a means to access services, political rights, and social integration programs. Unfortunately, such registration also made people more visible, increasing their risk of discrimination and victimization and discouraging service uptake (Ojeda & Murad, 2005; Popay et al., 2011). Similar barriers to

EXHIBIT 4-8 Social Determinants in the HIV Epidemic

The HIV and AIDS epidemic illustrates the complex relationship between exclusionary processes and health disadvantage. To start with, the epidemic has been most prevalent in countries (particularly in sub-Saharan Africa) that already experience social, economic, cultural, and political disadvantages. In the initial stages of the epidemic in east and southern Africa, HIV transmission was associated with more mobile, urban populations and adult HIV prevalence was higher in wealthier, more educated, and urban groups. The transmission reflected social differentials, however, such as in power and income between older men and younger women. As a reflection of the gender inequalities described earlier in this chapter, the HIV prevalence in young women age 15–24 years was more than twice that of their male counterparts in southern Africa in 2003. The changes in both HIV transmission and survival reflected common social differentials: The distribution of prevention and treatment interventions followed typical social differentials in the region, with lower coverage levels among rural, poorer people with lower education levels, raising their susceptibility and vulnerability. By comparison, urban, wealthier people lived longer lives with HIV—sustaining their prevalence rates—due to their better nutritional status and uptake of services and greater mobility.

A number of studies revealed the negative impact on these SDH on the epidemic, with households affected by AIDS facing higher levels of falling household assets, income, social security, food consumption, and increased social exclusion. The stigma and discrimination associated with HIV, combined with other exclusionary processes, exacerbated social, economic, political, and cultural inequalities, interfering with prevention, diagnosis, and treatment. The initial social characterization of HIV/AIDS as a moral issue exacerbated a culture of silence and denial by people living with HIV because of fear of rejection and isolation. Household survey data from four sub-Saharan African countries with different epidemic patterns (Ghana, Kenya, Swaziland, and Zambia) showed that AIDS increased the share of the population living in income poverty, with the strongest effects occurring in those just above the poverty line, placing demands on and further eroding the community-level safety nets and social support networks on which disadvantaged communities depend on (EQUINET, 2012; Rispel, Molomo, & Dumela, 2008; Salinas & Haacker, 2006; Southern African Development Community [SADC], 2003).

uptake have been found in programs that profile conditions that have been associated with social discrimination (e.g., HIV).

State-led responses to social exclusion are generally organized through a range of approaches. For example, *universalist policies*, reflecting values of social solidarity, extend rights to publicly funded services to all citizens, such as publicly funded national health services that are free at point of care. Universalist policies may address SDH and processes that drive social exclusion, as, for example, is intended in the health-promoting schools described in **EXHIBIT 4-9**. As noted earlier, they may need to be complemented by specific measures to address barriers to uptake or coverage in different social groups.

A second approach is through *policies targeting specific social groups*. These policies are specifically designed for disadvantaged groups, and intended to remedy health disadvantages. In some cases, they take the form of c*onditional targeted transfers*, involving some form of testing so as to target scarce services and health resources to those groups most in need, or to incentivize particular behaviors. They include social approaches, as exemplified in Exhibit 4-7. Targeted transfers include conditional transfers of cash or services, or a link between the two. As an incentive for service uptake in Brazil's Bolsa Familia, for example, recipients meeting the means test for a cash transfer must also ensure that their children attend health centers and school. Whether providing cash and/or services, these measures can bring resources to local services providers and promote economic inclusion. They are also documented to have limitations, however. For example, they are often insufficient to provide sustainable pathways out of low-income living, may impose bureaucratic barriers for some persons

EXHIBIT 4-9 Universal Measures to Address Social Exclusion: The Healthy Schools Program in South Africa

In South Africa, schools have been a focus of community-based prevention of negative social factors—including social power imbalances, poverty, violence, and substance abuse—that place learners at risk, including for health problems such as HIV and AIDS. South Africa has the paradox of relatively poor health outcomes despite the country's relatively high aggregate gross domestic product and level of health expenditure. Inequalities in health stem from a history of racial and gender discrimination, a migrant labor system, and high levels of income inequality. The country has sought to address social, racial, and gender disparities and realize social rights through redistributive policies, largely in the social sector.

The percent of adults in South Africa without schooling fell dramatically from 18% in 2001 to 7% in 2010, although with continuing inequality in access to education by region and racial group. Western Cape province, for example, despite being the second richest province in South Africa, has high rates of poverty, economic inequality, unemployment, crime, commercial sex work in teens, and substance abuse, contributing to social exclusion and poor health outcomes. As one response to these conditions, education was seen by the community as a key asset supporting social inclusion and autonomy, promoting the sexual autonomy of girls, and reducing their risk of contracting HIV. Despite this, a large proportion of high school students drop out before completing secondary school and school-age pupils have been found to engage in cigarette smoking, drug and alcohol abuse, interpersonal violence, and unsafe road-related and sexual behavior.

Education policy in the country has sought to strengthen universal access to education, and to support schools as sites for reaching adolescents and helping them acquire health-promoting skills and orientation. Based on the national health-promoting schools policy and approach, a program was initiated Western Cape in 1995 in which program leaders worked with community leaders, school staff members, school health nurses, parents, and the community to make schools a healthy and health-promoting environment for young people. The program integrates broad health promotion and education services, and promotes individual and social well-being for those at school and their families and community. It tackles content issues affecting social determinants, such as road safety, personal hygiene, substance abuse, HIV, and nutrition, and develops and supports alternative youth (and youth-led) activities such as teenage clubs, cultural activities, and camps. It incorporates local social and cultural resources, integrating local and indigenous knowledge, mentoring, and storytelling.

McNab (2013) found that this type of work depends more on a change in mindset rather than on the provision of major new resources, and on making a link between the "top down" influence of public policy and the "grass roots" approach of constructive community engagement. The social dimensions of the Western Cape program were seen to be a central element of its success: At a 2006 national conference, many presenters and participants highlighted the need to listen to and strengthen all voices in the program, and particularly the marginalized voices of local communities and indigenous wisdom, and of the learners themselves (Loewenson, 2013b; University of Western Cape, 2006).

with high health needs, and, as individual-based strategies, may have limited impact on promoting sociopolitical rights or building more inclusive political cultures (Lauthier, 2005).

A third strategy, *market approaches*, seeks to use private or state subsidies to support choices in the consumption of services by poor people to address economic or social barriers to such choices for the most marginalized households. For example, disadvantaged groups may obtain subsidies in insurance schemes to support their access to services or subsidies to address health determinants such as housing. This kind of market support is argued to face similar barriers as those detailed in relation to targeted transfers, and may potentially lead to poorer-quality services for subsidized groups and neglect deeper determinants of such disadvantage (Popay et al., 2011).

Whatever the overall approach, as the example in Exhibit 4-9 suggests, involvement of the people who are the intended beneficiaries of policies and actions that aim to reverse exclusionary processes is essential, as both as a resource and as a right. This implies building a deeper understanding of social power and agency as both determinants and key dimensions of successful policy and action. It suggests that actions that seek to tackle exclusionary processes should strengthen the power and capacity of those affected to engage in meaningful participation, while addressing resistance from professional workers who see it as a challenge to their technical power.

The demand for social participation, power, and agency in health has grown with increased social literacy and with democratic pressures for people's values and preferences to be incorporated into public decision making. The CSDH (2008) identified participatory approaches as a critical feature of health system action or intersectoral action for health to tackle SDH. Social participation has been integrated as a demand-side "intervention" to address barriers to access and acceptability of services and as a means to involve people as active participants in their own health care (Mittler, Martsolf, Telenko, & Scanlon, 2013). As a pragmatic measure, participation is seen to contribute to a range of areas necessary for the effective functioning of health systems:

- Community health literacy, public information, and the use of community knowledge on health
- Assessment, identification, and prioritization of health needs
- Health planning and decision making on spending priorities, resources, and budgets
- Policy deliberation and formulating strategic policy direction
- Co-design, coordination, and coproduction/implementation of health actions
- Oversight, monitoring, evaluation, review, and improvement of services, actions, and quality, including measures of social outcomes (Cornwall, 2008; Gilson et al., 2008; Loewenson & Gilson, 2012; Loewenson et al., 2014)

In a pragmatic and sometimes depoliticized approach, people as consumers or as volunteers may have minimal control over policies affecting structural determinants, but assume responsibility for implementation of compensatory actions. There may be limited change in current norms, processes, and mechanisms. While people may exercise oversight by monitoring, reviewing, and holding services accountable for their performance, they may not have the power to make the decisions that shape these services and may as volunteers take on unpaid burdens of care (Loewenson, 2016).

Participation has also been a product of political drivers, including through the actions of popular movements that have generated demand and sometimes conflict around health and around codetermination and accountability (Amar Amar & Echeverria Molina, 2008). As noted earlier, participation is viewed as the right to guide the duty bearers (states) in their implementation of other human rights. Citizens, including those engaging as health activists or in social movements, have pursued these rights through collective demands for change in areas of power, discrimination, beliefs, policies, and practices that are perceived as inequitable or as limiting their well-being, whether in relation to access to services or the SDH that affect the health of different social groups, or as a claim for shared decision making and oversight (Cornwall & Leach, 2010; Loewenson, 2016). The power, voice, and agency of affected groups can be found in invited spaces—created by external actors and in which people are invited to participate—or in *claimed/organic spaces*—created by people themselves, often as collective and popular self-organizing spaces around a common cause (Cornwall & Leach, 2010).

There are thus different conceptualizations of participation, taking different forms that express the different power relations that exist and the control that people have over events and resources. For example, a group may exert power "over" another group (such as when men control women's use of reproductive health services), may claim the power to be part of processes (such as when people living with HIV participate in price setting for treatment), may exert a shared power "with" each other (such as when communities organize as a group to engage service providers on delivery), or

may build a shared consciousness of themselves and their conditions in the form of a power "within" (such as when youth identify and call for changes in the way mental health services are organized to reflect their needs). In participatory processes where communities express their own collective voice and agency to an increasing degree, they build shared power as well as a consciousness and confidence to act. In contrast to fears that this paradigm diminishes the power of technical actors, it can be shown to create complementary forms of power that enable positive change, especially for community-level and frontline services, as exemplified in the case of Zambia described in **EXHIBIT 4-10**.

A growing number of studies using experimental or quasi-experimental designs in LMICs have demonstrated the potential of participatory interventions to generate health and health equity gains (Pronyk et al., 2006; Wallerstein, 2006). It is not easy, however, to be prescriptive about the specific mechanisms and measures to achieve these changes, as they are context

EXHIBIT 4-10 Social Participation and Power in Improving SDH in Lusaka, Zambia

Zambia is a lower-middle-income country; Lusaka is its capital city and home to 1.7 million people. The Lusaka District Health Office (LDHO) is the local health authority within Zambia's Ministry of Health. For more than a decade, LDHO has used participatory reflection and action (PRA) approaches to strengthen health services delivery at the primary care level and the role of neighborhood health committees (NHCs) as a more formal, sustained space for participatory interactions. As a matter of policy, NHCs involving elected community members are expected to identify and voice community health needs and to support information exchange between health services and communities.

In 2005, however, planning and budgeting for PHC was not participatory, and these committees were often poorly organized or not effective in their role. Growing public health problems in Lusaka motivated community members, health workers, and LDHO to address these shortcomings. Various participatory tools were used to identify needs, system barriers, and areas for change, and to build a shared identification of problems and actions to remedy them across health workers and community members. Information sharing between health workers and community members increased, community members became more confident in approaching health workers for information, and health workers provided the needed information to them for planning and resource allocation purposes. A range of actions were taken, including a cholera prevention strategy that led to significantly fewer cholera cases and deaths than in previous years, despite the heavy rains. By the end of this first phase of work, the community members had developed greater confidence and enthusiasm in seeing the process move forward. Likewise, health workers noted the benefits of improved communication with the community, though some feared losing power.

In 2008–2009, the same PRA approaches were used to scale up the health system by establishing new health centers in the city, while consolidating and building capacities to institute the approaches in existing ones, and building partnerships with a range of other sectors in areas such as clean-up campaigns, solid waste management, healthy food production, and services performance and uptake. These efforts were complemented by a range of approaches to organize and use community evidence. Photovoice (community photography), for example, presented evidence on community concerns with blocked sewers, health facility corruption, poor water supplies, and solid waste management during clinic and district health meetings held to trigger actions to resolve these problems. A wider community-level health literacy process with Lusaka residents drew evidence from their lived experience, facilitating community-level diagnosis, providing relevant health information, and stimulating action and engagement with health systems, in partnership with service providers, civic leaders, and the local authority. These approaches positioned the communities as change agents. Every three months, those working on the action plans met to review their work, reflect on their experience, and evaluate their progress.

Pre- and post-intervention surveys and reports from health workers and communities about the initial 2006–2008 PRA work found improved interactions between health workers and community members and partnerships with other actors, and increased confidence of community members in providing inputs to planning processes in support of action on health. There is also some evidence of positive health outcomes associated with this work. For example, the clinic catchment areas where these activities took place had no cases of cholera in 2012–2015, at a time when other areas still reported cases. While it is not possible to solely attribute the positive outcomes to the PRA program, the removal of waste sites that were once a site of disease and the measures to increase community health literacy and action were perceived to have made an important contribution. The health literacy work and the engagement with neighborhood health committees is now being expanded beyond Lusaka in a national scale-up in light of the positive outcomes from the work done in Lusaka District (LDHO & TARSC, 2016; Lusaka District Health Management Team &TARSC, 2015; Mbwili-Muleya, Lungu, Kabuba, Zulu Lishandu, & Loewenson, 2008; Ministry of Health Zambia, LDHMT, & TARSC, 2012).

dependent. Many of the approaches are triggered by or involve action outside the health system, in other sectors and organizations, and within the community itself. Various features of health systems have, however, been found to support the spaces and shifts in social power that enable effective participation. These approaches include the following:

- Setting health interventions in sites that are familiar to communities, such as markets or schools
- Integrating community mapping, monitoring, and preferences in health planning
- Producing accessible information (such as through newsletters, meetings, and social media) that shares local experience and responds to perceived needs
- Using socially appropriate and participatory methodologies that build on and validate local experience and knowledge
- Involving and supporting community-elected and -located community health workers to strengthen communication and linkages between health systems and communities
- Providing opportunities for dialogue on community perceptions of services, through community audits, health watches, community councils, participatory research, and legal action
- Devolving meaningful budgets to lower levels of the health system to facilitate and support social roles
- Enabling communities to shape the "rules of the conversation" by giving them the ability, resources, and opportunity to define the terms and processes in which they participate and the issues they want to address, and to have input into national laws and policies
- Investing time and resources in, and providing management support for, health worker competencies and incentives for participatory processes (Benequista, Gaventa, & Barrett, 2010; Cornwall & Leach, 2010; Gilson, 2007; Loewenson, 2016; Loewenson et al., 2014; Mbwili Muleya et al., 2008).

Implementing such measures depends on wider contexts, including the political freedom, spaces, and capacities that groups have to voice their views and to exercise influence; the constitutional, legal, and policy frameworks, social organizations, and systems that enable participation and support an active and democratic culture; the sociocultural conditions and relations within communities and families; and a comprehensive understanding of health and well-being that integrates and is accessible to society (Arenas-Monreal, Piña-Pozas, & Gómez-Dantés, 2015; Loewenson, 2016). These factors are not just shaped at local and national levels. That is, the local SDH and the social, institutional, and technical processes that affect them are increasingly influenced by global processes, and by decisions made by global and international actors. This topic is discussed in more depth in the next section.

Global SDH, Local Impact: Experiences of Extractive Industries in East and Southern Africa

As noted earlier, globalization is transforming many of the social conditions that have the greatest impact on health, while also affecting the public policies and programs that target these determinants. Many of the elements identified within the socioeconomic and political context in Figure 4-2 are now influenced by global policies, institutions, and processes.

A review of international evidence on globalization and SDH identified that while globalization has both positive and negative impacts on health, its benefits are unevenly distributed (Labonte & Shrecker, 2008). In this review, the benefits were reported to have been primarily obtained by countries and communities that already had financial, land, physical, institutional, and human capital assets. While globalization has certainly brought opportunities, such as the flow of information and new global funding to address health issues, the rules of current forms of globalization were found to favor already rich countries and people, which have greater resources and power to influence their design (Labonte & Shrecker, 2008; World Commission on the Social Dimensions of Globalisation 2004). In sub-Saharan Africa, globalization policies were associated with debt crises, capital flight, and structural adjustment programs that were reported to have contributed in part to increased inequality in health and reduced gains in life expectancy (Labonte & Shrecker, 2008). Examples of the mix of positive and negative outcomes are shown in **EXHIBIT 4-11**.

With the growing impact of global determinants on health, an analysis of globalization and SDH merits deeper attention than a subsection of a chapter. The distribution of positive and negative impacts of global determinants of health is further discussed in the *International Trade and Health*, and *Global Health Governance and Diplomacy* chapters in this text, and

EXHIBIT 4-11 Globalization, Women's Occupational Roles, and Nutrition in Sub-Saharan Africa

A 2010 review of 62 peer-reviewed publications on pathways between globalization, women's occupational roles, and nutrition found a mix of positive and negative outcomes, although with greater documentation of negative outcomes for women's health and nutrition.

At the global and national level, 5 papers reported largely positive outcomes for SDH, with globalization-related innovations in technology and information, global normative commitments, and improved incomes, markets, and opportunities for improved diets from investments where women are involved in export-oriented farm production. However, 24 papers reported largely negative outcomes, with trade liberalization, terms of trade, market access, and value-added food chains largely benefiting existing wealth and foreign corporations; gender inequity in access to investment, technology, and extension services due to "gender-neutral" trade policies; shifts to imported foods and processed foods displacing local food crops with livelihood losses for women; and diminishing healthcare resources raising the costs of care.

At the community and household levels, 7 papers noted largely positive impacts in terms of improved returns from new technologies in crop yields, increased nonfarm employment, and information flows that challenged gender-related discrimination improving women's conditions, with positive returns for household nutrition. Here, too, however, the larger number of papers (34) pointed to negative impacts at the community and household levels from determinants related to global trade, arising due to weak consideration of existing gender inequities in access to and control over production inputs in policy and program design; insecurity of and poor wages in new forms of employment; decreased time for household needs and neglect by women of their own well-being to meet time and resource demands; and decreased availability of local foods and increased consumption of higher-cost and poorer-quality imported and processed foods (Loewenson, Bambas Nolen, & Wamala, 2010).

so is not covered in detail here. This section explores global drivers of local SDH and the responses to them through the example of the extraction of mineral resources from Africa, largely through global corporations and interacting with global capital markets. It also highlights how global and regional norms can be positively applied to address negative health outcomes.

In 2009, Africa's oil, gas, and minerals exports were worth roughly five times the value of international aid to the continent ($246 billion versus $49 billion), with African oil and mineral resources sought after by high- and middle-income countries, including the emergent economies of China, Brazil, and India (Loewenson, Hinricher, & Papamichail, 2016). A surge in demand led to exploration and development of many new mining sites and new agreements exchanging investment in infrastructure for mining rights to oil, coal, and various strategic minerals (Besada & Martin, 2013; Shelton & Kabemba, 2012). As a consequence of this activity, African countries are increasingly engaging in global markets. At the same time, the rewards from these ventures are reported to be skewed toward those countries and individuals who have existing economic power (Birdsall, 2005). African countries that are rich in these mineral resources have experienced high levels of inequality and poverty—often referred to as "the resource curse" (Global Witness, 2012). A literature review of the health and SDH related to extractive industries in east and southern Africa found that they are significant economic actors in the region, but with limited forward or backward linkages into the national economy and limited job creation outside the sector, unless specifically stimulated (Loewenson et al., 2016).

This finding draws attention to how extractive industry (EI) activities, as a determinant of health, meet their responsibilities to protect against harm to health and make fiscal contributions to health care for the populations who work in or are affected by them. Corporations have duties to assess the potential risks of their activities to workers and surrounding communities, and to prevent and manage these risks. In recent years, more attention has been given to environmental risks associated with EIs, to EI company's duties to populations displaced by their activities, and to the duties that they have to inform and enable co-determination and participation of those affected in making the decisions that affect their health (International Labour Organisation, 2014; Murombo, 2013).

The health benefits of EIs have largely come from employment, income, and some service provision, generally for those persons who are directly employed by the mining companies and their families, and from EI fiscal contributions to public services. Nevertheless, tax exemptions often reduce their contributions to social funding. EIs also bring risks to health: They have been found to lead to accidents, hazardous

working conditions, environmental hazards, degradation of ecosystems, poor environmental health infrastructure, and displacement of local people—all of which increase the risk of disease. The literature reports the spread of communicable diseases (such as tuberculosis) and cholera epidemics from poor environmental health infrastructure, sexually transmitted diseases, and HIV in communities surrounding the mines (Aaboe & Kring, 2013; Catholic Relief Services [CRS], 2011; Loewenson et al., 2016).

Poor communities living around the mines are reported to be particularly vulnerable to pollutants, given their poor living conditions, with less recognition or monitoring of their risks. They are also least able to obtain reliable information on these risks, or to register their concerns with decision makers (CRS, 2011; von der Goltz & Barnwal, 2014). Mining is associated with (sometimes forced) displacement of communities, with reports indicating several thousand families having been resettled to facilitate mining in some countries (Global Environment Facility [GEF], Open Society Initiative for Southern Africa [OSISA], & United Nations Development Programme [UNDP], 2013; Human Rights Watch [HRW], 2013). As a health determinant, poor planning of such displacement is reported to have led to loss of livelihoods; loss of access to water, flora, arable land, and pastures for livestock for these communities; and substandard living conditions. Affected communities were poorly consulted on plans discussed with government, making it difficult for them to hold companies accountable when they did not deliver on commitments, and generating social frustration (HRW, 2013). For those living near mines or displaced by mining, exclusion from decision making on measures to address the health and social impacts, noted in many reports, is reported to have led these groups to "bear a disproportionate share of the costs of mineral development without adequate compensation, and to receive an inappropriately small share of the economic and social benefits" (International Institute for Environment and Development, 2002, p. 208).

As a key structural determinant, the policies and actions of the state play a vital role in protecting the health of populations caught in the midst of these global processes. For the health sector, this calls in part for public health leadership to leverage cross-sectoral health-promoting actions to manage health risks. **EXHIBIT 4-12** overleaf tells the story of one such response—that is, the experience of Tete in Mozambique. Other examples exist of state action in the east and southern Africa region, including using environment and health impact assessments to assess health risks, implementing audits to review the performance of EIs as in Zambia, and setting contract/lease renegotiations or renewals to review, require, or impose new obligations in Democratic Republic of the Congo, Zimbabwe, South Africa, and Mozambique (De Backer, 2012; GEF et al., 2013; Kabemba, 2014). While EI contributions for health and social welfare are generally identified as areas of voluntary corporate social responsibility (CSR), Democratic Republic of the Congo introduced a micro-levy on EIs in September 2014 to fight chronic malnutrition; in the same year, Zimbabwe lifted the exemption on the sector that had allowed EIs to avoid contributing to the AIDS Levy Fund (Loewenson et al., 2016).

As multinational EIs are themselves global actors, managing health in conjunction with EIs also calls for global governance arrangements to recognize and integrate social obligations for addressing SDH. At the global level, there are now numerous international standards, codes, or guidance documents related to the practices of EIs and multinational enterprises. These exist at the UN multilateral level, from Organisation for Economic Co-operation and Development (OECD) countries, at the African Union level, from financial institutions, and in CSR standards developed by international business and by civil society. Some regions, such as the Economic Community of West African States (ECOWAS) in West Africa, have moved toward a "strength in numbers" approach, with efforts to harmonize laws relating to EIs, including laws targeting health and social protection at the subregional level (Loewenson et al., 2016). The inclusion of these global norms as legal provisions in east and southern Africa and their strengthened implementation and oversight with accountability at the global, regional, and national levels, including in the source countries of EIs, have been identified in the literature as important measures to address the local health impacts of such global-level determinants (Loewenson et al., 2016).

The focus on EIs in this section presents one example of global-level SDH. Many others could be cited, including in the various dimensions of trade and health and climate change discussed in other chapters in this text. The experience in Tete described in Exhibit 4-12 and that of other countries in east and southern Africa in managing these challenging contexts suggests that while local communities and services may be organizing cross-sectoral responses to such global drivers, more attention is needed to determine how best to ensure that global standards are enforced, including by transnational, private actors.

EXHIBIT 4-12 Managing SDH in the Context of a Growing Extractive Sector: Tete, Mozambique

Tete province's geographical positioning, significant natural assets, and influx of investment, people, and capacities offer opportunities for addressing health needs in this region. Investments in hydroelectric power and coal mining have dominated recent investments in Tete, such that the province had the second highest provincial GDP growth in Mozambique between 2002 and 2008, largely driven by extractive industries. While these investments have brought economic benefit, they have also placed demands on infrastructure and services, led to displacement of people, and increased risks of injury, ill health, and social risks from displacement, harmful alcohol use, and road traffic accidents. Within the context of the communicable, reproductive, and other health burdens in the province, there is evidence of high social deficits in the very districts that host these large projects. The population in Cahora Bassa, a district with significant new power projects, has poorer health status outcomes than the populations in districts with less investment; Moatize, where the mines are concentrated, has fewer health service resources and poorer coverage outcomes; likewise, Tete City, a richer area and the focus of economic activity, also has poor health status outcomes. Tete City, Cahora Bassa, and Moatize had calorie-adjusted poverty rates in 2007 that were higher than those in most other districts in the province, including districts that were worse off than these three districts in 1997.

These findings raise the question of how such economic investment can be better linked to improved population well-being. One route for managing these SDH is through upstream entry points, linking EI activities to value-added local production, including in small and household enterprises, and providing links to their markets and infrastructures. Further upstream entry points include improved corporate fiscal and other resource contributions for public spending on education, health, and other social services, and the integration of health and social impacts in environmental impact assessments prior to granting a mining license.

Within this context, an assessment of the SDH, opportunities, and deficits carried out in 2014–2015 with the Tete provincial department of health (DPS Tete) identified key areas of action for the five-year provincial health strategy, including the following:

1. Closing the gap in improved household living conditions implemented through an intersectoral mechanism; ensuring all schools and health centers are connected to the power network; promoting small enterprise production of clean technologies for household cooking and energy; earmarking budgets and levies from local businesses for rapid improvements in rapidly growing settlements, such as Tete City and Moatize; and monitoring risks, interventions, and health impacts.
2. Promoting and protecting worker and community health in production activities, with measures outlined to promote occupational and community health and protect against production-related risks, including through environment, social, and health impact assessments; involving health personnel in oversight of corporate duties related to resettlement; and integrating company roles and resources in health services and activities in their districts.
3. Coordinated multisectoral strategies to improve food security and nutrition.
4. Expanding and improving equity in access to PHC, and positioning the health sector as a key redistributive sector in inclusive development, including through entry points in schools and workplaces, through private-sector contributions to healthcare services, and by enhancing community health literacy.

The province is now building support within other sectors for wider cross-sectoral implementation of these actions under the leadership of the governor of the province and with representatives of sectors at the national, provincial, and district levels (Loewenson & Simpson, 2015; Ministério da Planificação e Desenvolvimento, 2013; Ministério de Saúde, Mozambique, 2014; Republic of Mozambique et al., 2016; Salvucci, 2014).

Evaluating Action on the Social Determinants of Health and Health Equity

This chapter has presented evidence on the diversity of SDH and on the analysis and actions being applied to tackle them, including measures to address health equity. WHO's CSDH was a culmination of longstanding efforts by policy makers, state officials, academics, practitioners, and civil society to improve knowledge and evidence on the nature and extent of social determinants of health equity and to identify and implement policy or program actions to tackle them (Simpson, Kelly, & Morgan, 2013). Evaluation of such interventions has been important to gather evidence and build learning from their implementation about options to address the social determinants of health inequities; one such evaluation is described in **EXHIBIT 4-13**.

One of the recommendations of the CSDH to "close the gap in a generation" was to "measure the

EXHIBIT 4-13 Evaluating Interventions to Reduce Health Inequities

A six-year program was implemented in the Netherlands in the late 1990s to systematize learning from policies and interventions to reduce health inequities. Twelve studies were commissioned to evaluate a range of interventions to tackle SDH known to contribute to health inequities *and* to reduce exposure of lower-socioeconomic groups to these factors (e.g., poorer working conditions). Methods used ranged from observational to quasi-experimental to experimental studies. Evaluation of seven interventions gave positive results, with the researchers finding an improvement in at least one health outcome and/or an intermediate outcome. For example, a rotation of tasks among dustmen (i.e., garbage collectors) reduced physical workload and sickness absenteeism. Overall, the program sought to contribute to an evidence-informed approach for developing a national response to health inequities. It recommended that evaluation studies using a variety of approaches be embedded in all future interventions on health inequities (Stronks & Mackenbach, 2005; Simpson et al., 2013).

DEFINITION

A *realistic evaluation* approach combines assessment of design, process, and outcome and attempts to provide answers not only about which interventions work to address SDH, but also how they work and in which context(s) (Kelly et al., 2007).

problem, evaluate action, expand the knowledge base, develop a workforce that is trained in the social determinants of health, and raise public awareness about the social determinants of health" (CSDH, 2008, p. 20). This includes ensuring routine local, national, and international monitoring of SDH and health equity; investing in and building capacities to generate and share new evidence on the relationship between SDH and population health and health equity, and on the effectiveness of interventions to tackle these SDH; and raising public awareness on the findings (CSDH, 2008).

To date, a significant body of evidence has been developed on the "problem," on the link between specific SDH and health outcomes (particularly for intermediary SDH), and on the associations between SDH and health equity. By comparison, progress has been slower in evaluating the effectiveness and equity impact of interventions, and in generating attributions of impact to health or health equity and structural SDH. In part, this is due to the complexity of the methods used and the evidence needed to assess the commonly multicausal and context-dependent nature of the interventions for the "problem." While there is increasing recognition of the value of implementation and health systems research using new methodological approaches, there has been a continued predominance of a biomedical paradigm in health research and the use of traditional evidence-grading systems to value evidence. This has led to intervention studies on SDH, and the qualitative methods used to understand contexts for and impacts of complex interventions on SDH, often being classified as "unscientific" and "value-laden," despite their reliance on accepted social science methods (Baum, 2010; Bonnefoy et al., 2011). That this situation is changing is evidenced, for example, by the growing efforts to measure action on SDH among and within countries, by the documentation of country-level action on SDH, and by the Cochrane Equity Methods Group guidance on explicitly addressing equity in systematic reviews (Campbell & Cochrane Equity Methods Group, 2017; O'Neill, Tabish, Welch, Petticrew, Pottie, Clarke et al., 2014).

The CSDH Measurement and Evidence Knowledge Network emphasizes that no single approach to the generation of evidence should be favored over others and that evidence should not be appraised on the basis of adherence to a single evidence hierarchy or method (Kelly, Morgan, Bonnefoy, Butt, & Bergman, 2007). Health impact and equity impact tools, such as those outlined earlier, and more comprehensive approaches such as "realistic evaluation," support evidence on SDH and health equity because they move beyond simply estimating the likely effect of policy, to constructing logic models that show the implementation chain from policy to implementation in practice (Kelly et al., 2007). A range of methods and evidence enable practitioners to understand the complexity of mechanisms and the multifactorial nature of determinants that apply in practice and can make explicit the theories upon which programs are based. Realistic evaluation, for example, helps to capture the linkages between the context (the necessary conditions for an intervention to trigger mechanisms), the mechanisms (the aspects of a particular intervention that lead to a particular outcome in a given context), and the outcomes (the practical effects produced by causal mechanisms being triggered in a given context) (Kelly et al., 2007).

As argued earlier in this chapter, evidence and analysis contribute to the building of the relationships and complexity of interventions often needed to address SDH, including for intersectoral action and HiAP. Having an explicit and shared conceptual framework and theory of change at the outset helps to clarify the pathways for change, to identify shared outcome measures for assessing performance and impact, to prioritize action, and to test the thinking, thereby informing subsequent IAH work (Loewenson, 2013a). Analytic frameworks that explicitly include the assessment of equity in the design, implementation, and assessment of outcomes, while important, are not always present in evaluations of intersectoral action on SDH (Shankardass, Solar, Murphy, Greaves, & O'Campo, 2012).

The South Australian program on HiAP described in Exhibit 4-2 has, for example, taken a two-phase approach to evaluating its work: The first is an evaluation of perceptions of, interests in, and processes for HiAP implemented through interviews with senior South Australian public servants from the beginning of the initiative and continued in tandem with the health lens analysis. The second is a more comprehensive approach implemented in three overlapping stages over a five-year period (2012–2016) and using mixed methods, including test of a logic model of and theories of change for the work, shown in **FIGURE 4-7** (Baum et al., 2014).

Semi-structured interviews were implemented with key political figures, officials, and other actors who have knowledge of the HiAP process to generate knowledge to support its development and implementation. In the second stage of the evaluation, the theories developed in the first stage were tested, and the program logic and practice of implementation of the HiAP examined in eight health lens analysis projects, with detailed analyses in two of these health lens analyses and online surveys of policy actors administered regularly over the five-year period. A final stage synthesized the evidence to produce transferable knowledge and disseminate findings in annual research forums for peer review and discussion (Baum et al., 2014).

The features of the South Australia evaluation, which included multiple methods, stages, and areas of

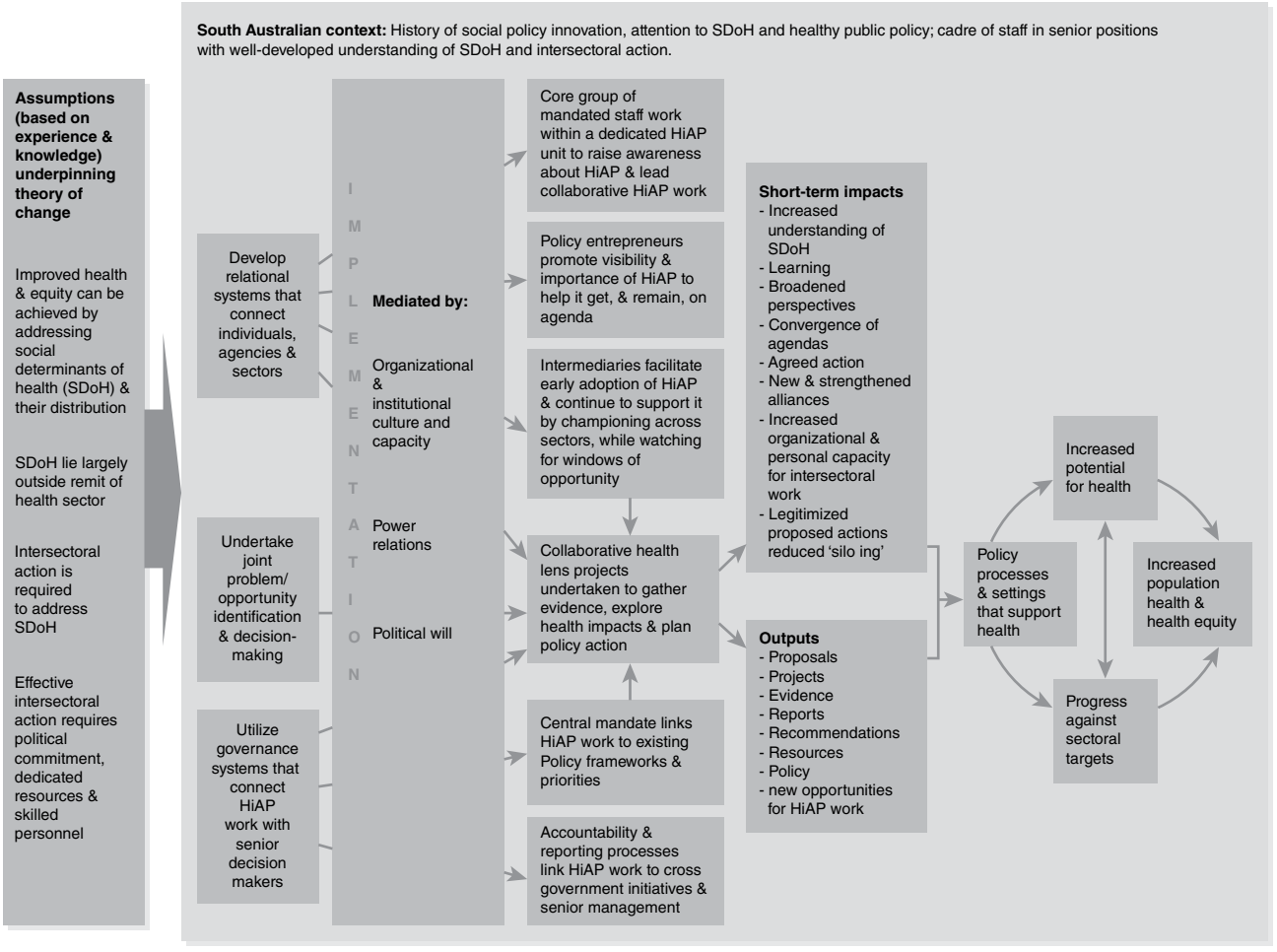


FIGURE 4-7 A preliminary program logic approach for evaluating HiAP in South Australia.

focus in the evaluation and multiple forms of evidence generated to support strategic reflection, are reflected in other processes for evaluating SDH. WHO has developed Innov8, an eight-step analytic process to be undertaken by a multidisciplinary review team to identify improvements in program performance and actions on SDH, health equity, and gender equality, as well as progressive realization of universal health coverage and the right to health. Innov8 applies realistic evaluations and integrates areas of human rights, health systems strengthening, and gender into its assessments. It builds a theory of change into the review analysis, including how and why interventions and activities are expected to produce results in these areas (WHO, 2016). A series of Innov8 publications and resources, including country case studies, are available that demonstrate application of this approach in areas such adolescent sexual and reproductive health, maternal and child health, and cervical cancer screening (WHO, 2016).

Routine collection, review, and reporting of data on SDH are as important as specific evaluations for effecting changes in and ensuring accountability of policy commitments made. Further, while increasingly sophisticated methods are available for measuring and analyzing social gradients in health and the impact of interventions, actually *using* such information in policy and stakeholder forums is critical to realize improvements in SDH and health equity. WHO (2013b) provides a resource for ministries of health to support the monitoring of social determinants of health equity, through a series of eight online presentations and a link to the Health Equity Assessment Toolkit (HEAT; a software application that facilitates assessment within and across countries using available data).

Equity Watch reports provide a further example of use of evaluations—in this case, analyzing routine, household, and available community data on SDH and health equity in Africa (**EXHIBIT 4-14**). Almost all ESA countries have made policy commitments to promote equity in health. In 2007, the Regional Network on Equity in Health in East and Southern Africa (EQUINET)—a network of professionals, civil society, state, parliament, and academic actors promoting health equity—analyzed and reported on health equity in the region, building a cooperation with the east, central, and southern Africa (ECSA) health community to gather and report on evidence on health equity (Rasanathan, 2011). In 2012, an "Equity Watch framework" of 25 priority indicators of health equity was used to organize evidence from 16 countries in east and southern Africa, complemented by Equity Watch work in countries (EQUINET, 2012). National teams in six countries in the region involving state and nonstate actors and working with EQUINET organized, analyzed, and presented quantitative and qualitative evidence to assess progress in addressing health inequities, to evaluate social determinants and health care, and to inform social dialogue on proposals for strengthening health equity. The national reports have been used to monitor implementation of equity in the implementation of the national health strategic plans (in Zambia and Zimbabwe) and to inform policy dialogue (in Kenya and Zimbabwe) and research (in Mozambique and Zimbabwe) (EQUINET, 2012; Rasanathan, 2011). The Equity Watch process is an evidence-driven approach to strengthening rights to participation and accountability on equity in relation to SDH and universal health systems.

A range of indicators are used in such processes for relative and absolute measures of health inequities (e.g., rate ratios and rate differences across two groups) and to gain insight into the patterns of health inequities in populations, such as through the Gini coefficient or the concentration index (Kelly & Doohan, 2014). The indicators used may be related to the following issues:

- Political and legal factors, such as the presence of constitutional rights to health or, conversely, the presence of constitutional or legislative barriers to health, such as early marriage
- Economic factors, such as the gross domestic product, level of tax revenues, and household poverty or wealth, often expressed in quintiles of the population

EXHIBIT 4-14 The Equity Watch in East and Southern Africa

An Equity Watch is a means of monitoring progress on health equity by gathering, organizing, analyzing, reporting, and reviewing evidence identified from policy commitments made and from research evidence deemed relevant for improving equity in health. In addition to areas of importance for specific countries, 25 progress markers are included in all Equity Watch reports:

- Five markers of advancing equity in health
- Seven markers of access to national resources and SDH
- Eight markers of resourcing redistributive health systems
- Five markers of a more just return from a global economy

Information and country and regional equity watch reports are available at www.equinetafrica.org (EQUINET, 2012).

EXHIBIT 4-15 The EU's SOPHIE Project

The EU's "Evaluating the Impact of Structural Policies on Health Inequalities and their Social Determinants, and Fostering Change" (SOPHIE) project sought to generate new evidence on the impact of structural policies (macroeconomy, welfare state, employment relations, unemployment, built environment, and housing policies) on health inequalities, and to develop innovative methodologies for the evaluation of these policies in Europe. The Gender group of SOPHIE developed guidance for assessing how gender intersects with other social, cultural, and economic determinants, and evidence on how gender-oriented and immigration-related policies affect health inequalities. This work enables strengthened and more comprehensive evaluations of the health equity impact of policy and interventions. The main findings of the SOPHIE project include the following: (1) Evaluating structural policies is a new methodological challenge; (2) mixed methods are essential to the evaluation of structural policies; and (3) realistic approaches help to address new and vital questions, while noting the need for a clear and efficient working protocol to be developed. Further information is available on the SOPHIE project at www.sophie-project.eu/project.htm (Palència, Malmusi, & Borrell, 2014).

- Services and entitlements, such as education completion, health service coverage, or social protection
- Living standards and material conditions, such as access to clean water, air pollution, traffic density, housing, and sanitation
- Social features, such as culture, residence, class, occupation, ethnic groups, age, gender, disability, and religion (Kelly & Doohan, 2014; Solar & Irwin, 2010)

Some key factors, such as power relations, social control/autonomy, and social support are less easily measured and often less effectively included in assessments. Factors such as gender may be intertwined with power relations and other SDH in generating health inequities. As noted at the beginning of this chapter, sociopolitical processes are themselves structural determinants that can persistently impact on health outcomes, such as the processes of colonization that negatively impact indigenous peoples' health (Axelsson et al., 2016). For this reason, research and evaluation of policies affecting determinants of health equity need to take all these dimensions and their intersections into consideration, as in the European Union's SOPHIE project (described in **EXHIBIT 4-15**).

It is important that the processes for evaluation, monitoring, and analysis strengthen the social power of those affected by these SDH, and do not alienate them. In part, this means explicitly including SDH prioritized by these groups, including those determinants that are difficult to measure, and directly involving affected communities as researchers and agents of change in analysis and learning from action, such as in participatory action research and participatory evaluation processes (Loewenson et al., 2014; SHaPeS Thematic Working Group of Health Systems Global et al., 2016).

Conclusion

This chapter has explored a range of country experiences and published reviews and papers to develop a conceptual understanding of the different levels and types of SDH; their relationship to health equity, human rights, and gender equality; and their implications for health systems and for intersectoral and whole-of-government actions directed toward health and health-in-all-policies systems. While clearly a "work in progress," the growing body of evidence on all levels of SDH and their association with health outcomes is opening a deeper understanding of the "causes of the causes" in health outcomes. It is also building knowledge on how SDH at a structural level generate social stratification and social differentials in a range of intermediary SDH with an impact on health. Where these social differences in health are avoidable and remediable, they call for action on the SDH underlying them as a matter of social justice and human rights. The SDH perspective draws attention to measures to frame, understand, and integrate into interventions factors such as social exclusion, social agency, and power.

Through these approaches, those working in health-related fields can promote health and well-being and address differences in exposure and vulnerability. They can take a leadership or facilitating role in intersectoral action, and mediate or mitigate the financial, social, and physical consequences of illness. This is not just a matter for local and national levels to address. As this chapter has pointed out, globalization is influencing structural and intermediary SDH at the national and local levels, driving the need for normative and other actions on such SDH from the local level to the global level, to ensure fair benefits and prevent harms.

This chapter concluded with observations on how the positive or negative impact of these actions may be monitored and evaluated, including in terms of their distributional impacts. Evidence on and analysis of the relationship between SDH and health (equity) outcomes, while raising measurement and method challenges, is important to inform action, and to support evaluation of interventions and processes that seek to "close the gap."

At the same time, these policies and actions need to be understood in terms of how they affect the social power and status of those communities affected by them, and how they impact procedural justice. Participatory processes, wherein communities document and express their collective experience, voice, and agency, can build forms of shared power and social support that enable and sustain action on SDH, and are a social asset and a right in processes that build knowledge and action on SDH. Addressing unfair, avoidable, or remediable differences in SDH and improving health equity outcomes among population groups is not simply a technical issue: It is an ethical issue, a matter of social justice, and an outcome of social power and action.

Acknowledgments

The authors acknowledge complementary work by Kelly and Doohan (2014) in a chapter on SDH in an earlier edition of this text, together with the contributions to this field of the work made by the Commission on the Social Determinants of Health, by WHO, by regional organizations such as EQUINET, and by the many country actors cited in the chapter. We are grateful to Victoria Saint for her peer review of an early version of the chapter.

Discussion Questions

1. What is the difference between health inequality and health inequity?
2. What are the key features and dimensions of conceptual frameworks that explore the relationship between social determinants of health and health equity?
3. What are the different ways that action on SDH can reduce health inequities? Provide an example of each.
4. What role does gender play in health equity? How does a human rights approach affect the way this issue is addressed? Provide examples to demonstrate your response to both questions.
5. What can health systems do to support the measures for successful implementation of intersectoral action for health? Which features of health systems have been found to facilitate these roles? Provide examples to demonstrate your response to both questions.
6. "Social power affects the ability people have to influence and make choices over health inputs and to use these inputs to improve their well-being." Explain, with concrete examples, the pathways through which social power may positively or negatively affect health equity.
7. What are the key features of approaches used to monitor and evaluate interventions on the social determinants of health equity?

References

Aaboe, E., & Kring, T. (2013). *Natural resource management and extractive industries in Mozambique: A UN Mozambique Study.* Maputo, Mozambique: UN Country Team.

Amar Amar, J. J., & Echeverria Molina, J. (2008). Participación comunitaria para el control ciudadano de los servicios públicos domiciliarios, *Revista de Derecho, 29,* 129–152. Retrieved from http://www.scielo.org.co/pdf/dere/n29/n29a06.pdf

Amerson, R. (2013). Contributing to family health using a *promotora* programme in Guatemala. *Journal of Community Engagement and Scholarship, 6*(1), 3.

Arenas-Monreal, L., Piña-Pozas, M., & Gómez-Dantés, H. (2015). Aportes y desafíos del enfoque de género en el estudio de las enfermedades transmitidas por vector. *Salud Pública de México, 57*(1). 66–75. Retrieved from http://www.scielo.org.mx/scielo.php?pid=S0036-36342015000100010&script=sci_arttext

Australian Institute of Health and Welfare (AIHW). (2012). *Australia's health 2012*. Australia's Health Series No. 13. Cat. No. AUS 156. Canberra, Australia: Author.

Axelsson, P., Kukutai, T., & Kippen, R. (2016). The field of Indigenous heath and the role of colonisation and history. *Journal of Population Research, 33*, 1–7.

Baum, F. (2010). Overcoming barriers to improved research on the social determinants of health. *MEDICC Review, 12*(3), 36–38.

Baum, F., Lawless, A., Delany, T., Macdougall, C., Williams, C., Broderick, D., . . . Marmot, M. (2014). Evaluation of Health in All Policies: Concept, theory and application. *Health Promotion International, 29*(suppl 1), i130–i142.

Benequista, N., Gaventa, J., & Barrett, G. (2010). *A policy-maker's guide to the research of the Development Research Centre on Citizenship, Participation and Accountability.* London, UK: Department for International Development.

Besada, H., & Martin, P. (2013). *Mining codes in Africa: Emergence of a "fourth" generation?* Ottawa, Canada: North-South Institute.

Birdsall, N. (2005). *The world is not flat: Inequality and injustice in our global economy.* WIDER Annual Lecture 9. Helsinki, Finland: UNU World Institute for Development Economics Research.

Bonnefoy, J., Morgan, A., Doohan, E., Popay, J., Mackenbach, J., & Kelly, M. P. (2011). Measuring the social determinants of health: Theoretical and empirical challenges. In J. H. Lee & R. Sadana (Eds.), *Improving equity in health by addressing social determinants* (pp. 263–293). Geneva, Switzerland: World Health Organization.

Braveman, P., & Gruskin, S. (2003). Poverty, equity, human rights and health. *Bulletin of the World Health Organization, 81*(7), 539–545.

Bucket, K., Williams, C., & Wildgoose, D. (2011). *Health in All Policies: South Australia's country case study on action on the social determinants of health.* Draft Background Paper No. WCSDH/BCKGRT/5/2011. Retrieved from http://www.who.int/sdhconference/resources/draft_background_paper5_australia.pdf?ua=1

Calma, T. (2007). *Social determinants and the health of Indigenous peoples in Australia: A human rights based approach.* Workshop paper presented by Mr Darren Dick on behalf of Mr Tom Calma, Aboriginal and Torres Strait Islander Social Justice Commissioner, International Symposium on the Social Determinants of Indigenous Health, Adelaide, 29–30 April 2007. Retrieved from https://www.humanrights.gov.au/news/speeches/social-determinants-and-health-indigenous-peoples-australia-human-rights-based

Campbell and Cochrane Equity Methods Group. (2017). Cochrane Methods: Equity website. Retrieved from http://methods.cochrane.org/equity/

Catholic Relief Services (CRS). (2011). *Extractives and equity: An introductory overview and case studies from Peru, Angola and Nigeria.* Baltimore, MD: Author.

Commission on the Social Determinants of Health (CSDH). (2008). *Closing the gap in generation: Final report of the Commission on Social Determinants of Health.* Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/social_determinants/en/

Cornwall, A. (2008). Unpacking participation: Models, meanings and practices. *Community Development Journal, 43*(3), 269–283.

Cornwall, A., & Leach, M. (2010). *Putting the politics back into public engagement: Participation, mobilisation and citizenship in shaping of health services*. Mimeo, UK: Institute of Development Studies.

Cornwall, A., & Shankland, A. (2008). Engaging citizens: Lessons from building Brazil's national health system. *Social Science and Medicine, 66*(10), 2173–2184. doi:10.1016/j.socscimed.2008.01.038

Dahlgren, G., & Whitehead, M. (2007). *Levelling up (part 2): A discussion paper on European strategies for tackling social inequities in health,* Copenhagen, Denmark: WHO Regional Office for Europe.

De Backer, S. (2012). *Mining investment and financing in Africa: Recent trends and key challenges.* Johannesburg, South Africa: Webber Wentzel.

Delany, T., Lawless, A., Baum, F., Popay, J., Jones, L., McDermott, D., . . . Marmot, M. (2016). Health in All Policies in South Australia: What has supported early implementation? *Health Promotion International, 31*(4), 888–898. doi:10.1093/heapro/dav084

European Commission (EC). (2016), Employment, social affairs and inclusion: Social protection & social inclusion. Retrieved from http://ec.europa.eu/social/main.jsp?langId=en&catId=750

Friel, S., Harris, P., Simpson, S., Bhushan, A., & Baer, B. (2015). Health in All Policies approaches: Pearls from the Western Pacific region. *Asia and the Pacific Policy Studies*, *2*(2), 324–337.

Gesellschaft für Internationale Zusammenarbeit (GIZ) & World Health Organization (WHO). (2011a). *Health and human rights e-learning course. Module 1: Health and human rights norms and principles*.

Gesellschaft für Internationale Zusammenarbeit (GIZ) & World Health Organization (WHO). (2011b). *Health and human rights e-learning course. Module 2: The legal framework*.

Gilson, L. (2007). Acceptability, trust and equity. In G. Mooney & D. McIntyre D (Eds.), *The economics of health equity* (pp. 124–148). Cambridge, UK: Cambridge University Press.

Gilson, L., Doherty, J., & Loewenson, R. (2011). Challenging inequity through health systems. In J. H. Lee & R. Sadana (Eds.), *Improving equity in health by addressing social determinants* (pp. 196–230). Geneva, Switzerland: World Health Organization.

Gilson, L., Doherty, J., Loewenson, R., & Francis, V. (2008). *Challenging inequity through health systems: Final report of the Knowledge Network on Health Systems.* Geneva, Switzerland: WHO Commission on Social Determinants of Health. Retrieved from http://www.who.int/social_determinants/resources/csdh_media/hskn_final_2007_en.pdf

Global Environment Facility (GEF), Open Society Initiative for Southern Africa (OSISA), & United Nations Development Programme (UNDP). (2013). *Land, biodiversity and extractive industries in southern Africa: How effective are legal and institutional frameworks in protecting people and the environment?* Washington, DC: GEF; Johannesburg, South Africa: OSISA; & New York, NY: UNDP.

Global Witness. (2012). Extractive sector transparency: Why the EU needs a strong set of rules. Retrieved from https://www.globalwitness.org/sites/default/files/Extractive%20industry%20transparency%20briefing.pdf

Human Rights Watch (HRW). (2013). *What is a house without food? Mozambique's coal mining boom and resettlements*. New York, NY: Author.

International Institute for Environment and Development. (2002). *Breaking new ground: Mining minerals, and sustainable development.* London, UK/Sterling, VA: Earthscan Publications.

International Labour Organisation. (2014). *Tripartite declaration of principles concerning multinational enterprises and social policy (MNE Declaration)* (4th ed.). Geneva, Switzerland: Author.

Kabemba, C. (2014). *Myths and mining: The reality of resource governance in Africa.* Johannesburg, South Africa: Southern Africa Resource Watch & Open Society Initiative for Southern Africa.

Kelly, M., & Doohan, E. (2014). The social determinants of health. In M. H. Merson, R. E. Black, & A. J. Mills (Eds.), *Global health: Diseases, programs, systems and policies* (3rd ed., pp. 75–114). Burlington, MA: Jones & Bartlett Learning.

Kelly, M. P., Morgan, A., Bonnefoy, J., Butt, J., & Bergman, V. (2007). *The social determinants of health: Developing an evidence base for political action.* Final Report to World Health Organization Commission on the Social Determinants of Health from Measurement and Evidence Knowledge Network. The Measurement and Evidence Knowledge Network. Chile: Universidad del Desarrollo; United Kingdom: National Institute for Health and Clinical Excellence.

Kringos, D. S., Boerma, W. G. W., Van der Zee, J., & Groenewegen, P. (2013). Political, cultural and economic foundations of primary care in Europe. *Social Science and Medicine, 99*(0), 9–17.

Labonte, R., & Schrecker, T. (2008). *Towards health-equitable globalisation: Rights, regulation and redistribution.* Final Report to the Commission on Social Determinants of Health of the Globalisation Knowledge Network. Ottawa, Canada: University of Ottawa Institute of Population Health.

Lauthier, B. (2005). Una protección social mutualista y universal: Condición para la eficacia de la lucha contra la pobreza. In P. Salama, J. Marques-Pereira, B. Lauthier, Y. Le Bonniec, O. Rodríguez, & C. Giraldo (Eds.), *Sistemas de protección social: Entre la volatilidad económica y la vulnerabilidad social* (pp. 95–146). Bogotá, Colombia: Centro de Investigaciones para el Desarrollo (CID), Facultad de Ciencias Económicas, Universidad Nacional de Colombia (Colección estudios sobre protección social, tomo 1).

Loewenson, R. (2013a). *Evaluating intersectoral processes for action on the social determinants of health: Learning from key informants*. Social Determinants of Health Discussion Paper 5: Policy & Practice. Geneva, Switzerland: World Health Organization.

Loewenson, R. (2013b). *The Healthy Schools programme in South Africa*. Brazzaville, Republic of Congo: WHO Regional Office for Africa.

Loewenson, R. (2016, April). *Understanding and organising evidence on social power and participation in health systems*. Oxford, UK: Training and Research Support Centre.

Loewenson, R., Bambas Nolen, L., & Wamala, S. (2010). Globalisation and women's health in sub-Saharan Africa: Would paying attention to women's occupational roles improve nutritional outcomes? *Scandinavian Journal of Public Health, 38*(suppl 4), 6–17.

Loewenson, R., & Gilson, L. (2012). The health system and wider social determinants of health. In R. Smith & K. Hanson (Eds.), *Health systems in low- and middle-income countries: An economic and policy perspective* (pp. 219–242). Oxford, UK: Oxford University Press.

Loewenson, R., Hinricher, J., & Papamichail, A. (2016). *Corporate responsibility for health in the extractive sector in east and southern Africa.* EQUINET Discussion paper 108. Harare, Zimbabwe: Training and Research Support Centre & Regional Network on Equity in Health in East and Southern Africa.

Loewenson, R., Laurell, A. C., Hogstedt, C., D'Ambruoso, L., & Shroff, Z. (2014). *Participatory action research in health systems: A methods reader.* Ottawa, Canada/Geneva, Switzerland/ Harare, Zimbabwe: Training and Research Support Centre, Alliance for Health Policy and Systems Research, World Health Organization, & International Development Research Centre.

Loewenson, R., & Simpson, S. (2014). *Strengthening primary care to improve health: Learning for the USA from high- and middle-income countries.* Oxford, UK: Training and Research Support Centre.

Loewenson, R., & Simpson, S. (2015). Situational analysis on health equity and social determinants of health, Tete Province, Mozambique. Produced by TARSC in cooperation with DPS Tete and Embassy of Denmark, Mozambique. Retrieved from http://www.tarsc.org/publications/documents/Tete%20SitAn%20on%20SDHE%20Eng2015.pdf

Loewenson, R., & Whitehead, M. (2012). Robust and equity oriented health systems. Venice, Italy: WHO European Office for Investment for Health and Development.

Lusaka District Health Management Team (LDHMT) & Training and Research Support Centre (TARSC). (2015). *Health centre committee involvement in local government planning for health in Zambia*. EQUINET Case Study Brief. Harare, Zimbabwe: Regional Network on Equity in Health in East and Southern Africa.

Lusaka District Health Office (LDHO) & Training and Research Support Centre (TARSC). (2016). *Using photovoice to strengthen health centre committees as a vehicle for social participation in health system in east and southern Africa: Zambia report.* Harare, Zimbabwe: Regional Network on Equity in Health in East and Southern Africa.

Mackintosh, M. (2007, March; revised April 2008). *Planning and market regulation: Strengths, weaknesses and interactions in the provision of less inequitable and better quality health care.* Paper prepared for the Health Systems Knowledge Network of the World Health Organization's Commission on Social Determinants of Health. Retrieved from http://www.who.int/social_determinants/resources/csdh_media/planning_market_regulation_2007_en.pdf

Mackintosh, M., & Koivusalo, M. (2005). *The commercialisation of health care: Global and local dynamics and policy responses.* Basingstoke, UK: Palgrave Macmillan.

Mbwili-Muleya, C., Lungu, M., Kabuba, I., Zulu Lishandu, I., & Loewenson, R. (2008). *Consolidating processes for community–health centre partnership and accountability in Zambia: An EQUINET PRA project report.* Lusaka District Health Team and Equity Gauge Zambia, EQUINET Participatory Research Report. Harare, Zimbabwe: Regional Network on Equity in Health in East and Southern Africa.

McNab, A. (2013). The Stellenbosch consensus statement on health promoting schools. *Global Health Promotion, 20*(1), 78–81.

Micikas, M., Foster, J., Weis, A., Lopez-Salm, A., Lungelow, D., Mendez, P., & Micikas, A. (2015). A community health worker intervention for diabetes self-management among the Tz'utujil Maya of Guatemala. *Health Promotion Practice, 16*(4), 601–608.

Ministério da Planificação e Desenvolvimento (MPD). (2013). *Mozambique Integrated Growth Poles Project (P127303). Environmental and Social Management Framework (ESMF). Draft final.* Maputo, Mozambique: National Directorate of Services Planning, Ministry of Planning and Development.

Ministério de Saúde, Mozambique (MISAU). (2014). *Plano Estratégico do Sector da Saúde PESS 2014–2019.* Maputo, Mozambique: Author.

Ministry of Health Zambia, Lusaka District Health Management Team (LDHMT), & Training and Research Support Centre (TARSC). (2012). *National launch meeting of the Zambia Health Literacy Programme. Workshop report, Lusaka, 26 July 2012.* Lusaka, Zambia: LDHMT, TARSC, & Regional Network on Equity in Health in East and Southern Africa.

Mittler, J. N., Martsolf, G. R., Telenko, S. J., & Scanlon, D. P. (2013). Making sense of "consumer engagement" initiatives to improve health and health care: A conceptual framework to guide policy and practice. *Milbank Quarterly, 91*(1), 37–77.

Murombo, T. (2013). Regulating mining in South Africa and Zimbabwe: Communities, the environment and perpetual exploitation. *Law, Environment and Development Journal, 9*(1), 31–49.

Nguyen, H. T., Passmore, J., Cuong, P. V., & Nguyen, N. P. (2012). Measuring compliance with Viet Nam's mandatory motorcycle helmet legislation. *International Journal of Injury Control and Safety, 20*(2), 192–196.

Norwegian Ministry of Health and Care Services. (2007). *Report No. 20 (2006–2007) to the Storting. National strategy to reduce social inequalities in health.* Oslo, Norway: Norwegian Ministry of Health and Care Services. Retrieved from https://

www.regjeringen.no/contentassets/bc70b9942ea241cd90029989bff72d3c/en-gb/pdfs/stm200620070020000en_pdfs.pdf

Ojeda, G., & Murad, R. (2005). Salud sexual y reproductiva en zonas marginals: Situación de las mujeres desplazadas 2005. Bogotá, Colombia: Profamilia, USAID.

Ollila, E., Baum, F., & Peña, S. (2013). Introduction to Health in All Policies and the analytical framework of the book. In K. Leppo, E. Ollila, S. Peña, M. Wismar, & S. Cook (Eds.), *Health in All Policies: Seizing opportunities, implementing policies* (pp. 3–24). Helsinki, Finland: Ministry of Social Affairs and Health.

O'Neill, J., Tabish, H., Welch, V., Petticrew, M., Pottie, K., Clarke, M., . . . Tugwell, P. (2014). Applying an equity lens to interventions: Using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *Journal of Clinical Epidemiology, 67*(1), 56–64.

Oxman, A., Lavis, J., Lewin, S., & Fretham, A. (2009). Support tools for evidence-informed health. Policymaking (STP) 10: Taking equity into consideration when assessing the findings of a systematic review. *Health Policy Research and Systems, 7* (Suppl 1), S10.

Palència, L., Malmusi, D., & Borrell, C. (2014). *Incorporating intersectionality in evaluation of policy impacts on health equity: A quick guide.* Barcelona, Spain: Agéncia de Salut Pública de Barcelona, CIBERESP, SOPHIE.

Palmary, I. (2015). Reflections on social cohesion in contemporary South Africa. *Psychology in Society, 49*, 62–69. Retrieved from http://www.scielo.org.za/pdf/pins/n49/05.pdf

Passmore, J., Tu, N. T. D., Luong, M. A., Chinh, N. D., & Nam, N.P. (2010). Impact of mandatory motorcycle helmet wearing legislation on head injuries in Viet Nam: Results of preliminary analysis. *Traffic Injury Prevention, 11*(2), 202–206.

Passmore, J., Nguyen, L. H., Nguyen, P. N., & Olivé, J.-M. (2010). The formulation and implementation of a national helmet law: A case study from Viet Nam, *Bulletin of the World Health Organization, 88*, 783–787.

Pennington, A., & Whitehead, M. (2014). *Learning from promising primary care practice models for the USA: Country case study UK: With a focus on England.* Liverpool, UK: Training and Research Support Centre & University of Liverpool.

Perez Jarauta, M., Goya Arive, M. A., & Merino Merino, B. (2010). Spain: Health promotion among Navarre ethnic minorities programmed. In T. Koller (Ed.), *Poverty and social exclusion in the WHO European Region: Health systems respond*. Copenhagen, Denmark: WHO Regional Office for Europe. Retrieved from http://www.euro.who.int/__data/assets/pdf_file/0006/115485/E94018.pdf

Popay, J., Escorel, S., Hernández, M., Johnston, H. B., Mathieson, J., & Rispel, L. (2011). Social exclusion and health inequalities: Definitions, policies and actions. In J. H. Lee & R. Sadana (Eds.), *Improving equity in health by addressing social determinants.* pp. 88-115. Geneva, Switzerland: World Health Organization.

Pronyk, P., Hargreaves, J., Kim, J., Morison, L., Phetla, G., Watts, C., . . . Porter, J. D. (2006). Effect of structural intervention for the prevention of intimate partner violence and HIV in rural South Africa: a cluster randomised trial. *Lancet, 368*(9551), 1973–1983.

Public Health Agency of Canada (PHAC). (2007). *Crossing sectors: Experiences in intersectoral action, public policy and health.* Paper prepared in collaboration with the Health Systems Knowledge Network of the WHO Commission on Social Determinants of Health and EQUINET. Ottawa, Canada: Public Health Agency of Canada.

Rasanathan, K. (2011). *Closing the gap: Policy into practice on social determinants of health: Discussion paper for the World Conference on Social Determinants of Health.* Geneva, Switzerland: World Health Organization.

Rede Brasil de Direitos Humanos. (n.d.). Curso sistematico de derechos humanos [Systematic course of human rights]. Online course by Instituto de Estudios Politicos para America Latina y Africa. Retrieved from http://www.dhnet.org.br/dados/cursos/edh/interdisciplinario/ddhh_c.htm

Regional Network on Equity in Health in East and Southern Africa (EQUINET). (2012). *Regional Equity Watch 2012: Assessing progress towards equity in health in East and Southern Africa.* Harare, Zimbabwe: Author.

Regional Network on Equity in Health in East and Southern Africa (EQUINET) Steering Committee (SC). (2007). *Reclaiming the resources for health: A regional analysis of equity in health in east and southern Africa*. Harare, Zimbabwe/Kampala, Uganda/Johannesburg, South Africa: EQUINET, Weaver Press, Fountain Publishers, & Jacana Media.

Republic of Mozambique, Provincial Government of Tete, Provincial Health Directorate, with Embassy of Denmark and Training and Research Support Centre (TARSC). (2016). *Tackling the social determinants of health equity in Tete Province, Mozambique.* Mozambique: DPS Tete.

Rispel, L., Molomo, B., & Dumela, S. (2008). *South African case study on social exclusion: Social Exclusion Knowledge Network Background Paper No. 3.* Johannesburg, South Africa.

Sadana, R., Simpson, S., Popay, J., Albrecht, D., Hosseinpoor, A. R., & Kjellstrom, T. (2011). Strengthening efforts to improve health equity. In J. H. Lee & R. Sadana (Eds.), *Improving equity in health by addressing social determinants* (pp. 7–22). Geneva, Switzerland: World Health Organization.

Salinas, G., & Haacker, M. (2006). HIV/AIDS: The impact on poverty and inequality. IMF Working paper No. 06/126. Retrieved from http://www.imf.org/en/publications/wp/issues/2016/12/31/hiv-aids-the-impact-on-poverty-and-inequality-19175

Salvucci, V. (2014). *District poverty maps for Mozambique: 1997 and 2007.* Maputo, Mozambique: Mimeo.

Sen, G., & Ostlin, P. (2011). Gender inequity in health. In J. H. Lee & R. Sadana (Eds.), *Improving equity in health by addressing social determinants* (pp. 59–87). Geneva, Switzerland: World Health Organization.

Shankardass, A., Solar, O., Murphy, K., Greaves, L., & O'Campo, P. (2012). A scoping review of intersectoral action for health equity involving governments. *International Journal of Public Health, 57*(1), 25–33.

Shankardass, A., Solar, O., & O'Campo, P. (2012, November). *WHO Database Intersectoral Action for Health Equity Case Studies. Draft WHO Database Update Protocol, Version 1.7* (rev2).

SHaPeS Thematic Working Group of Health Systems Global, Regional Network on Equity in Health in East and Southern Africa (EQUINET), & Emerging Voices for Global Health; Daniels, K., Loewenson, R., & 68 other signatories. (2016). Fair publication of qualitative research in health systems: A call by health policy and systems researchers. *International Journal for Equity in Health, 15*, 98. doi:10.1186/s12939-016-0368-y.

Shelton, G., & Kabemba, C. (Eds.). (2012). *Win–win partnership? China, Southern Africa and the extractive industries.* Johannesburg, South Africa: Southern Africa Resource Watch.

Simpson, S., Kelly, M. P., & Morgan, A. (2013). Defining principles for good practice: Using case studies to inform health systems

action on health inequalities. *Evaluation and Program Planning, 36,* 191–197.

Simpson, S., Mahoney, M., Harris, E., Aldrich, R., & Williams, S. J. (2005). Equity-focused health impact assessment: A tool to assist policy makers in addressing health inequalities. *Environmental Impact Assessment Review, 25*, 772–782.

Solar, O., & Irwin, A. (2010). *A conceptual framework for action on the social determinants of health.* WHO Discussion Paper 2 (Policy and Practice), Geneva, Switzerland: World Health Organization.

Southern African Development Community (SADC). (2003) *SADC HIV/AIDS framework and strategic programme of action 2003–2007*. Gaborone, Botswana: Author.

Stronks, K., & Mackenbach, J. P. (2005). Evaluating the effects of policies and interventions to address inequalities in health: Lessons from a Dutch programme. *European Journal of Public Health, 16*(4), 346–353.

Training and Research Support Centre (TARSC). (2014). Strengthening primary care in the USA to improve health: Learning from high- and middle-income countries. Retrieved from http://www.tarsc.org/publications/documents/Learning%20on%20PC%20from%20HMICs%20TARSC%20RWJF.pdf

Turrell, G., Oldenburg, B., McGuffog, I., & Dent, R. (1999). *Socioeconomic determinants of health: Towards a national research program and a policy and intervention agenda.* Canberra, Australia: Queensland University of Technology, School of Public Health & Ausinfo.

UNICEF. (2014). *Ending child marriage: Progress and prospects*. New York, NY: Author.

UNICEF global databases. (2016). Based on Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), and other nationally representative surveys, 2008–2014. Retrieved from https://data.unicef.org/topic/child-protection/child-marriage/

United Nations. (1948). *Universal Declaration of Human Rights. Adopted and proclaimed by General Assembly resolution 217 A (III) of 10 December 1948*. New York, NY: Author.

United Nations. (1976). The International Covenant on Economic, Social and Cultural Rights. Retrieved from http://www.unhchr.ch/html/menu3/b/a_cescr.htm

United Nations. (2000). General Comment No 14: The Right to the Highest Attainable Standard of Health Committee on Economic, Social, and Cultural Rights. Twenty-second Session, 25 April–12 May 2000, Geneva, Switzerland: Author.

United Nations Human Rights Based Approach (UNHRBA). (2003). The human rights based approach to development cooperation: Towards a common understanding among UN agencies. Retrieved from http://hrbaportal.org/the-human-rights-based-approach-to-development-cooperation-towards-a-common-understanding-among-un-agencies

University of Western Cape. (2006). Creating a decade of health promoting schools, strengthening whole school development: Executive summary report on the National Conference on Health Promoting Schools held at the University of Western Cape, Bellville, Western Cape, South Africa, 14–16 September 2006.

von der Goltz, J., & Barnwal, P. (2014). *Mines: The local wealth and health effects of mining in developing countries.* Working paper. New York, NY: School of International and Public Affairs, Columbia University.

Wallerstein, N. (2006). *What is the evidence on effectiveness of empowerment to improve health?* Copenhagen, Denmark: WHO Regional Office for Europe.

Whiting, D., Unwin, N., & Roglic, G. (2010). Diabetes: Equity and social determinants. In E. Blas & A. Sivisankara Kurup (Eds.), *Equity, social determinants and public health programmes.* pp. 77-84. Geneva, Switzerland: World Health Organization.

WHO Europe. (2013). *Health 2020: A European policy framework and strategy for the 21st century.* Copenhagen, Denmark: World Health Organization.

WHO Regional Office for Europe (WHO Europe). (2011). *Developing the new European policy for health 2020.* Concept paper. First meeting of the European Health Policy Forum for High-Level Government Officials, Andorra, 9–11 March 2011. Retrieved from http://www.ndphs.org/documents/2506/Health2020_European%20Health%20Policy%20formulation_eng.pdf

World Commission on the Social Dimensions of Globalisation. (2004). *A fair globalisation: Creating opportunities for all.* Geneva, Switzerland: International Labour Organisation

World Health Organization (WHO). (2005). Sustainable health financing, universal coverage and social health insurance. World Health Assembly Resolution 58.33, Geneva, Fifty-eighth World Health Assembly, 16–25 May 2005. Retrieved from http://apps.who.int/medicinedocs/documents/s21475en/s21475en.pdf

World Health Organization (WHO). (2006). *Constitution of the World Health Organisation.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2008). *Primary health care: Now more than ever.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2011). *Gender mainstreaming for health managers: A practical approach. Facilitators' guide.* Paris, France: Author.

World Health Organization (WHO). (2013a). *Demonstrating a health in all policies analytic framework for learning from experiences: Based on literature reviews from Africa, South-East Asia and the Western Pacific*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2013b). *Handbook on health inequality monitoring with a special focus on low- and middle-income countries*. Luxembourg: Author.

World Health Organization (WHO). (2014, September). Adolescent pregnancy. Fact Sheet No. 364. Retrieved from http://www.who.int/mediacentre/factsheets/fs364/en/

World Health Organization (WHO). (2015). What is health in all policies? [Infographic on air pollution]. Retrieved from http://who.int/social_determinants/publications/health-policies-manual/HiAP_Infographic.pdf?ua=1

World Health Organization (WHO). (2016). *Innov 8: The Innov8 approach for reviewing national health programmes to leave no one behind.* Geneva, Switzerland: Author

World Health Organization (WHO). (2017a). Social Determinants of Health Sectoral Briefing Series. Retrieved from http://www.who.int/social_determinants/corner/en/

World Health Organization Regional Office for Europe. (2016). *Women's health and well-being in Europe: Beyond the mortality advantage.* Copenhagen, Denmark: Author.

CHAPTER 5

Reproductive Health

M. Omar Rahman, Jane Menken, and Dipak Kumar Mitra

Reproductive health in low- and middle-income countries (LMICs) has long been addressed primarily through family planning and maternal and child health programs, and through programs to prevent and treat sexually transmitted infections (STIs), particularly human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS).

Reproductive health is tied to policy concerns about population growth as well as health. In 1994, the United Nations (UN) sponsored the third decennial International Conference on Population and Development (ICPD) in Cairo. The previous two conferences had emphasized family planning and economic development, respectively, as the major focus of population policy—policy that was intended to reduce fertility and, thereby, population growth. The rationale for support of family planning programs included both the right of individuals to control their own fertility and the belief that reduced fertility would lead to reduced population growth, which would have benefits for individuals, nations, and the world. In 1994, for the first time, women's health advocates, including many from nongovernmental organizations (NGOs), played a key role in the ICPD and brought to the fore issues of reproductive health that went beyond family planning. They called for a fundamental redefinition of population policy that focused on the status of women and gave "prominence to reproductive health and the empowerment of women while downplaying the demographic rationale for population policy" (McIntosh & Finkle, 1995).

The 1994 ICPD adopted a Programme of Action that included the following definition of reproductive health:

> Reproductive health is a state of complete physical, mental and social well-being, and not merely the absence of disease or infirmity, in all matters relating to the reproductive system and its processes. Reproductive health therefore implies that people are able to have a satisfying and safe sex life and that they have the capability to reproduce and the freedom to decide if, when, and how often to do so. Implicit in this last condition are the right of men and women to be informed and to have access to safe, effective, affordable and acceptable methods of family planning of their choice, as well as other methods of their choice for the regulation of fertility which are not against the law, and the right of access to appropriate health-care services that enable women to go safely through pregnancy and childbirth and provide couples with the best chance of having a healthy infant. . . . It also includes sexual health, the purpose of which is the enhancement of life and personal relations, and not merely counseling and care related to reproduction and sexually transmitted diseases (United Nations, 1994).

This vision of reproductive health has—perhaps not unexpectedly—proved controversial and has not been achieved to the extent hoped for by its

proponents (see, for example, "Conference Adopts Plan on Limiting Population," 1999; "Population Control Measures to Aid Women Are Stumbling," 1999). Nor did donors uphold the pledges made at the Cairo conference. In 2003, the UN estimated that the international donor community had contributed far less than the pledges for reproductive health made at Cairo. Of the $6.1 billion promised by 2005, slightly more than $3 billion was provided through 2003 (Population Reference Bureau, 2004; United Nations Commission on Population and Development, 2004). However, since the initiation of Millennium Development Goals (MDGs) and other allied efforts, such as the "Global Strategy for Women's and Children's Health" and "Every Woman, Every Child" campaign, the situation has significantly improved. Commitments worth of approximately $45 billion were made to achieve MDG targets, of which $31 billion had been disbursed as of May 2015. Annual disbursement for women's and children's health increased from 2010 to 2013 by 25%, particularly in previously underfunded areas including family planning, where the increase was 50.5% from 2010 to 2013 (World Health Organization [WHO], 2015a).

A less controversial and more limited version of the vision for reproductive health guided the 1997 U.S. National Academy of Sciences report on reproductive health (Tsui, Wasserheit, & Haaga, 1997):

1. Every sex act should be free of coercion and infection.
2. Every pregnancy should be intended.
3. Every birth should be healthy for both mother and child.

Even so, no country, according to Tsui, Wasserheit, and Haaga (1997), met these more limited goals by 1997, and the problems were greatest in LMICs.

In 2000, the United Nations Millennium Declaration was adopted as a commitment to "making the right to development a reality for everyone and to freeing the entire human race from want" (United Nations General Assembly, 2000). Eight MDGs were established as part of the "Roadmap Toward the Implementation of the United Nations Millennium Declaration" (United Nations General Assembly, 2001). Goal 5 was to improve maternal health, with a specific target of reducing "by three quarters, between 1990 and 2015, the maternal mortality ratio." Despite significant improvements in maternal health during the MDG era (the world maternal mortality ratio [MMR] declined from 385/100,000 live births in 1990 to 216/100,000 live births in 2015), final estimates show that vast majority of LMICs failed to achieve the three-fourths reduction in maternal mortality required by MDG 5 (UNICEF & WHO, 2015). A wide variation exists in the reductions in MMR achieved in specific world regions. South Asian countries experienced the highest reduction, 67% (from 558/100,000 live births in 1990 to 182/100,000 live births in 2015), and sub-Saharan countries had the lowest reduction over the same period, 37%. The Sustainable Development Goals (SDGs) established as follow-ons to the MDGs have already been set, with the SDG being to bring down global MMR to 70/100,000 live births by 2030 (United Nations, 2016).

This chapter emphasizes both the older and the newer views of family planning and reproductive health. In the next section, which examines demographic trends, the focus is on population growth and change and the transitions under way around the world from situations of high fertility and high mortality to those of low fertility and low mortality. The ways in which people control their fertility and indices of the effects of various fertility determinants on overall fertility in a range of countries are then considered. The third section examines family planning programs and their role in the reduction of fertility and unintended pregnancy. The next two sections consider the role of fertility patterns in the health of children and women, respectively. The final section presents brief recommendations for future research and programs. Because the *Infectious Diseases* chapter focuses on sexually transmitted infections (STIs), including HIV/AIDS, these crucial aspects of reproductive health are mentioned only briefly here.

Demographic Trends and Fertility Determinants

History of Population Growth

To understand the context of the concerns about population and reproductive health in the world today, it is instructive to review the history of population growth and its associated impacts. **FIGURE 5-1** shows the growth of the world's population and, in particular, the extraordinary changes in growth rates over the past 200 years. The world population reached 1 billion just after 1800. By the turn of the twentieth century, it had reached 1.6 billion, and before 1930 it had surpassed 2 billion. Thus, it took less than 125 years to add the second billion people—an astounding feat given the long sweep of history needed to reach the first billion. World population passed the 3 billion mark in 1960. Each additional billion people has taken less time to add, so that by the year 2000, population had reached 6.1 billion. Thus, population doubled between 1960

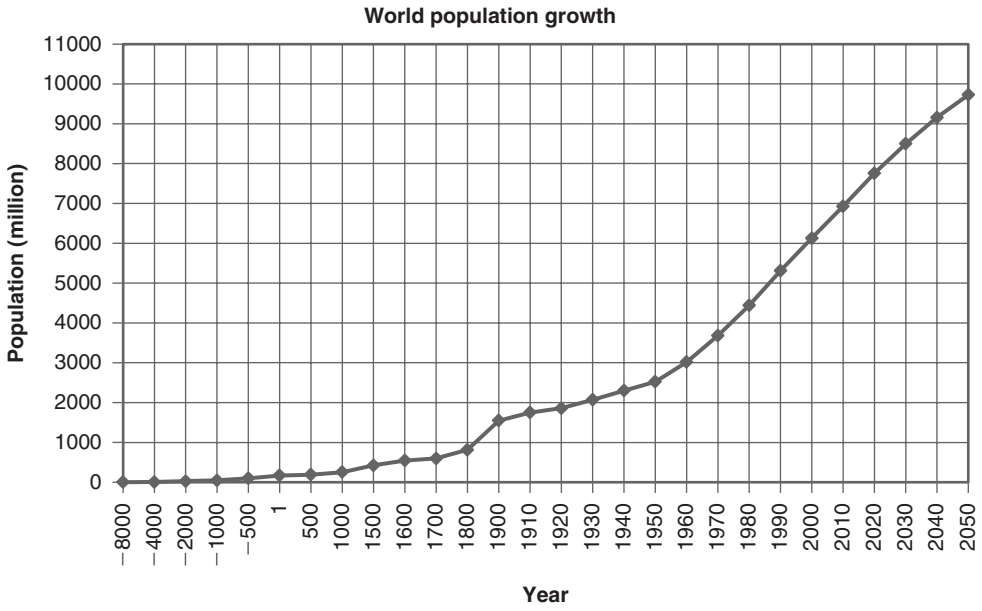


FIGURE 5-1 The growth of population and, in particular, the extraordinary change of the past 200 years.

Data from U.S. Census Bureau, International Programs, International Data Base, Revised: September 27, 2016 for the years 8000 BC to 1940 AD and from United Nations, Department of Economics and Social Affairs, Population Division (2015). *World population prospects: The 2015 revision, volume I: Comprehensive tables (ST/ESA/SER.A/379)* for the years 1950–2050 AD.

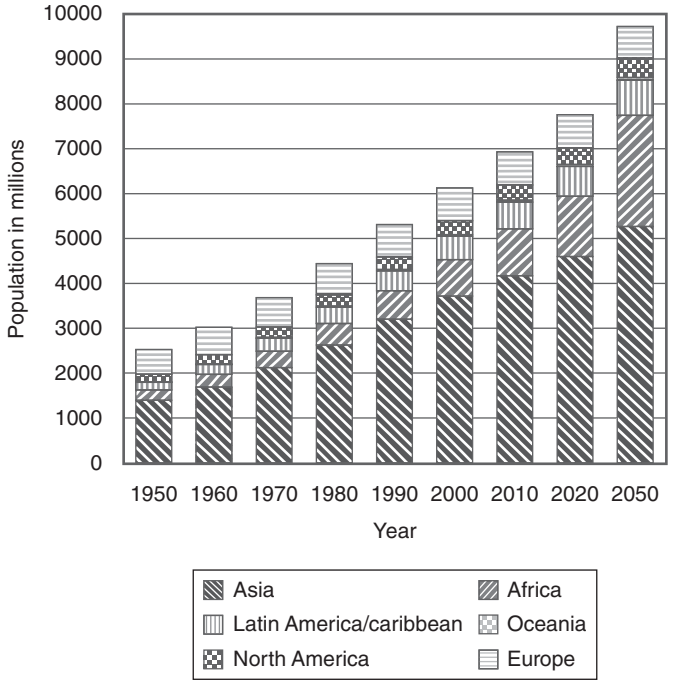


FIGURE 5-2 Population size by continent, 1950–2050.

Data from United Nations, Department of Economics and Social Affairs, Population Division (2015). *World population prospects: The 2015 revision, volume I: Comprehensive tables (ST/ESA/SER.A/379)*. Accessed on November 10, 2016 from https://esa.un.org/unpd/wpp/publications/Files/WPP2015_Volume-I_Comprehensive-Tables.pdf

and 2000, adding 3 billion people in only 40 years. By 2010, almost 1 billion people were added in just 10 years' time, so that world population in 2010 was 6.9 billion (United Nations Population Division, 2009); Figure 5-1). The global population projection until 2050 shows that every 10 years 1 billion people will be added in this planet (United Nations Population Division, 2015b).

The majority of this population expansion has taken place in the low- and middle-income regions of Asia, Africa, Latin America, and Oceania (**FIGURE 5-2**), which accounted for 84% of the world's population in 2010, compared with 71% in 1950. The share of world population in these regions is expected to continue increasing, reaching 88% in the middle of the twenty-first century (United Nations Population Division, 2015b).

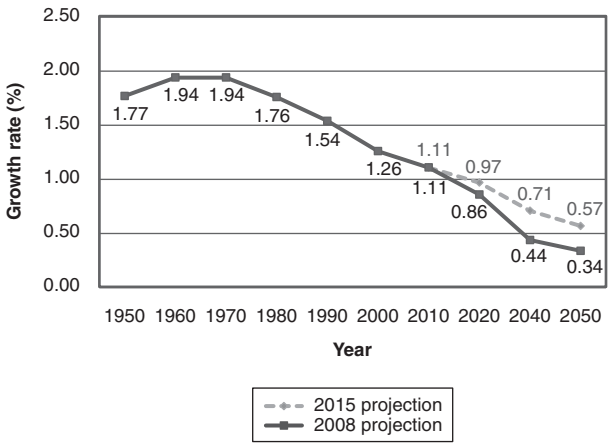


FIGURE 5-3 World population growth rate, 1950–2050, projected in 2008 (1950–2050) and 2015 (2010–2050).

Data from United Nations, Department of Economics and Social Affairs, Population Division (2009). *World population prospects: The 2008 revision, volume I: Comprehensive tables (ST/ESA/SER.A/287).* Accessed on November 10, 2016 from http://kczx.shupl.edu.cn /download/786444c9-20c1-4b5a-b0d6-d7544569a2ee.pdf and United Nations, Department of Economics and Social Affairs, Population Division (2015). *World population prospects: The 2015 revision, volume I: Comprehensive tables (ST/ESA/SER.A/379).* Accessed on November 10, 2016 from https://esa.un.org/unpd/wpp/publications/Files/WPP2015_Volume-I_Comprehensive-Tables.pdf

The encouraging news is that the rate at which the world's population is growing has declined continuously since about 1960 (**FIGURE 5-3**), although the absolute *number* of people added in each decade has continued to increase. After 2000, that number was projected to decline. Population growth rates are declining in most LMICs, albeit at an uneven pace; that is, some countries and regions are experiencing much more rapid change than others. China is a particularly prominent example: Its growth rate was 2.1% in 1960, increased to 2.6% in 1965, but then steadily dropped to less than 1% by 1995 and is expected to reach zero growth by 2030.The United Nations estimates that world population in the middle of the twenty-first century will total between 8 billion and 11 billion people. By that time, the growth rate will be significantly less than 1% but will not yet have achieved the no-growth stable situation.

It is important to note that population projections involve complex statistical procedures with variable uncertainty, dependent on the stability of the mortality and fertility trend lines. This is particularly true as the projection period becomes longer. In hindsight, actual changes can vary significantly from the projections if mortality and fertility trends do not match expectations. For example, the UN projection made in 2008 estimated that the world population growth rate would decrease faster, reaching 0.34 in 2050. However, the updated projection in 2015 shows that the decline in world population growth will be relatively slower than earlier estimates and will reach only 0.57 by the year 2050 (Figure 5-3). This outcome reflects the combined effects of lower than expected fertility reduction and greater than expected mortality reduction, particularly in Africa. Fertility reduction was slower than expected due to lack of proper implementation of family planning programs, and mortality reduction was faster than expected due to better control of the HIV/AIDS epidemic in affected countries (May et al., 2014).

Because of reduced fertility and increased life expectancy, the proportion of the population consisting of people older than 65 years will increase over time from 8% in 2010 to 16% in 2050, while the proportion attributable to children younger than age 15 will decrease from 26.7% in 2010 to 21.3% in 2050 (United Nations Population Division, 2015b).

The Demographic Transition

What explains the historic experience of low initial population growth, followed by an explosive increase and finally by a steady decline in growth? The classic theory of demographic transition proposed by Notestein (1953) and others postulated that all societies initially start off with high fertility and high mortality levels. At some point in societal development, mortality rates fall due to public health advances, while fertility rates remain high. This combination results in explosive population growth, with birth rates far exceeding death rates, until at some point birth rates also start to decline and a new equilibrium is reached at low fertility and low mortality levels.

Until fairly recently, the classic theory of demographic transition held sway, and all societies were supposed to go through it in a lock-step manner. In the early to mid-1970s, however, an international team of researchers participated in the Princeton University European Fertility Project and carefully examined historical fertility declines in Europe. They came to a somewhat surprising conclusion: The process of demographic transition is quite varied and does not always follow the path suggested by classic theory (Coale & Watkins, 1986). Under that scenario, a certain level of socioeconomic development was required for the initial mortality decline, which was followed, at some later point, by fertility decline. Researchers in the European Fertility Project, however, found that mortality decline took place in different societies at different levels of development and that there was no magic threshold of mortality above which fertility decline would not take place.

The current consensus about the demographic transition is that there is no specific sequence in which fertility and mortality decline. They can decline together, or one can decline before the other. Furthermore, no specific thresholds of development are required for either process to start. Moreover, the intervals between a high-fertility and high-mortality regime and a subsequent low-fertility and low-mortality regime

are also not fixed and can vary considerably. The experience of the LMICs has borne out this new consensus. Demographic transitions have taken place at different rates in different places and with different sequences. A common thread, however, is that the transition has often been considerably more rapid than those seen in the European or North American historical record.

Population growth rates and fertility and mortality for different parts of the world since the mid-twentieth century to the mid-twenty-first century are shown in **TABLE 5-1**, and for the specific cases of Bangladesh and Kenya, in **EXHIBIT 5-1**. Growth rates are given as the percent change in population per year for the 5 years subsequent to the specified date. When a population stops changing in size, it becomes *stationary* and its growth rate is zero. The total fertility rate (TFR) is the number of children a woman would bear, on average, if they lived to the end of the reproductive period under the childbearing pattern of a particular year—for example, if they had, at age 15, the birth rate of 15-year-olds in 1970; at age 16, the birth rate of 16-year-olds in 1970; and so on. Life expectancy at birth is the average number of years people would live if their entire life were spent under the age-specific mortality conditions of a particular year—for example, if they experienced the infant mortality of 1970, the death rate at age 1 of 1970, and so on. The numbers resulting from these examples would be the 1970 TFR and the 1970 life expectancy. A TFR of approximately 2.1 is usually referred to as *replacement-level* fertility. If, over the long run, women have that number of children on average, the population will become stationary, neither growing

TABLE 5-1 Growth Rate, Total Fertility Rate, and Life Expectancy for Regions of the World by Time Period

Region	1950	1970	1990	2000	2010	2020	2050
World							
Growth rate	1.77	1.96	1.54	1.24	1.18	0.97	0.57
Total fertility rate	4.96	4.48	3.04	2.62	2.51	2.43	2.25
Life expectancy	46.8	58.0	64.5	67.1	70.5	72.7	77.1
Low- and Middle-Income Countries							
Growth rate	2.03	2.39	1.84	1.46	1.36	1.12	0.66
Total fertility rate	6.08	5.42	3.39	2.83	2.65	2.52	2.30
Life expectancy	41.5	54.8	62.5	65.3	68.8	71.2	76.0
Higher-Income Countries							
Growth rate	1.19	0.77	0.44	0.34	0.29	0.17	–0.02
Total fertility rate	2.82	2.15	1.67	1.58	1.67	1.72	1.82
Life expectancy	64.7	71.1	74.1	75.6	78.3	79.9	83.5
Asia							
Growth rate	1.91	2.29	1.63	1.20	1.04	0.75	0.19
Total fertility rate	5.82	5.06	2.96	2.39	2.20	2.09	1.92
Life expectancy	42.1	56.4	65.1	68.5	71.6	73.9	78.3

(continues)

TABLE 5-1 Growth Rate, Total Fertility Rate, and Life Expectancy for Regions of the World by Time Period *(continued)*

Region	1950	1970	1990	2000	2010	2020	2050
Africa							
Growth rate	2.08	2.61	2.63	2.45	2.55	2.31	1.77
Total fertility rate	6.60	6.67	5.73	5.10	4.71	4.14	3.11
Life expectancy	37.3	46.4	51.7	53.3	59.5	62.9	69.9
Latin America and Caribbean							
Growth rate	2.69	2.43	1.73	1.36	1.12	0.85	0.26
Total fertility rate	5.89	5.03	3.01	2.52	2.15	1.96	1.78
Life expectancy	51.2	61.2	68.4	72.1	74.5	76.8	81.7
Oceania							
Growth rate	2.23	1.76	1.49	1.43	1.54	1.23	0.79
Total fertility rate	3.84	3.23	2.49	2.43	2.42	2.29	2.06
Life expectancy	60.4	66.4	72.5	75.1	77.5	79.2	82.1
North America							
Growth rate	1.67	0.95	1.05	0.92	0.78	0.69	0.38
Total fertility rate	3.35	2.01	2.00	1.99	1.86	1.87	1.90
Life expectancy	68.6	71.4	75.8	77.4	79.2	80.6	84.3
Europe							
Growth rate	0.98	0.60	0.19	0.07	0.08	–0.04	–0.21
Total fertility rate	2.66	2.17	1.57	1.43	1.60	1.66	1.79
Life expectancy	63.6	70.6	72.6	73.8	77.0	78.6	82.2

Note: Growth rate (% per year), total fertility rate, and life expectancy are for the subsequent 5-year period (e.g., 1950–1955). Life expectancy is for both sexes combined. All estimates are from the medium variant projections.

Data from United Nations, Department of Economics and Social Affairs, Population Division. (2015). *World population prospects: The 2015 revision, volume I: Comprehensive tables (ST/ESA/SER.A/379).* Accessed on November 10, 2016 from https://esa.un.org/unpd/wpp/publications/Files/WPP2015_Volume-I_Comprehensive-Tables.pdf

nor declining. Women will be contributing to the next generation one child for themselves and one for their partner, and a bit more for girls who were born but did not survive to reproduce.

As shown in Table 5-1, life expectancy rose continuously over the latter half of the twentieth century. The exceptions (not shown) occurred in the countries of Africa hardest hit by AIDS, where, by 2000, this disease had wiped out much, if not all, earlier gains (see Exhibit 5-1 for the specific case of Kenya). According to the United Nations, nine countries had adult HIV prevalence of 10% or more in 2009: Swaziland,

Botswana, Lesotho, Zimbabwe, South Africa, Namibia, Zambia, Malawi, and Mozambique (UNAIDS Fact Sheets, 2009). The U.S. Census Bureau (2004) estimated that, on average, life expectancy in 2010 will be significantly less in these countries than it would have been in the absence of AIDS. Much of the increase in every country in life expectancy is due to improvements in infant and child survival; by contrast, the HIV-related declines are primarily the result of increased adult mortality. However, the latest WHO (2016) estimate shows that African countries have overcome the setbacks posed by AIDS epidemic and that life expectancy dramatically improved between 2000 and 2015. With increasing access to antiretroviral treatment (ART), the number of AIDS-related deaths declined by 39% between 2005 and 2013 in sub-Saharan Africa, increasing life expectancy in that area by 10 years (Kharsany & Karim, 2016).

Total fertility rate had declined by 1970 in all parts of the world, with the exception of Africa. That decline was, in low- and middle-income regions, overwhelmed by increases in life expectancy, so that these

EXHIBIT 5-1 Demographic Change in Kenya and Bangladesh

Kenya (population = 46 million) and Bangladesh (population = 161 million) will be used as case studies to illustrate demographic change in this chapter for a number of reasons. There are both similarities and differences in their experiences.

First, similar to Bangladesh, Kenya experienced high population growth rates that were the result of continuing high fertility during a period when mortality was declining quite sharply. Thus, in the period 1950–1955, life expectancy was 40.5 years for males and 44.2 years for females. By 1980–1985, almost 16.5 years had been added to life expectancy for both males and females in Kenya, raising life expectancy there to 57 years for males and 60.7 years for females. In comparison, although life expectancy also rose sharply during the same period for Bangladesh, the increase was notably less than in Kenya (10 years for males and 14 years for females). Because Bangladesh started from a lower base in 1980–1985, life expectancy in Bangladesh was actually significantly (almost 10 years) lower than that in Kenya (male = 48.9 years; female = 50.2 years). As a result of the sharp decline in mortality coupled with continuing high fertility rates (during that period, the total fertility rate in Kenya exceeded 7 births per woman), population growth rates in Kenya increased from 2.77% per year in 1950–1955 to 3.78% per year in 1980–1985. Similarly, for Bangladesh, due to the sharp drop in mortality and the continued high fertility (TFR remained between 6 and 7 births per woman), there was also a sharp increase in population growth rates, from 2.11% in 1950–1955 to 2.61% in 1980–1985.

Second, as was the case in Bangladesh, a successful family planning program in Kenya managed to bring about a fertility decline despite relatively little improvement in socioeconomic indicators (Toroitich-Ruto, 2001). Fertility dropped sharply over the next 15 years, from 7.22 births per woman in 1980–1985 to 5.07 births per woman in 1995–2000. Fertility subsequently appears to have declined slowly; it was approximately estimated at 4.96 births per woman in 2005–2010 and 4.54 births per woman in 2010–2015. Both the initial sharp decline and the subsequent slower decline in total fertility rates are similar to the experience of Bangladesh, although TFR in Bangladesh remained at a significantly lower level—approximately 3.3 births per woman in 1995–2000—and subsequently dropped very slowly to an estimated 2.36 births per woman in 2005–2010 and 2.2 births per woman in 2010–2015.

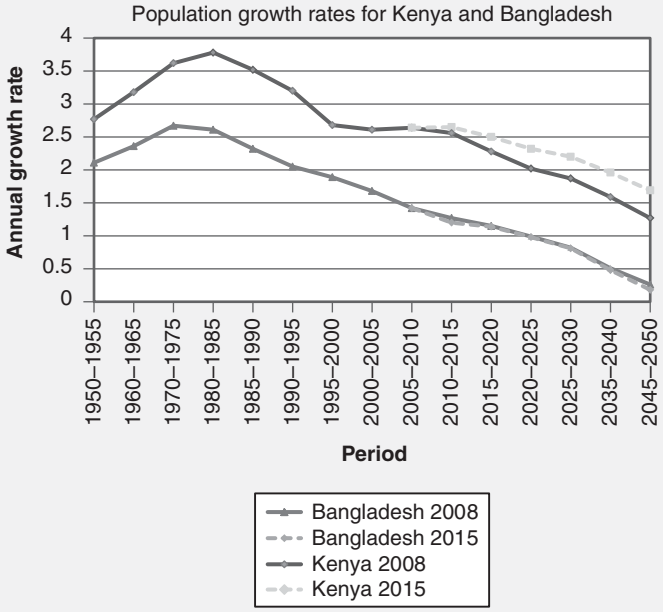
Kenya, like other countries severely affected by AIDS, saw its life expectancy fall in the latter part of the twentieth century, from 57 to 51 years for males and from 62 to 52.3 years for females in 2000. Since then, the situation in Kenya has improved due to a significantly faster than expected morality decline resulting from more effective control of HIV in the post-2008 period. Additional gains of 4 to 5 years in life expectancies have been observed in both males and females for Kenya. However, unlike mortality reductions, fertility rates in Kenya declined much less than expected as per the 2008 projections. Instead of declining from 4.9 births per woman in 2010 to 2.39 births per woman in 2050, TFR will decline to only a projected 2.85 births per woman in 2050 as per 2015 projections. Thus, the annual population growth rate in Kenya was predicted to be 1.27% in 2050 by the 2008 projections and is now 1.69% according to the 2015 projections.

In the case of Bangladesh, since 2008, mortality reductions have been slightly greater than expected (an additional 2 years of life expectancy is projected through 2050) and fertility reductions have also been greater than expected, thus overriding the mortality decline. TFR, instead of declining from 2.36 births per woman in 2010 to 1.85 births per woman in 2050, is now expected to decline to a projected 1.67 births per woman in 2050. Thus, the predicted growth rate in 2050 for Bangladesh was 0.26% as per the 2008 projections and is now 0.18% according to the 2015 projections.

In conclusion, Kenya experienced significantly lower than expected declines in population growth rates in the post-2008 period due to a combination of higher than expected mortality decline along with significantly lower than expected fertility decline. Bangladesh, by contrast, experienced basically a flat picture, with a slightly higher than expected fertility decline matching the slightly higher than expected mortality decline.

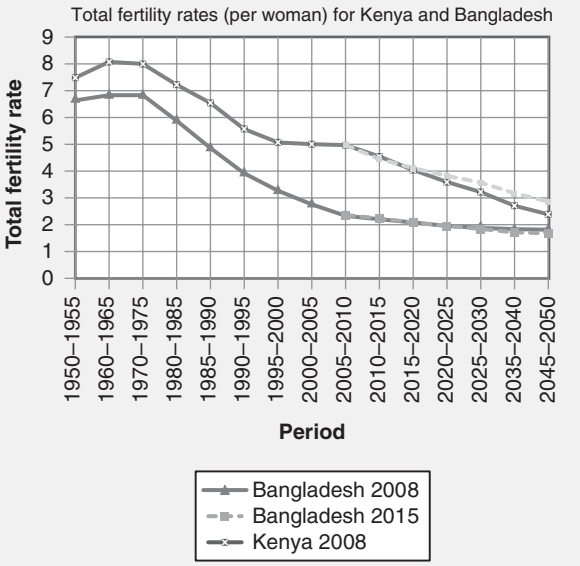
(continues)

EXHIBIT 5-1 Demographic Change in Kenya and Bangladesh *(continued)*



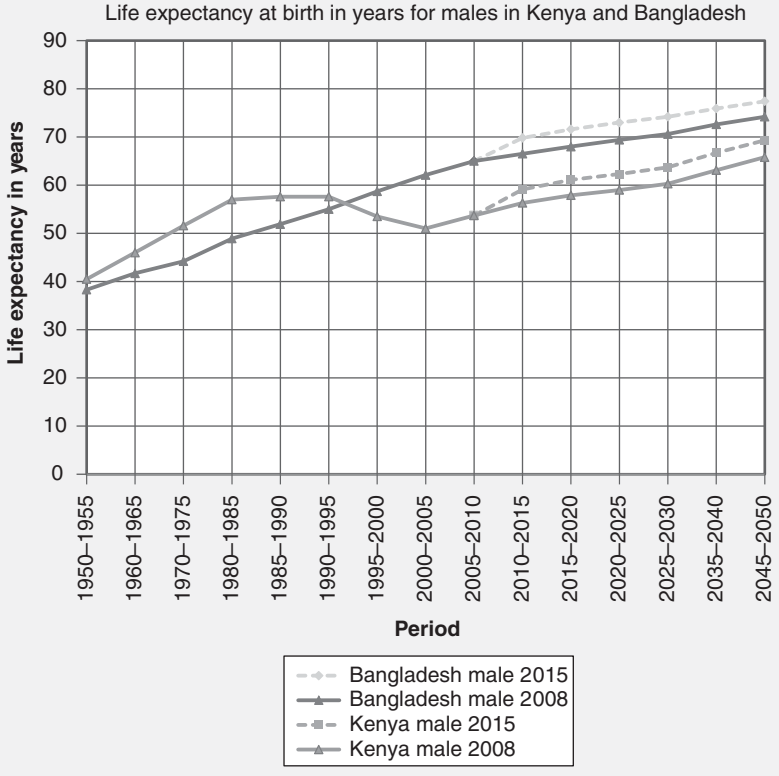
Population growth rates for Bangladesh and Kenya, 1950–2050, projected in 2008 (1950–2050) and in 2015 (2010–2050). All estimates are from the medium variant projection.

Data from United Nations, Department of Economics and Social Affairs, Population Division (2009). *World population prospects: The 2008 revision, volume I: Comprehensive tables (ST/ESA/SER.A/287).* Accessed on November 10, 2016 from http://kczx.shupl.edu.cn/download/786444c9-20c1-4b5a-b0d6-d7544569a2ee.pdf. United Nations, Department of Economics and Social Affairs, Population Division (2015). *World population prospects: The 2015 revision, volume I: Comprehensive tables (ST/ESA/SER.A/379).* Accessed on November 10, 2016 from https://esa.un.org/unpd/wpp/publications/Files/WPP2015_Volume-I_Comprehensive-Tables.pdf



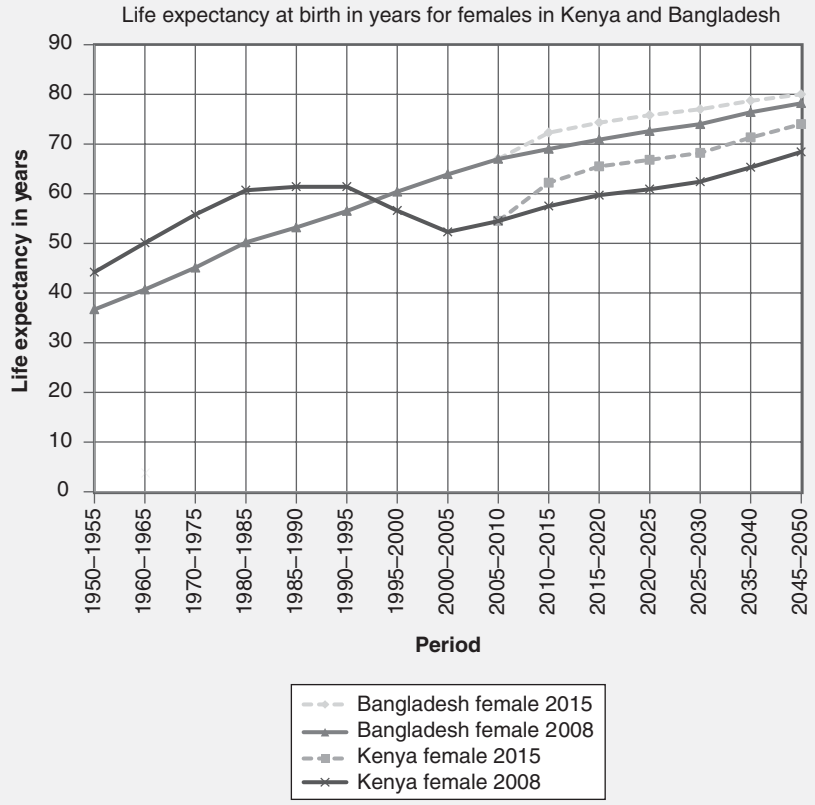
Total fertility rate (per woman) for Bangladesh and Kenya, 1950–2050, projected in 2008 (1950–2050) and 2015 (2010–2050). All estimates are from the medium variant projection.

Data from United Nations, Department of Economics and Social Affairs, Population Division (2009). *World population prospects: The 2008 revision, volume I: Comprehensive tables (ST/ESA/SER.A/287).* Accessed on November 10, 2016 from http://kczx.shupl.edu.cn/download/786444c9-20c1-4b5a-b0d6-d7544569a2ee.pdf. United Nations, Department of Economics and Social Affairs, Population Division (2015). *World population prospects: The 2015 revision, volume I: Comprehensive tables (ST/ESA/SER.A/379).* Accessed on November 10, 2016 from https://esa.un.org/unpd/wpp/publications/Files/WPP2015_Volume-I_Comprehensive-Tables.pdf



Life expectancy at birth in years for males in Bangladesh and Kenya, 1950–2050, projected in 2008 (1950–2050) and 2015 (2010–2050). All estimates are from the medium variant projection.

Data from United Nations, Department of Economics and Social Affairs, Population Division (2009). *World population prospects: The 2008 revision, volume I: Comprehensive tables (ST/ESA/SER.A/287).* Accessed on November 10, 2016 from http://kczx.shupl.edu.cn/download/786444c9-20c1-4b5a-b0d6-d7544569a2ee.pdf. United Nations, Department of Economics and Social Affairs, Population Division (2015). *World population prospects: The 2015 revision, volume I: Comprehensive tables (ST/ESA/SER.A/379).* Accessed on November 10, 2016 from https://esa.un.org/unpd/wpp/publications/Files/WPP2015_Volume-I_Comprehensive-Tables.pdf



Life expectancy at birth in years for females in Bangladesh and Kenya, 1950–2050, projected in 2008 (1950–2050) and 2015 (2010–2050). All estimates are from the medium variant projection.

Data from United Nations, Department of Economics and Social Affairs, Population Division (2009). *World population prospects: The 2008 revision, volume I: Comprehensive tables (ST/ESA/SER.A/287).* Accessed on November 10, 2016 from http://kczx.shupl.edu.cn/download/786444c9-20c1-4b5a-b0d6-d7544569a2ee.pdf. United Nations, Department of Economics and Social Affairs, Population Division (2015). *World population prospects: The 2015 revision, volume I: Comprehensive tables (ST/ESA/SER.A/379).* Accessed on November 10, 2016 from https://esa.un.org/unpd/wpp/publications/Files/WPP2015_Volume-I_Comprehensive-Tables.pdf

regions' growth rates—and their population growth—increased. Fertility continued to fall, however, and at sufficient rates to counteract the continuing increased in life expectancy. This fertility transition is serving to bring growth rates down in all low- and middle-income regions today. HIV/AIDS may be contributing to recent declines in fertility in the parts of the world hardest hit by this disease; a mounting body of evidence indicates that infected women have reduced fecundity (Lewis, Ronsmans, Ezeh, & Gregson, 2004; United Nations Population Division, 2002). The changes in population growth rates, total fertility rates, and life expectancy are illustrated in Exhibit 5-1, which provides information for Bangladesh and Kenya from 1950 to 2050 using the medium-variant UN population estimates (United Nations Population Division, 2015b).

To understand the different types of fertility transitions that have taken place, we need an understanding of the determinants of fertility and fertility change in different contexts. An extensive body of literature has examined the impact of socioeconomic factors on desired family size (Bankole & Westoff, 1995; Bulatao & Lee, 1983; Rutstein, 1998). Much of this discussion centers on the costs and benefits of children and the notion that couples desire additional children as long as the benefits are greater than the costs. These benefits and costs are, in turn, determined by a range of factors, some of which are *structural* (e.g., wages, rates of return on investments, opportunity costs), and some of which are attitudinal (i.e., changes in values and expectations). Improvements in the educational status of women, for example, are thought to decrease desired family size because such trends increase the potential wages that women can earn and, therefore, raise the opportunity costs of childbearing. Education may, in addition, lead to attitudinal change about quantity–quality tradeoffs in numbers of children—for example, having fewer children so that greater investment in the education of each child is feasible.

Implicit in this theoretical framework is the idea that couples weigh a variety of alternatives, with childbearing being just one of the possible behavioral choices available. Other structural factors affecting fertility rates include trends such as increasing landlessness, which decreases the benefits of the labor provided by children and thereby tends to reduce family sizes. More recent research emphasizes attitudinal

change as affecting fertility rate. It posits that values and expectations can change as a result of outside influences. Thus, exposure to messages in which small families are treated as a marker for modernity may motivate couples to reduce their desired family sizes even in the absence of any changes in the structural costs and benefits of children.

Although this chapter focuses on LMICs, in almost none of which has fertility declined to replacement level, it is worth noting that high-income countries, especially those in Europe, are concerned about their very low fertility rates and population declines. For Europe as a whole, TFR fell to less than 1.9 births per woman before 1980 and has continued to decline. It is expected that the entire continent will have a negative growth rate for 2000–2020 (United Nations Population Division, 2009). Understanding which factors maintain below-replacement fertility and which factors cause it to increase is an important issue for high-income countries.

How Do People Control Their Fertility?

In addition to considering *why* people control their fertility, we need to understand *how* people actually do so. It is useful first to consider the *proximate determinants* that lead to variation in fertility in the absence of deliberate family planning (Bongaarts, 1978, 1983; Menken & Kuhn, 1996; Sheps & Menken, 1973). These proximate determinants can be divided into those that affect the *reproductive span* and those that influence the *intervals between successive births* within that span. As shown in **FIGURE 5-4**, the *effective reproductive span* exists within boundaries set by both the *biological span* and the *social reproductive span.* The biological span is the time during which a woman is capable of childbearing because she has the biological capacity to ovulate and to carry a pregnancy to a live birth. It is usually marked by menarche and menopause, but first ovulation may occur well after menarche and last ovulation precedes menopause.

In no society, however, do women devote their full biological span to reproduction. Were they to do so, according to Bongaarts (1978), women who survive to sterility would bear more than 15 children on average. This figure is well beyond the maximum ever recorded for any population.

Every society has social controls on initiation and cessation of sexual activity. We will refer to entry into sexual activity as *marriage* and cessation as *marriage dissolution*. We use these terms as social markers rather than as representing legal ceremonies and arrangements of the state. Specifically, marriage dissolution can occur through breakup of the relationship or through widowhood. The social reproductive span is, therefore, the interval between initiation and cessation of sexual activity. The effective reproductive span is the overlap of the biological and the social spans. It begins with the later of menarche and marriage and ends with the earliest of sterility, death, and cessation of sexual activity. In many societies, the effective reproductive span is interrupted by time between successive unions or by temporary separation of spouses.

Within the effective reproductive span, the pace of childbearing is determined by the lengths of the successive intervals between births (B1, B2, and so on in Figure 5-4). We will first discuss birth intervals in the

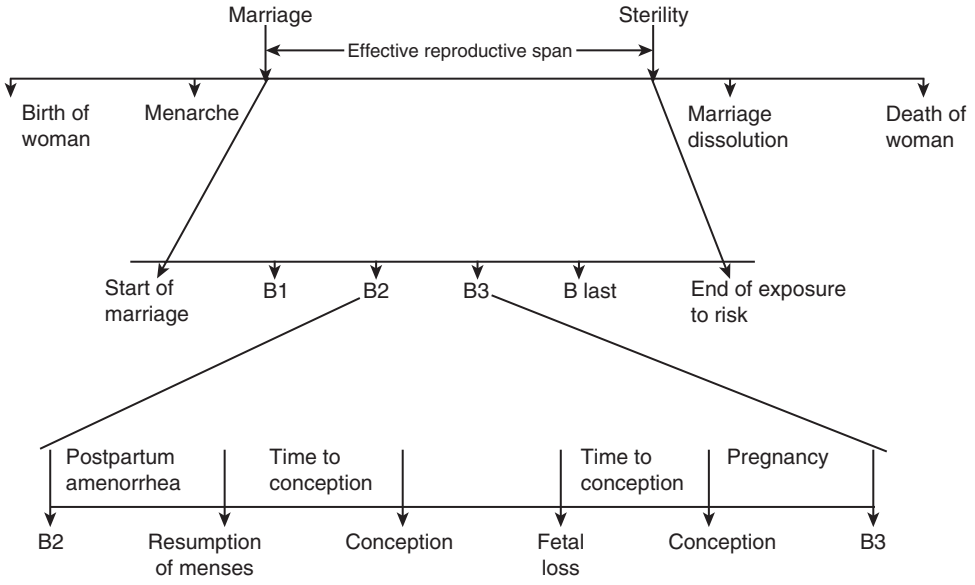


FIGURE 5-4 Reproductive span and birth intervals.

absence of deliberate family planning. The birth interval may be divided into several segments:

- The postpartum period after a birth until both ovulation and sexual relations resume
- The time to conception
- Additional time due to fetal loss through spontaneous abortion
- The pregnancy leading to the next live birth

Fertility in the Absence of Contraception and Induced Abortion

The postpartum period ends when both ovulation and intercourse have resumed. It is largely determined by the duration and intensity of breastfeeding and by postpartum taboos against intercourse by a nursing mother. Women who do not breastfeed usually menstruate for the first time approximately two months after giving birth (Salber, Feinleib, & Macmahon, 1965). Frequent, intense breastfeeding, however, can postpone average time of ovulation to more than 20 months (Wood, Lai, Johnson, Campbell, & Maslar, 1985). Some populations, particularly in sub-Saharan Africa, have traditionally had taboos against intercourse that can lengthen the postpartum period beyond the resumption of ovulation, but these practices are rare outside this region and their observance is believed to be decreasing.

Breastfeeding not only provides the child with nourishment (see the *Nutrition* chapter), but also, depending on the pattern of suckling, can postpone the return of ovulation for many months. It exerts this effect through a maternal response to suckling that suppresses the secretion of gonadotropins. The classic studies of McNeilly (1996) and his colleagues showed that "if the frequency of suckling is maintained above five times a day and the duration is maintained above 65 minutes a day, amenorrhea will often be the consequence." Others have found that night feeds are particularly important in maintaining amenorrhea (Jones, 1988). Women who fully breastfeed their infants and are amenorrheic are highly unlikely to conceive in the first six months after a birth. Indeed, one multinational study estimated that fewer than 2% would do so (Labbok et al., 1997). In addition, demographic studies suggest that the duration of lactational amenorrhea increases with the age of the woman (Wood, 1994). The effects of breastfeeding patterns are so important that they are the major factor in explaining differences in fertility among populations in which no family planning is practiced.

The time to conception depends on the monthly probability of conception in the absence of birth control and can vary among populations, by age, and according to the frequency and pattern of intercourse. The monthly probability of conception, known as *fecundability*, is extremely difficult to measure. Estimates of this probability differ in part because of the methods used to determine whether a conception has occurred. If early fetal loss occurs before the woman is aware she is pregnant, then the estimates of conception are biased downward. According to Wood (1994), the best study of early fetal loss was conducted by Wilcox and associates (1988), who followed a group of women ages 20 to 35 and collected blood and urine samples from them regularly. There were 198 pregnancies detected by assays of these samples; 43 pregnancies, or approximately 22%, were lost before the woman realized she was pregnant and before clinical diagnosis was made (Wilcox et al., 1988; Wood, 1994).

More traditionally, fecundability has been measured by accepting a woman's report of her pregnancy. This measure of *apparent* fecundability has yielded rates that range from approximately 0.10 to 0.30 for relatively young women (Menken, 1975; Wood, 1994). The waiting time to conception is, on average, the inverse of fecundability; for younger women, this time is 3 to 10 months on average. In many cases, fecundability has been estimated from reported waiting times to conception using a variety of mathematical models (some of which are summarized by Wood, 1994).

Fecundability depends in part on a couple's frequency and pattern of intercourse, which determines the likelihood that coitus will occur during the woman's fertile period. Both those couples wishing to conceive a wanted child and those couples hoping to avoid pregnancy without the use of hormonal or barrier contraceptives depend on knowledge of the woman's cycle to time their sexual activity, thereby changing their probability of conception. Whether ovulation is discerned by changes in cervical mucus or basal body temperature, conception is most likely to occur when intercourse takes place shortly before ovulation (Colombo & Masarotto, 2000).

Fecundability declines with the age of the woman, although increasing evidence indicates that this decline does not take place until the late thirties, on average, if patterns of intercourse remain unchanged. Decline with age of the man is quite slow.

Lactation has a fertility-reducing effect even after a woman has resumed menstruation and ovulation (Wood, 1994). Apparently, continued suckling beyond ovulation reduces fecundability through a response that interferes with the functioning of the corpus luteum. Although fertilization can occur, the corpus luteum may not produce sufficient progesterone to enable the pregnancy to continue (McNeilly, 1996).

Infection with HIV also affects fertility. Notably, it reduces conception rates beginning in the earliest asymptomatic stages of infection (Gray et al., 1998; Ross et al., 2004; United Nations Population Division, 2002).

Spontaneous abortion occurs frequently; the estimated rate, as already described, depends on how early the pregnancy is detected. From the time conceptions are recognizable by virtue of late menses, approximately 24% end in spontaneous abortion (French & Bierman, 1962). Even higher proportions of fertilized ova do not lead to live births. Wilcox and associates (1988), in the study cited earlier in which urine specimens were collected so that early pregnancy could be detected, found that 31% of the pregnancies ended in fetal loss. Rates of spontaneous abortion increase with the age of the woman (Nybo Andersen, Hansen, Andersen, & Davey Smith, 2004; Wood, 1994) and, more slowly, with the age of the man (Nybo Andersen et al., 2004; Slama et al., 2003). The time added to the birth interval by a recognized fetal loss is the sum of the time from pregnancy to the next ovulation (usually estimated to be slightly more than 3 months, on average, as the vast majority of spontaneous abortions occur very early in pregnancy) plus the time to the next conception. There is little evidence that rates of spontaneous abortion vary to any great extent among populations.

Gestation leading to a live birth does not vary much, usually lasting between 35 and 40 weeks, with few differences noted between population groups (Wood, 1994, p. 207). In some populations with outstanding care of premature infants or increasing maternal risks, gestation may be somewhat shorter on average. In the United States, the percentage of live births reported as preterm (fewer than 37 weeks' gestation) increased from 9.4% in 1981 to 12.1% in 2002 (a jump of 29%) and subsequently declined to 9.6% in 2015 (March of Dimes, 2016; Martin et al., 2003). In 2010, the highest rates of preterm births occurred in Southeastern Asia (13.5%), South Asia (13.3%), Africa (12.3%), and North America (12.0%), with Northern Europe (5%) having the lowest rate (Blencowe et al., 2016).

A cross-cutting issue that affects fertility rates is that of infertility and sterility. Especially in parts of Africa, infertility and early sterility are major factors affecting fertility. This effect may take the form of absolute sterility that causes the effective reproductive span to end early, or decreased fecundability or increased risk of spontaneous abortion (Larsen, 2000). In most cases, however, the concern is with early sterility, much of which is believed to be due to sexually transmitted infections.

Thus, the main reasons that populations not practicing family planning vary in terms of their fertility rates are differences in the effective reproductive span and the duration of the postpartum amenorrheic period, with variations in time to conception and infertility playing lesser roles.

Deliberate Control of Fertility

People can deliberately reduce their fertility in three ways: (1) by reducing the effective reproductive span through postponement of marriage or interrupted marriage or by sterilization that ends reproductive capacity early; (2) by using contraception, which increases the time to conception; and (3) by induced abortion, which increases the time added to the birth interval by ensuring that pregnancies do not lead to a live birth. Family planning programs can both promote the motivation to reduce fertility and provide the means to do so. Many governments encourage or enforce later marriage explicitly through changes in the legal age of marriage and implicitly through programs that foster female education. Family planning programs have traditionally focused on education regarding methods of fertility control, motivation to reduce the number of wanted children, and provision of family planning methods themselves. Such methods include promotion of breastfeeding, both for the health of the infant and to prolong the postpartum period; contraception, which is intended to prolong the time to conception; and, except where there is opposition for religious reasons, abortion, which increases the time added to the birth interval.

Although specific family planning methods are discussed in greater detail in the section on family planning programs later in this chapter, it seems appropriate to consider two important general issues here. First, why do populations in which the desired number of children is low still have high proportions of unintended births? Second, why is reliance on abortion an inefficient approach to family planning?

Contraception, even when highly effective, may not prevent all unintended pregnancies. A simple calculation makes this problem clear. Suppose a woman is using a highly effective method (0.5% failure)—one that reduces her probability of conceiving to about 1 in 200 per month. She begins using this method at age 30 and wants no more children before she reaches menopause at age 45. We can calculate the probability that she has no pregnancy in each of 13 lunar months over the next 15 years, or a total of 195 lunar months. The probability of succeeding (not getting pregnant) is 0.995 each month. The probability of not getting

pregnant in 195 months is 0.995^{195}, which equals 0.38. In other words, this woman has only a 38% chance of avoiding pregnancy for 15 years. Among women like her, 62% will have at least one unintended pregnancy in that time period. For this reason, even women who are very serious users of contraception are at high risk of unintended pregnancy. Family planning program and health planners need to be aware of this high risk when they are developing their programs.

In a population that relies primarily on induced abortion to reduce fertility, if a woman becomes pregnant unintentionally, she may choose to have an abortion. Approximately three months after the abortion, she again begins ovulating and is capable of conceiving. Suppose her time to conception averages 10 months. Then 13 months after the first abortion, she is again pregnant and must have another abortion if she is not to have an unwanted birth. Another 13 months later, an abortion is again performed, and so on. Preventing a birth for 15 years may, therefore, require 15 abortions.

Given this pattern, it is not surprising that many women in Eastern European countries who relied primarily on abortion for birth control reported numbers of abortions in the double digits. Women in the former Soviet Union, for example, are believed to have had six or more abortions, on average, over the course of their lifetime (David, 1992). While abortion rates in Eastern Europe have declined sharply in the last decade or so with increased access to modern contraception services, they remain among the highest in the world (42 women undergoing abortions per 1,000 women of childbearing age; 105 abortions for every 100 births) (Alan Guttmacher Institute, 2009a, 2016). Abortion, is, however, extremely effective as a backup to effective contraception. A woman who has an abortion and subsequently uses extremely effective contraception is unlikely to have more than one or two unintended pregnancies—but our previous analysis shows that she may, indeed, have these one or two.

For these reasons, it is not surprising that sterilization—the one method that has a failure rate near zero—is so widely selected by women and couples who want no additional children. In the United States in 2002, sterilization of the woman or of the man was the model method used by women older than age 30 (Mosher, Martinez, Chandra, Abma, & Willson, 2004). Reliance on the sterilization method in LMICs has declined over time. Approximately 26% of married women relied on sterilization to prevent additional births (22% of wives, and an additional 4% of their husbands had been sterilized) in 2002, but the rates declined to 12.8% female sterilization and 1.3% male sterilization in 2012 (Population Reference Bureau, 2002; Ross, Keesbury, & Hardee, 2015).

The Effect on Fertility of the Proximate Determinants: Bongaarts' Indices

Bongaarts (1978) developed a set of indices to measure the effects on fertility of some proximate determinants. These indices are based on the assumption that there is some maximum potential fertility, *TF*, for women. This figure is usually estimated to be slightly more than 15 children.

Ci	The **index of postpartum infecundity** varies from 0 to 1. It represents the proportion of potential fertility, *TF*, remaining when the average postpartum period of the population of interest is taken into account. Therefore, $Ci = 1$ if the population does not breastfeed at all. The fertility-reducing effect of postpartum infecundity is $(1 - Ci)$.
CA	The **index of abortion** is the proportion of *TF*, after postpartum infecundity is first taken into account, remaining when the effect of induced abortion in reducing live births is taken into account. Spontaneous abortions are included in the original estimate of *TF*, because they are treated as a purely biological occurrence. Few countries have sufficient information available on abortion to make reasonable estimates of *CA*, so it usually must be disregarded in application.
CC	The **index of contraception** is the proportion of *TF*, after the effects of postpartum infecundity and induced abortion are taken into account, remaining after contraceptive use is considered.
Cm	The **index of marriage** is the proportion of *TF*, after the first three factors are considered, remaining when the particular marriage or sexual union pattern is taken into account.

Reproduced from United Nations (1994). Programme of action of the 1994 International Conference on Population and Development. Paper presented at the (A /CONF.171/13), http://www.un.org/popin/icpd/conference/offeng/poa.html

Thus, in the Bongaarts decomposition of the total fertility rate,

$$TFR = TF \leftrightarrow Ci \leftrightarrow CA \leftrightarrow CC \leftrightarrow Cm$$

Both *CC* and *Cm* contain adjustments for infertility and sterility. In the first index, the adjustment takes into account infertility and sterility and assumes no use of contraception by infertile and sterile couples. In the second index, a weighting factor is present, in that nonmarriage has a greater effect on fertility reduction when the woman is young (e.g., the effect of nonmarriage is much greater for a 25-year-old woman than for a 42-year-old woman).

TABLE 5-2 presents these indices, except for *CA*, which is assumed to be 1 because of the lack of data, for a number of populations around 1970 and for several historical populations (Bongaarts & Potter, 1983). The major impact of breastfeeding can be seen through the values of the index *Ci* in countries whose populations used little contraception during the period in question. All of the South and East Asian countries, as well as Kenya, have indices that do not exceed 0.67; thus,

TABLE 5-2 Estimates of Total Fertility Rate and Bongaarts' Proximate Determinants Indices*

Region (Year)	Total Fertility Rate (*TFR*)	Index of Postpartum Infecundity (*Ci*)	Index of Marriage (*Cm*)	Index of Contraception (*CC*)
Low and Middle-Income Countries				
Bangladesh, 1975	6.34	0.54	0.85	0.90
Colombia, 1976	4.57	0.84	0.58	0.61
Dominican Republic, 1975	5.85	0.61	0.60	1.00
Indonesia, 1976	4.69	0.58	0.71	0.75
Jordan, 1976	7.41	0.80	0.74	0.81
Kenya, 1976	8.02	0.67	0.77	1.00
Korea, 1970	3.97	0.66	0.58	0.68
Lebanon, 1976	4.77	0.78	0.58	0.69
Sri Lanka, 1975	3.53	0.61	0.51	0.74
Syria, 1973	7.00	0.73	0.73	0.86
Thailand, 1975	4.70	0.66	0.63	0.74
Industrialized Countries				
Denmark, 1970	1.78	0.93	0.55	0.23
France, 1972	2.21	0.93	0.52	0.30
Hungary, 1966	1.80	0.93	0.62	0.21
United Kingdom, 1967	2.38	0.93	0.61	0.27
United States, 1967	2.34	0.93	0.63	0.26

Historical Populations				
Bavarian villages, 1700–1850	4.45	0.85	0.37	0.91
Grafenhausen, 1700–1850	4.74	0.67	0.44	1.00
Hutterites	9.50	0.82	0.73	1.00
Quebec, 1700–1730	8.00	0.81	0.63	1.00

* Each index represents the proportion of potential fertility remaining after the particular factor is taken into account, in the following order: postpartum infecundity, marriage, contraceptive use.
Adapted with permission of Population Council, from John Bongaarts, "The Fertility-Inhibiting Effects of the Intermediate Fertility Variables," Studies in Family Planning, 13 (6/7), pp. 182–183, 1982; with permission of BLACKWELL PUBLISHING, INC. in the format Book via Copyright Clearance Center.

EXHIBIT 5-2 Proximate Determinants of Fertility in Kenya

The fertility rate in Kenya was estimated to be approximately 8.20 births per woman in 1978. This high fertility was due to the combination of very high rates of marriage for women of reproductive age (*Cm* = 0.91) and essentially no fertility control within marriage (C+ = 0.96). Fertility could, however, have been much higher except for prolonged amenorrhea, which is related to the long breastfeeding practiced in this country (*Cf* = 0.64). Total fertility, at 8.20 births per woman, was slightly more than half (0.56) of what it would have been if women had not breastfed at all (TFR = 14.7) (Toroitich-Ruto, 2001).

The focus of the Kenyan Family Planning Program has been to increase fertility control within marriage by increasing the use of contraception by married women. This emphasis paid off at least initially, resulting in sharp declines in TFR between 1978 and 1989 (from 8.20 births per woman to 6.60 births per woman). Estimates of Bongaarts' proximate determinants of fertility for Kenya in 1989 compared to 1978 support the notion that this sharp decline in TFR was essentially due to a rise in contraceptive prevalence rates. Bongaarts' estimates show a significantly lower index of contraception (*CC*) of approximately 0.80 in 1989 compared to 0.96 in 1978, reflecting the rapid rise in contraceptive prevalence over a period of 12 years (33% in 1989 versus 7% in 1978); a relatively unchanged index for marriage (*Cm* = 0.86 in 1989 versus 0.91 in 1978), reflecting the continuing high rates of marriage; and a relatively unchanged index of postpartum infecundity (0.66 in 1989 versus 0.64 in 1978), reflecting continued long breastfeeding durations (Toroitich-Ruto, 2001).

Total fertility rates continued to decline relatively rapidly until 1993 (TFR = 5.4 births per woman). Subsequently, fertility decline stalled. In 2014 (the most recent data available), TFR in Kenya was 3.9 births per woman, with contraceptive prevalence rates having risen to 58%, up from 33% in 1989 (Macro International, 2014).

their potential fertility is reduced by at least one-third by long postpartum periods. In fact, the long breastfeeding periods employed in Bangladesh and Indonesia reduced these countries' fertility rates to only about half of their potential. In Europe, the demographic transition to low levels of fertility was caused, to a great extent, by very late marriage and a relatively high degree of nonmarriage. The index of marriage is far lower, on average, for high-income countries around 1970 than for LMICs. Two historical populations shown, however, had indices of marriage less than 0.45, indicating that nonmarriage reduced their potential fertility by at least 55%. By 1970, fertility in all the developed countries shown was reduced by contraceptive use to no more than 30% of its potential level.

Thus, in high-income countries, breastfeeding has little effect on fertility, but nonmarriage and use of contraception reduce the TFR to relatively low levels. In the 1970s, contraceptive use had little impact on fertility in LMICs included in Table 5-2; lower fertility was achieved in some of these countries through long-term breastfeeding (**EXHIBIT 5-2** provides information on the example of Kenya). Populations that had very high TFRs achieved these rates through a combination of high indices of marriage and breastfeeding and little or no contraception.

Stover's Revision of Bongaarts' Indices

The indices of the effects of the proximate determinants on fertility have been revised a number of times by Bongaarts and others to take advantage of more recent detailed and reliable data and to substitute more realistic assumptions. Stover (1998), for example, dropped the original indices of marriage and contraceptive use; he also attempted to treat infertility and sterility more directly and to deal with sexual activity rather than using the proxy of marriage or stable sexual union incorporated in the earlier version of the indices (Stover, 1998). He includes, instead, three new indices:

Cx	The **index of sexual activity** depends on the reported proportion of women in the population who are sexually active. Its interpretation is, therefore, the proportion of potential fertility remaining after celibacy is taken into account.
Cf	The **index of infecundity** reflects the effect on fertility of infecundity among sexually active women, and is simply $1 - f$, where f is the proportion who report that they believe themselves to be infecund.
Cu	The **index of contraceptive** use reflects actual contraceptive use by women who believe themselves to be fecund and who are not experiencing postpartum amenorrhea.

Thus, the fertility-reducing effect of the proximate determinants is given by

$$Cx \leftrightarrow Ci \leftrightarrow CA \leftrightarrow Cu \leftrightarrow Cf$$

where the index of abortion (CA) is rarely estimated and Ci is the index of postpartum infecundity as previously defined by Bongaarts.

These results show that the two main factors producing lower fertility in Latin America are relatively low participation in sexual activity and relatively high contraceptive use (**TABLE 5-3**). By contrast, in most countries of Africa, there is much higher participation in sexual activity and lower use of contraception. Fertility would be even higher were it not for the effects of postpartum infecundity, which reduces fertility by at least 40% in the countries represented here. What is striking is the documentation of the rather large impact on overall fertility (approximately 20%) in

TABLE 5-3 Estimates of TFR and Bongaarts' Revised Proximate Determinants Indices

Region (Year)	TFR	Index of Sexual Activity (*Cx*)	Index of Postpartum Infecundity (*Ci*)	Index of Infecundity (*Cf*)	Index of Contraceptive Use (*Cu*)
Africa					
Burkina Faso, 1993	6.9	0.66	0.49	0.88	0.94
Cameroon, 1991	5.8	0.69	0.57	0.81	0.84
Ghana, 1993	5.5	0.64	0.55	0.86	0.81
Madagascar, 1992	6.1	0.71	0.61	0.83	0.89
Namibia, 1992	5.4	0.59	0.59	0.86	0.70
Niger, 1992	7.4	0.86	0.58	0.78	1.00
Nigeria, 1990	6.0	0.73	0.53	0.80	0.93
Rwanda, 1992	6.2	0.60	0.56	0.85	0.86
Senegal, 1993	6.0	0.63	0.56	0.81	0.93
Zambia, 1992	6.5	0.69	0.60	0.86	0.90
Latin America and Caribbean					
Brazil, 1991	3.7	0.59	0.83	0.88	0.39
Colombia, 1990	2.9	0.53	0.77	0.89	0.36
Dominican Republic, 1991	3.3	0.54	0.81	0.88	0.41
Paraguay, 1990	4.7	0.49	0.76	0.86	0.47
Peru, 1992	3.5	0.55	0.68	0.89	0.53

Modified from Population Council from Stover, J. (1998). Revising the proximate determinants of fertility framework: What have we learned in the past 20 years? *Studies in Family Planning, 29*(3), p. 263.

some countries due to infertility among sexually active women.

Family Planning Programs

A fundamental rationale for family planning programs is to reduce unintended fertility because of its negative health and welfare consequences and because control of fertility has been recognized as a human right of women and couples. Over the last 50 years, societal changes have included reduced infant mortality, increased urbanization, improved education for women, increased economic opportunities, and the dissemination and adoption of modern ideas about small families. The response by couples in LMICs has been accelerated change in expectations about both the number and the timing of births (Bongaarts, 1983; Freedman, 1987).

In 1997, 155 countries had programs that subsidized the cost of family planning services (Gelbard, Haub, & Kent, 1999). However, in part due to the lack of available, accessible, and effective contraception, the gap between observed and desired fertility grew, leading in turn to an increase in unintended fertility (Bankole & Westoff, 1995; Bulatao, 1998). Out of 213 million pregnancies in 2012, 85 million women become pregnant unintentionally (Sedgh, Singh, & Hussain, 2014). Due to concerted family planning efforts in the last three decades to increase access to contraceptive services (worldwide, such access increased from approximately 10% in the early 1960s to 59% in 2000 and 63.6% in 2015), many more women are now able to avoid unintentional pregnancies (United Nations Population Division, 2015a; WHO, 2005). For example, in Kenya, 35.5% of women expressing a desire to control their fertility were not able to meet this need in 1989. By 2003, this figure had decreased to 24.5% and by 2014 to 17.5% (**EXHIBIT 5-3**).

Unintended Fertility

A number of definitional issues complicate the process of estimating unintended fertility and its distribution. In general, data on unintended fertility come from representative population surveys in which women who are pregnant at the time of the survey or who have had at least one birth in the five years prior to the survey are asked whether each of those births (including the outcome of the current pregnancy) was intended, *mistimed* because it came too early but was still within the desired number of births, or *unwanted* in that no more children were desired. Unfortunately, usage of these terms is not completely consistent; some authors include mistimed births as part of their estimates of unwanted births, while others count only those births that exceed the desired family size as unwanted (Brown & Eisenberg, 1995). The major weakness of this approach is that women may be reluctant to classify specific births as unwanted, leading to artificially low estimates of unwanted births.

Measurement of fertility intentions has been criticized because it relies exclusively on mothers' intentions and not the intentions of other family members—most importantly, fathers—to gauge unintended fertility.

EXHIBIT 5-3 Desired Family Size and Unmet Need for Contraception in Kenya

One rationale for family planning programs is that couples who want to have fewer children may be unable to do so either because they lack knowledge of the means of fertility control or because they lack access to those means, owing to an absence of supplies or services.

In the case of Kenya, even in the earliest surveys carried out, a high proportion of married women reported that they wanted no more children. At that time, the mean ideal number of children was reported to be 4.4, and 45.5% of women reported they did not want to have another child. The desired TFR (wanted fertility) was 4.5 births per woman, while the actual TFR was 6.7 births per woman. Thus, there was a big gap between desired fertility and actual fertility despite the inception in late 1970s of a nationwide family planning program. Among the barriers to achieving this low desired fertility were (1) the economic costs of access to services, including the cost of transportation, and supplies; (2) the social costs, including travel by women whose mobility was traditionally constrained; (3) the psychic costs of contraceptive use in a society that offered little social or familial support for low fertility; and (4) the health costs of side effects, whether subjective or objective, from contraceptive use.

These barriers have been overcome to the extent that in 2015, 58% of married women of reproductive age were current contraceptive users, compared to 43% in 2003 and 33% in 1989. Also, actual TFR in Kenya declined from 6.7 births per woman in 1989 to 4.9 births per woman in 2003 and then to 3.9 births per woman in 2014. However, wanted fertility did not remain at the 1989 level: It fell to 3.6 births per woman in 2003 and 3.0 births per woman in 2014. The unmet need for contraception among women at risk of pregnancy decreased from 35.5% in 1989 to 24.5% in 2003 and 17.5% in 2014.

It has been argued that the intentions of the mother, especially in many LMICs, may not accurately reflect the desirability of a birth. Evidence suggests that intergenerational differences in family size goals (i.e., preferences of grandparents versus parents) may be more pronounced than interspousal differences (Caldwell, 1986; Mason & Taj, 1987). Ultimately, the justification for relying on the stated preferences of the mother in determining desired family size and unintended or unwanted births stems from the fact that the mother is the person most responsible for the birth and child care (Tsui et al., 1997).

Demographic and Health Surveys (DHS) have been conducted since 1984 in approximately 75 LMICs (https://dhsprogram.com/). This research is intended to provide comparable information on a variety of subjects related to health and fertility issues. DHS calculates, for each survey, both the total fertility rate (the number of children a woman would bear were she to live her reproductive life under the fertility conditions just prior to the survey) and the unwanted total fertility rate (the number of those children who would be unwanted) (Westoff, 2001). A birth is counted as unwanted only when the mother states she wanted no more children at the time of the pregnancy.

The results of DHS surveys conducted over the period 2010–2015 in 35 countries are shown in **TABLE 5-4**. On average, 17.7% of total fertility was unwanted, with the percentage varying considerably by region. The association between fertility level and proportion of unwanted births has been reversed in the last decade. Previously the regions with the lowest and highest fertility rates (Eastern Europe and sub-Saharan Africa, respectively) had the lowest

TABLE 5-4 Total Fertility Rate and Percentage of Births Unwanted, by Mother's Education, in Countries with DHS Surveys Around 2010 and Beyond

	No Education		Primary Education		Secondary or Higher Education		Total	
	TFR	Percentage Unwanted	TFR	Percentage Unwanted	TFR	Percentage Unwanted	TFR	Percentage Unwanted
Sub-Saharan Africa								
Benin, 2011–2012	5.6	17.9	4.6	19.6	3.8	13.2	4.9	18.4
Cameroon, 2011	6.8	10.3	5.9	11.9	3.8	10.5	5.1	11.8
Chad, 2014–2015	6.5	4.6	7.4	4.1	4.8	4.2	6.4	4.7
Congo, 2011–2012	6.8	11.8	6.6	7.6	4.5	6.7	5.1	5.9
Democratic Republic of the Congo, 2013–2014	7.4	9.5	7.5	13.3	5.6	14.3	6.6	13.6
Ethiopia, 2011	5.8	19.0	4.6	26.1	1.6	6.3	4.8	20.8
Ghana, 2014	6.2	11.3	4.9	16.3	3.5	14.3	4.2	14.3
Guinea, 2012	5.7	8.8	5.1	9.8	3.0	10.0	5.1	9.8
Kenya, 2014	6.5	6.2	4.4	27.3	3.0	20.0	3.9	23.1
Lesotho, 2014	1.9	31.6	4.0	32.5	2.9	24.1	3.3	30.3
Liberia, 2013	5.9	11.9	5.1	11.8	3.4	11.8	4.7	10.6

Madagascar, 2008–2009	6.4	10.9	5.3	13.2	3.1	12.9	4.8	12.5
Malawi, 2010	6.9	18.8	5.9	20.3	3.6	16.7	5.7	21.1
Mali, 2012–2013	6.5	13.8	5.9	15.3	4.0	10.0	6.1	13.1
Namibia, 2013	5.3	26.4	4.8	27.1	3.3	15.2	3.6	19.4
Niger, 2012	8.0	2.5	7.0	5.7	4.9	4.1	7.6	2.6
Nigeria, 2013	6.9	2.9	6.1	6.6	4.2	7.1	5.5	5.5
Rwanda, 2014–2015	5.1	25.5	4.5	26.7	3.0	16.7	4.2	26.2
Senegal, 2014	6.2	11.3	4.4	9.1	3.2	6.3	5.0	10.0
Tanzania, 2010	7.0	10.0	5.6	14.3	3.0	10.0	5.4	13.0
Uganda, 2011	6.9	23.2	6.8	25.0	4.8	18.8	6.2	24.2
Zambia, 2013–2014	7.2	13.9	6.3	14.3	3.8	13.2	5.3	15.1
North Africa/West Asia/Europe								
Armenia, 2010	—	—	1.9	15.8	1.7	5.9	1.7	5.9
Egypt, 2014	3.8	21.1	3.6	22.2	3.5	20.0	3.5	20.0
Jordan, 2012	3.0	36.7	3.9	30.8	3.5	28.6	3.5	28.6
South and Southeast Asia								
Bangladesh, 2014	2.4	37.5	2.4	29.2	2.2	18.2	2.3	26.1
Cambodia, 2014	3.3	12.1	3.1	9.7	2.3	4.3	2.7	11.1
Indonesia, 2012	2.8	10.7	2.9	13.8	2.6	15.4	2.6	15.4
Nepal, 2011	3.7	32.4	2.7	29.6	1.9	21.1	2.6	30.8
Pakistan, 2012–2013	4.4	20.5	4.0	20.0	2.9	17.2	3.8	21.1
Latin America and Caribbean								
Colombia, 2010	4.3	39.5	3.2	34.4	2.0	20.0	2.1	23.8
Dominican Republic, 2013	5.1	25.5	3.2	28.1	2.3	13.0	2.5	20.0
Haiti, 2012	5.4	40.7	4.3	37.2	2.6	19.2	3.5	34.3
Honduras, 2011–2012	4.1	29.3	3.5	25.7	2.4	16.7	2.9	24.1
Peru, 2012	4.7	46.8	3.5	37.1	2.3	21.7	2.6	30.8

Data from Macro International Inc., 2010–2015, MEASURE DHS STAT compiler. Accessed on November 10, 2016 from http://www.measuredhs.com

proportions of unwanted births (the mean percentages were 8% for Eastern Europe and 16% for sub-Saharan Africa). Now the low-fertility region (Latin America and Caribbean) with average TFR of 2.72 births per woman has a 26.6% rate of unwanted births compared to a 14.8% rate of unwanted births in sub-Saharan countries with average TFR of 5.16 births per woman. This change is possibly due to faster decline of desired fertility than actual fertility in low-fertility regions, while in high-fertility countries both desired and actual fertilities decreased in the same pace. Variations in this regard are also evident by countries within and between regions. Some countries with high total fertility rates had low proportions of unwanted births (e.g., Niger), whereas Uganda has high proportion of unwanted births despite a relatively high TFR. The highest percentages of unwanted births (one-fourth or more) were found in countries with TFRs between 2.5 and 5.5 births per woman. Finally, in low-fertility countries where the TFR is less than 4 births per woman, a substantial proportion of those births remain unwanted (Macro International, 2015).

Note that this and earlier evidence suggest that increases in contraception prevalence rates do not necessarily cause a decline in the proportion of unwanted births and, in fact, may initially be associated with an increase in the proportion of unwanted births (Tsui et al., 1997). This scenario can happen if desired fertility rates drop faster than the compensating rises in contraceptive prevalence rates, and if the use of methods of fertility control is so effective that unintended births rarely occur. Thus, when couples have very high fertility desires, it is difficult to exceed those desires, so nearly all children are wanted. As desired fertility falls, use and effective use of contraception and abortion may not increase quickly enough to avoid unwanted births. Finally, at low levels of fertility, women want so few children that there remain many years after the last wanted child during which an unwanted pregnancy and birth can, and frequently do, occur. The differences by education within a country can, in part, be explained by this phenomenon: More-educated women within a society frequently are earlier adopters of contraception and, because they desire few children, a higher proportion of their children are unwanted.

With regard to mistimed births, DHS data from selected countries around 2005 (2002–2010) suggest that approximately 16% of births in LMICs were mistimed—that is, they came too early (Bradley, Croft, & Rutstein, 2011). Highest proportion of mistimed births is reported from Latin America (29%), followed by Africa (20%), Asia (15%), the Middle East (11.5%), and Europe (10%) (Bradley et al., 2011). There is no clear association between contraceptive prevalence and the proportion of mistimed births. Thus, even in a region such as sub-Saharan Africa, where contraceptive use rate is low, considerable mistimed fertility is still observed. Contraceptive failures contributed only 29% of total mistimed births, while 71% of such births resulted from lack of contraceptive use (Bradley et al., 2011). This finding suggests that there is a need for contraception to delay first births and control spacing of subsequent births even when there may be little demand for contraception to control the number of births (Bankole & Westoff, 1995; Rafalimanana & Westoff, 2001).

Consequences of Unintended Pregnancies and Births

Aside from helping individual couples fulfill their desires and expectations, why should we care about unintended pregnancies and births and, moreover, try to reduce them? A compelling public health reason is their negative consequences. Unintended pregnancies increase the lifetime risk of maternal mortality simply by increasing the number of pregnancies (Alan Guttmacher Institute, 2009a; Koenig, Fauveau, Chowdhury, Chakraborty, & Khan, 1988). Unintended pregnancies can lead to unsafe abortion, poor infant health, and lower investment in the child.

Abortion

In many cases, unintended pregnancies are terminated by abortion, which was the outcome for an estimated 25% of all pregnancies in 2010–2014 (Alan Guttmacher Institute, 2016). Most abortions, especially those in LMICs, continue to be conducted illegally under unhygienic conditions and pose significant health risks to the mother.

The legal status of abortion and, consequently, access to safe abortion services is highly variable. Roughly 26% of women of childbearing age (15 to 44 years) in 66 countries currently live in countries where abortion is highly restrictive (Center for Reproductive Rights, 2013). Women in LMICs (except for China and India) are much more likely to live under restrictive abortion laws than women in high-income countries. As of 2013, out of 199 countries, only 61 countries allowed abortion without any restriction; in contrast, in 66 countries, abortion remained illegal even in the face of incest or rape (Center for Reproductive Rights, 2013). There have been some positive

changes in access to legal abortion since 1995, with more than 30 new countries broadening the grounds under which abortion can be performed, and only a few counties tightening legal restrictions (Center for Reproductive Rights, 2014; Reed, Koblinsky, & Mosley, 2000).

Despite these differences in legal access, there is little difference in the likelihood of having an abortion. Thus, in 2010–2014, the abortion rate was 34 per 1,000 women of childbearing age (15 to 44 years) in Africa, where abortion is fairly restrictive, compared to 30 per 1,000 women in Europe, where it is very accessible (Alan Guttmacher Institute, 2016). As a result, women in LMICs (excluding China) are more likely to have illegal abortions, and their abortion mortality rate is many times higher than that for women in high-income countries. Compared to the rate of 0.6 death per 100,000 abortions in the United States, the global abortion rate is 220 deaths per 100,000 procedures; it is a staggering 460 deaths per 100,000 abortions in Africa (Alan Guttmacher Institute, 2016).

Sedgh et al. (2016) estimated that approximately 56 million abortions were performed worldwide yearly between 2010 and 2014, which is about 4 million higher than the annual estimates of 52 million in the 2005–2009 period (Sedgh et al., 2016). Of these 56 million abortions, 34 million were legal abortions and 22 million were unsafe (illegal) procedures (WHO, 2017). The proportion of abortions that are unsafe decreased from 49% in 2008 to 39% in 2010–2014, while the overall rate of abortion has decreased from 40 abortions per 1,000 women in 1990–1994 to 35 abortions per 1,000 women in 2010–2014 (Sedgh et al., 2016). Overall, the decrease in the worldwide abortion rate was largely due to reductions in safe abortions, most likely associated with increases in contraceptive use in the high-income world. Most strikingly, in Eastern Europe, which had the highest abortion rates in 1995, abortion rates declined dramatically—from 90 per 1,000 women in 1995 to 44 per 1,000 women in 2003 and to 42 per 1,000 women in 2010—coincident with a substantial increase in contraceptive use. Western Europe continues to have the lowest abortion rate, which has not changed since 1995 (12 per 1,000 women) (Sedgh et al., 2012).

Even where abortion is legal, access is very limited, there are poor systems of referral, and very often services are of poor quality. For example, despite the fact that manual vacuum extraction is much safer, the most frequently used method for abortion in most countries is dilatation and curettage, which has significant associated morbidity. In recent years, some encouraging signs indicate that use of manual vacuum extraction and medications to end unwanted pregnancies is increasing (Alan Guttmacher Institute, 2009a).

The World Health Organization estimates that approximately 21.2 million unsafe abortions (i.e., those not attended by a trained health professional) were performed in 2008, resulting in nearly 46,800 deaths of women in LMICs. Globally, unsafe abortion accounts for 8% of maternal deaths (Say et al., 2014). Case fatality rates (i.e., deaths per 100,000 unsafe abortion procedures) vary tremendously by region, with the world average being 220 deaths per 100,000 unsafe abortions and ranging from 4 deaths per 100,000 unsafe abortions in the high-income world to 470 deaths per 100,000 unsafe abortions in sub-Saharan Africa (WHO, 2011). It is worth noting that case fatality rates for unsafe abortions are many times higher than the rates for safe abortions, again with significant regional differences—ranging from 10 times higher in the high-income world to 1,000 times higher in sub-Saharan Africa.

Another abortion-related issue is the recent rise in the male-to-female sex ratio at birth in Southeast Asia. Data from a number of countries, including China and Korea, show that the sex ratio at birth in these areas is unusually high compared with the expected ratio of 1.06 male birth to each female birth, and is steadily increasing. Rising sex ratios at birth suggest that selective abortion of female fetuses may be increasing (Arnold, Kishor, & Roy, 2002; Larsen, Chung, & Das Gupta, 1998; Westley, 1995). This practice is embedded in the context of strong societal preferences for male children and declining family size desires, which increase the incentives to have both the desired family composition and the desired number of children (Hesketh, Lu, & Xing, 2011).

In recent decades, the broader availability of modern technology has provided the means to actualize these preferences. There are basically three ways of determining the sex of a fetus: chorionic villus sampling, amniocentesis, and ultrasound. Ultrasound is the safest and cheapest method, but works reliably only 5 to 6 months into the pregnancy (i.e., the end of the second trimester). This technology is widely available in rural areas in India, China, and other parts of East Asia. Despite strong legal sanctions against sex-selective abortion, the availability of relatively cheap ultrasound technology has promoted this practice, leading to deleterious consequences for women who undergo late-term, riskier abortions and possibly for children due to the increasing sex imbalance in the population.

Consequences of Unintended Births for Infant Health

Data on the consequences of unintended births for infant health are available for the period before 1990. They show that these births were concentrated in demographically high-risk categories (**TABLES 5-5**–**5-7**); that is, the proportion of unintended births was much higher among young mothers and older mothers, among higher-parity mothers, and following a short birth interval (National Research Council, 1989). Why these demographic characteristics are associated with higher risks for infant health is discussed in greater detail in the next section. Notably, unwanted births have higher mortality even when the pregnancy fits into an otherwise demographically low-risk category (e.g., mothers age 24 to 29; parity of 2 to 4).

Human Capital Investments

In addition to deleterious health consequences for the mother and the index child, unintended births have spillover and long-term cumulative effects by reducing human capital investments (i.e., allocation of

TABLE 5-6 Percentage of Unintended Most Recent Birth or Current Pregnancy by Birth Order in Selected Countries

	Birth Order		
Country (year)	**1**	**2–4**	**5+**
Bolivia (1993)	32.7	50.1	78.6
Colombia (1986)	25.0	50.7	68.8
Egypt (1988)	3.8	39.5	67.2
Kenya (1993)	52.1	52.1	66.2
Nigeria (1990)	11.2	9.9	22.7
Philippines (1993)	22.3	47.4	63.6
Tanzania (1991)	18.7	25.0	39.5
Thailand (1987)	20.7	36.3	64.4

Reproduced with permission from Tsui, A. O., Wasserheit, J. N., & Haaga, J. G. (Eds.). (1997). *Reproductive health in developing countries: Expanding dimensions, building solutions*. Washington, DC: National Academy of Sciences. Copyright 1997 by the National Academy of Sciences.

TABLE 5-5 Percentage of Unintended Most Recent Birth or Current Pregnancy by Mother's Age in Selected Countries

	Mother's Age		
Country (year)	**Younger Than 20 Years**	**20–34 Years**	**35 Years or Older**
Bolivia (1993)	41.5	53.9	74.0
Colombia (1986)	34.9	48.3	60.5
Egypt (1988)	13.9	41.3	75.0
Kenya (1993)	61.2	55.3	65.4
Nigeria (1990)	13.7	12.5	21.6
Philippines (1993)	37.6	46.4	58.2
Tanzania (1991)	22.7	26.1	31.7
Thailand (1987)	28.3	38.4	47.2

Reproduced with permission from Tsui, A. O., Wasserheit, J. N., & Haaga, J. G. (Eds.). (1997). *Reproductive health in developing countries: Expanding dimensions, building solutions.* Washington, DC: National Academy of Sciences. Copyright 1997 by the National Academy of Sciences.

TABLE 5-7 Percentage of Unintended Recent Higher-Order Birth or Current Pregnancy by Interval from Previous Birth to Conception: Selected Countries

	Interval from Previous Birth to Conception	
Country (year)	**Birth Interval < 24 Months**	**Birth Interval ≥ 24 Months**
Bolivia (1993)	68.4	60.1
Colombia (1986)	67.6	49.2
Egypt (1988)	54.0	50.2
Kenya (1993)	67.2	56.1
Nigeria (1990)	21.1	13.0
Philippines (1993)	62.1	49.3
Tanzania (1991)	33.5	27.5
Thailand (1987)	53.4	36.5

Reproduced with permission from Tsui, A. O., Wasserheit, J. N., & Haaga, J. G. (Eds.). (1997). *Reproductive health in developing countries: Expanding dimensions, building solutions.* Washington, DC: National Academy of Sciences. Copyright 1997 by the National Academy of Sciences.

resources for education and health) in the family as a whole. A number of studies in LMICs have shown that older children, especially girls, suffer disproportionately in terms of lower educational attainment and health status as family size increases—the latter being a proxy for unwanted births (Bledsoe & Cohen, 1993; Desai, 1995; Frenzen & Hogan, 1982; Lloyd, 1994).

Cross-sectional associations between large family sizes and lower health and educational attainments need to be interpreted cautiously, given that parents who choose to have large families may also choose to invest differently in different children (Knodel, Chamratrithirong, & Debavalya, 1987). Nevertheless, some evidence, from a family planning quasi-experiment in which villages were assigned different levels of family planning services, suggests that part of the relationship between large family sizes and low human capital investments is causal (Foster & Roy, 1996).

Unmet Need for Contraception

A primary objective of family planning programs has been to reduce unintended births by addressing "the unmet need for contraception." This unmet need is conventionally estimated from representative population-based surveys of currently married women as the sum of the number of currently pregnant women who report that their pregnancy is unintended and the number of currently nonpregnant women who are not using contraception and would not like to have any more children or, at least, none in the next two years (Bankole & Westoff, 1995). On the basis of this definition, the unmet need for contraception (**TABLE 5-8**)

TABLE 5-8 Unmet Need for Contraception and Demand for Family Planning for Married Women in 37 Countries, Based on DHS Survey in 2010–2015

Country	Unmet Need	Current Use of Contraception	Demand for Family Planning	Percentage of Total Demand Satisfied
Sub-Saharan Africa				
Benin, 2011–2012	32.6	12.9	45.5	28.4
Cameroon, 2011	23.5	23.4	46.9	49.9
Chad, 2014–2015	22.9	5.7	28.6	19.8
Congo, 2011–2012	18.4	44.7	63.1	70.9
Democratic Republic of the Congo, 2013–2014	27.7	20.4	48.1	42.5
Ethiopia, 2011	26.3	28.6	54.9	52.1
Ghana, 2014	29.9	26.7	56.6	47.2
Guinea, 2012	23.7	5.6	29.3	19.1
Lesotho, 2014	18.4	60.2	78.6	76.5
Liberia, 2013	31.1	20.2	51.3	39.4
Malawi, 2010	26.2	46.1	72.3	63.8
Mali, 2012–2013	26.0	10.3	36.3	28.5
Namibia, 2013	17.5	56.1	73.6	76.2

(continues)

TABLE 5-8 Unmet Need for Contraception and Demand for Family Planning for Married Women in 37 Countries, Based on DHS Survey in 2010–2015 *(continued)*

Country	Unmet Need	Current Use of Contraception	Demand for Family Planning	Percentage of Total Demand Satisfied
Sub-Saharan Africa				
Niger, 2012	16.0	13.9	29.9	46.5
Nigeria, 2013	16.1	15.1	31.2	48.5
Rwanda, 2014–2015	18.9	53.2	72.2	73.8
Senegal, 2014	25.6	22.2	47.7	46.4
Swaziland, 2006–2007	24.7	50.6	75.4	67.2
Tanzania, 2010	22.3	34.4	56.7	60.6
Uganda, 2011	34.3	30.0	64.3	46.7
Zambia, 2013–2014	21.1	49.0	70.2	69.9
Zimbabwe, 2010–2011	14.6	58.5	73.1	80.1
North Africa/West Asia/ Europe				
Armenia, 2010	13.5	54.9	68.4	80.3
Egypt, 2014	12.6	58.5	71.1	82.3
Jordan, 2010	11.7	61.2	72.9	83.9
Asia				
Bangladesh, 2014	12.0	62.4	74.4	83.9
Cambodia, 2014	12.5	56.3	68.8	81.9
Indonesia, 2012	11.4	61.9	73.2	84.5
Pakistan, 2012–2013	20.1	35.4	55.5	63.8
Latin America and Caribbean				
Colombia, 2010	8.0	79.1	87.1	90.8
Dominican Republic, 2013	10.8	71.9	82.7	86.9
Haiti, 2012	35.3	34.5	69.8	49.4
Honduras, 2011–2012	10.7	73.2	83.9	87.3

Data from Macro International Inc., 2010–2015, MEASURE DHS STAT compiler. Accessed on November 10, 2016 from http://www.measuredhs.com

is estimated to be 20% or less for married women in countries with high contraceptive prevalence rates (e.g., 55% or higher) and ranges up to 30% (Uganda) of women in countries with lower contraceptive prevalence rates.

The definition of "unmet need for contraception" has been criticized as an underestimate of actual need because it excludes both currently married women who are not pregnant and who are using inappropriate (because of health consequences or side effects) methods of contraception and sexually active women who are not currently married and who do not wish to become pregnant, at least in the next two years (Bongaarts, 1991; Dixon-Mueller & Germain, 1992; Pritchett, 1994a, 1994b). An additional criticism revolves around the issue of ineffective contraception. Large numbers of women currently use traditional methods, which have much higher failure rates than available modern contraceptives. Current estimates of unmet need do not include women who are using traditional (i.e., ineffective) contraceptive methods.

More recent estimates (2007) consider unmarried as well as married women. They show that for all LMICs, approximately 112.4 million women have an unmet need for contraception (108 million married and 4.2 million unmarried). These numbers translate into 14% of married women versus 7.7% of unmarried women having an unmet need for contraception. Women age 15 to 24 have significantly higher unmet need than 25- to 49-year-old women (Sedgh, Hussain, Bankole, & Singh, 2007). More than half of the total unmet need for contraception is attributed to the need to space births, and the other half to the need to limit births (Ross & Winfrey, 2002). The total number of women whose needs are unmet results from the combination of upward pressure due to population growth and downward pressure due to the success of family planning programs (Ross & Winfrey, 2002).

Family planning programs have played an important role in reducing the unmet need for contraception by making contraceptive options both physically accessible and financially affordable. Since the late 1950s, when the first national family planning programs in LMICs were established, there has been a significant increase in the prevalence of contraceptive use. This increase has played an important role in the significant reduction in fertility that has taken place especially over the last three decades in these regions, where the average number of births per couple declined from more than 6 to fewer than 3 in the latter half of the twentieth century.

Program success in improving contraceptive prevalence rates has, however, been somewhat uneven. It has depended on a number of factors, including a receptive social and family environment that accepts fertility control as legitimate behavior, a favorable political and bureaucratic climate, a management structure that pays close attention to both quality and quantity of services, and reliable sources of funding. Furthermore, those programs that have succeeded have invested considerable resources in evaluation, research, and monitoring of their services, and have had the flexibility to adapt to local conditions (Bongaarts, 1997; Bongaarts & Watkins, 1996; Bulatao, 1993, 1998; Freedman, 1987).

A long-standing debate has focused on the relative merits of "demand-side" versus "supply-side" interventions to reduce the unmet need for contraception (Bongaarts, 1997; Pritchett, 1994a, 1994b). Demand-side proponents argue that improvements in women's socioeconomic status are an essential and necessary prerequisite to the success of family planning programs. Thus, educated women with higher status, compared to their less-educated and lower-status peers, are more likely to know about contraception and to seek it out to actualize their latent fertility desires. Supply-side proponents, in contrast, posit that family planning programs, when properly managed, can increase access to and availability of contraception, even in the absence of changes in the socioeconomic status of women. Thus, they can lead to increased contraceptive prevalence rates and initiation of fertility decline.

The experience of many countries shows that the onset of fertility decline[1] is not dependent on any particular threshold in socioeconomic factors such as levels of urbanization, female education, or infant mortality. In fact, fertility decline appears to have started in a wide range of LMICs at quite varied levels of socioeconomic status. Bangladesh is frequently cited as the best example of improved contraceptive prevalence rates and dramatic fertility decline in the absence of socioeconomic improvements but in the presence of a well-run, sharply focused family planning program (Cleland, Phillips, Amin, & Kamal, 1994), although the absence of socioeconomic change has recently come under question (Caldwell, Barkat, Caldwell, Pieris, & Caldwell, 1999; Menken, Khan, & Williams, 1999). While there appears to be no magic

1 The onset of fertility decline is usually dated from an initial decline of at least 0.7 point in total fertility over a five-year period, following the practice employed by Bulatao and Elwan (1985).

threshold of socioeconomic development for initiation of fertility decline, the decline occurs more rapidly in countries with greater levels of socioeconomic development (Bongaarts & Watkins, 1996).

The demand-side versus supply-side debate is basically a false dichotomy. Neither development nor family planning programs are a necessary prerequisite, nor is either sufficient to induce fertility decline on its own (Ross & Mauldin, 1996). Rather, these factors work in a complementary fashion, with the time scales of their respective impacts being very different. On the one hand, investments in improving women's status and educational attainment certainly have an important impact in reducing unmet need, but it is a long-term impact. On the other hand, family planning programs can increase access to contraception in the short run, thereby enhancing knowledge about its use and availability and addressing many of the negative myths about particular methods of contraception. Appropriately crafted and focused media campaigns, when implemented as part of family planning programs, can also help legitimize contraception as an acceptable and desirable form of behavior. Moreover, it is important to note that access to the means to limit fertility in and of itself helps improve the status of women. Family planning programs work synergistically with improvements in socioeconomic status and are most effective when they are directed at an informed, educated, empowered client base (Freedman, 1987).

In summary, Bongaarts (1997) estimates that approximately 40% of the fertility decline in the last three decades of the twentieth century in LMICs (from a TFR of 6 to 3 births per woman) can be attributed to family planning programs, and approximately 60% to changes in socioeconomic status, particularly for women.

The Challenges Facing Family Planning

Despite significant family planning program success in reducing financial and logistic constraints to contraceptive access, the unmet need for contraception remains high in many countries. Studies by Bongaarts and Bruce (1995) and Casterline and associates (1996) reported that the major barriers to use of contraception appear to be lack of knowledge about contraception availability and use, concerns about the deleterious health consequences of contraception, and opposition from family and community to contraception use.

Given that physical access and financial constraints are not considered to be significant barriers to the use of contraception, the major challenges for family planning revolve around improving the quality of services, particularly in the areas of information exchange and method choice; integration with reproductive health services other than contraception; and last, but not least, financial sustainability.

Information Exchange

In upgrading the quality of family planning services, the major area of concern is information exchange between providers and clients. The fragmentary evidence that exists suggests that inadequate information is often provided about the proper use of contraceptives, alternatives in the event of non-optimal use, contraceptive side effects, and the appropriateness of the chosen method for women who have particular health problems (Winikoff, Elias, & Beattie, 1994). For example, quite a few women who are using oral birth control pills do not know that they can make up for a missed day by taking two pills the next day. Similarly, not enough women know that birth control pills should not be used if a woman is a smoker or has a heart condition (Trottier, Potter, Taylor, & Glover, 1994).

In general, far fewer than 50% of women have meaningful knowledge about contraceptive methods (Bongaarts & Bruce, 1995). This lack of specific knowledge often leads to exaggerated notions of the health risks of contraception (Casterline et al., 1996). It is worth reiterating that contraception is, by and large, very safe, especially when compared to the health risks deriving from an unplanned pregnancy. Ross and Frankenberg (1993) have estimated that the mortality risk of an unplanned, unwanted pregnancy is 20 times the risk of any modern contraceptive method and 10 times the risk of a properly performed abortion (Ross & Frankenberg, 1993). Although the last two decades have witnessed great success in social marketing, whereby most women have become aware of the benefits of small families and the existence and availability of contraception, much more needs to be done to educate women about method choice and associated health risks and benefits.

Although concern about information exchange in family planning programs is long-standing, progress in addressing this problem has been uneven. An issue that comes up repeatedly is whether there is a quality–quantity tradeoff. Program managers voice a common complaint that they have their hands full just providing physical access to contraceptives. Many feel they do not have the luxury of providing extensive information about contraception because of the time-intensive nature of this type of activity.

In reality, there is little contradiction or tradeoff between paying attention to quality issues and achieving quantity targets for numbers of users. The two are integrally linked in several ways. First, the key steps needed to improve quality—such as attention to logistics, adequate supervision, motivation of workers at every level, real feedback to managers and supervisors, and accountability for supplies and money—are exactly the same steps needed to improve quantity. Second, family planning services are inseparable from information provision. In fact, provision of information is one of the key services that a family planning program can offer. Third, attention to quality will improve efficiency and will allow the addition of new users to family planning services without new costs (Tsui et al., 1997).

Contraceptive Use and Method Choice

Modern contraceptive methods are now so widely available in LMICs that approximately 46% of all couples outside of China and approximately 62% of all couples in high-income countries in which the woman is of reproductive age use such methods (Population Reference Bureau, 2016). In 2003, the estimated number of contraceptive users in LMICs was 506 million. This number had grown to 600 million by 2008 and to 645 million by 2012 (Darroch & Singh, 2013). By 2030, the number of global users of contraceptives is expected to be approximately 800 million (United Nations Population Division, 2015a). This rapid growth rate is a result of population increases and a concomitant rise in contraceptive prevalence in LMICs, from 61% in 2000 to 64% in 2015, due to declines in desired fertility.

TABLE 5-9 shows the change in worldwide distribution of contraceptive use according to method from 2002 to 2015. Note that the global contraceptive use rate increased by only 3% between 2000 and 2015. The increase is almost uniform across the regions except Africa, where a 7% increase has been experienced (from 26% to 33%); this rate remains the lowest on a regional basis in the world. These findings support the notion that once the contraceptive use rate reaches 60% in any country, there is little scope for further increase.

In addition to demonstrating the great variation in use, these data are remarkable in that they show, for much of the world, a high proportion of those couples using contraception are using sterilization—the one method that has almost zero risk of failure. Note that the average rate of contraception use for LMICs in Asia (67.8%) is heavily influenced by the very high rate of contraceptive prevalence in China (83.4%). Without China, the average for all LMICs would drop from 61% to 54% (Population Reference Bureau, 2016). More recent data show that there is a consistent decline in permanent method use across the regions with increase in condom and pill use.

Another modern contraceptive that has not yet reached the mainstream is emergency contraception. A recent study in LMICs shows that nearly 20% of women age 15 to 49 years had heard about this method and approximately 2% of sexually active women had ever used it (United Nations, 1994), although wide variations exist among countries (Palermo, Bleck, & Westley, 2014). Another important observation is that there is no strong correlation between TFR and contraceptive use rate in many high-income countries. For example, Japan has one of the lowest TFR in the world (1.4 births per woman), but a modest contraceptive use rate of 56%, of which 80% involves use of condoms. Japanese couples rely on other fertility control measures that reduce exposure to risk of pregnancy, perhaps including more effective use of controlled abstinence (Moriki, Hayashi, & Matsukura, 2015); in Japan, the average age at marriage is 26.3 years for women and the average age at first birth is 27.5 years (Anonymous, 1996).

It is much more difficult to judge how much *choice* of method women in LMICs have, and how their array of choices has changed in recent years. With the exception of sub-Saharan Africa, where contraceptive prevalence rates are low, significant progress in overall method availability occurred in most countries between 1982 and 1994 (Ross & Mauldin, 1996). In all 22 countries (with the exception of Nigeria) where DHS surveys were conducted between 1990 and 1993, at least half of all women had heard of at least one modern contraceptive method; in 13 of these countries, more than 90% of women knew of at least one contraceptive method (Curtis & Neitzel, 1996).

TABLE 5-10 shows historical data on contraceptive method mix and alternative projections for contraceptive choices in 2015 in LMICs (Bongaarts & Johansson, 2002). Although the two projection methodologies resulted in varying estimates for specific methods, female sterilization appeared likely to remain the most popular method of choice in 2015, followed by intrauterine devices (IUDs) and pills. When the estimated method mix was compared with actual data collected between 1998 and 2012 from LMICs, the earlier projections were found to have largely overestimated the use of sterilization and IUDs (Ross et al., 2015). Conversely, injectable methods

TABLE 5-9 Contraception Use Pattern in 2000 and 2015: Percentage of Married Couples[a] in Which Women Are of Reproductive Age, by Region

	Sterilization													
	Female		Male		Pill		Intrauterine Device		Injectables		Condom		Total	
Region	2000	2015	2000	2015	2000	2015	2000	2015	2000	2015	2000	2015	2000	2015
World	21.0	19.2	4.0	2.4	7.0	8.8	15.0	13.7	—	4.6	5.0	7.7	61.0	63.6
Low- and Middle-Income Regions														
Africa	2.0	1.6	0.1	0.0	7.0	8.7	5.0	3.8	—	9.8	1.0	2.1	26.0	33.4
Asia	25.0	23.7	4.0	2.2	5.0	6.4	18.0	17.4	—	3.9	4.0	7.6	64.0	67.8
Latin America and Caribbean	31.0	25.7	2.0	2.6	13.0	15.0	8.0	6.4	—	6.8	4.0	9.6	70.0	72.7
Oceania	12.0	8.0	9.0	6.3	21.0	21.6	2.0	1.1	—	1.9	9.0	10.2	59.0	59.4
High-Income Regions														
Japan	3.0	1.7	0.6	0.5	0.8	1.1	1.5	1.0	—	0.0	43.0	46.1	56.0	56.5
Europe	4.0	3.7	2.0	3.3	16.0	21.9	15.0	11.3	—	0.4	10.0	16.5	67.0	69.2
North America	23.0	20.6	14.0	11.9	15.0	16.5	1.0	4.7	—	0.1	13.0	11.9	76.0	74.8
Australia	14.0	6.3	19.3	8.9	20.5	28.8	3.3	1.5	—	2.0	11.0	14.0	74.0	68.4

[a] Including, where possible, those in consensual unions.

Data from Population Reference Bureau, Family Planning Worldwide. (2002). *2002 Data Sheet*. Washington, DC: The Population Reference Bureau, Adapted with permission for 2002 data; United Nations, Department of Economic and Social Affairs, Population Division. (2015). *Trends in contraceptive use worldwide 2015 (ST/ESA/SER.A/349)*. Retrieved from http://www.un.org/en/development/desa/population/publications/pdf/family/trendsContraceptiveUse2015Report.pdf

TABLE 5-10 Estimates of Method Distribution in 1980 and 1993 and Alternative Projections for 2015, Low- and Middle-Income Countries

	Actual	Actual	Futures Group Projection	New Procedure Projection	Estimated
Method	**1980**	**1993**	**2015**	**2015**	**1998–2012***
Female sterilization	24	39	26	37	12.8
Male sterilization	13	8	3	5	1.3
Pill	13	11	22	17	22.3
Injectable/Implants	0	4	6	5	18.0
Intrauterine device	32	26	18	20	12.9
Vaginal methods	0	0.3	0.6	0.4	—
Condom	5	4	10	9	10.7
Traditional methods	12	9	14	7	22.1
Total	100	100	100	100	100

The figures for 1993 are proportions of total contraceptive use. An estimated 21% of married couples used female sterilization. As the total number of married couples using contraceptive methods was 55%, this translates into (21/55 = 39%) of total contraceptive use.
Modified from Population Council, Bongaarts, J., & Johannsson, E. (2002). Future trends in contraceptive prevalence and method mix in the developing world. *Studies in Family Planning, 33*(1), pp. 24–36.
*Estimates for 1998–2012 are taken from Ross, J., Keesbury, J., & Hardee, K. (2015). Trends in the contraceptive method mix in low- and middle-income countries: Analysis using a new "average deviation" measure. *Global Health: Science and Practice*, *3*(1), 34–55.

were used by a significant share of the population (18.0%), and—perhaps surprisingly—use of traditional methods also increased more than projected. An issue of considerable concern is that despite evidence of a spreading HIV epidemic in LMICs, condom use remains low in all estimates.

Largely due to a lack of funding and the long lead time required for new methods to gain acceptance, there has been relatively little innovation in contraceptive technology in the last 30 years. Thus, in an era of increasing expectations for contraception, the menu of choices has not expanded greatly. Given the AIDS pandemic, new contraceptive methods that are of high priority are vaginal microbicides, which protect women from STIs, in combination with spermicides, which provide contraceptive protection (Bongaarts & Johansson, 2002; Harrison & Rosenfield, 1996; Tsui et al., 1997).

In some situations, knowledge and use of existing technology have not been widely disseminated. One example is emergency contraception—that is, the prevention of pregnancy through the use of contraceptive methods after unprotected sex—for which appropriate technologies (e.g., a combination of oral contraceptive pills, progestin-only pills, and the copper-T IUD) have long been available but are not used by many women (e.g., victims of coercive sex) who could benefit from them (International Planned Parenthood Foundation, 1995; Trussell, Ellertson, & Stewart, 1996). Although this treatment was discovered in 1966, it was not until 1995 that the International Consortium for Emergency Contraception (ICEC, 2010) was formed to promote and mainstream emergency contraception worldwide. However, progestin-only preparations, commonly known as the "morning-after pill," are now available commercially in most countries as a dedicated emergency contraception product under many names worldwide (Trusell & Cleland, 2010).

Political, Social, and Financial Constraints

Although high unmet need is in part a function of specific management deficiencies in family planning programs, it is important to recognize that broader political and societal constraints also play a role.

Kenney (1993) has identified policies in a variety of countries that retard access to safe contraception:

- Health and safety regulations that restrict choice of methods or providers (e.g., the failure to approve oral contraceptives for use in Japan for more than nine years; a committee finally recommended their approval after the male impotence–relieving drug Viagra received endorsement within six months (Goldberg, 1999)
- Taxes and barriers to trade that affect importation of contraceptives
- Regulation of advertising (usually due to concerns about modesty and privacy)
- Restrictions on private-sector involvement in family planning

In addition, law and policies in many countries restrict or forbid access to abortion (Kenney, 1993).

In any discussion of family planning programs and their performance, the issue of financial sustainability is key. Family planning expenditures in LMICs as a whole are estimated to be approximately $14 billion annually, or roughly $1–2 per person per year in 2014 (High-Impact Practices in Family Planning [HIP], 2014). In the last two decades, most of this expense has been paid for by national governments (50%) and individual households (20%), with international donor assistance accounting for only 30% of the total ($4.2 billion). In the coming decades, rising demand for contraception and increasing budget constraints will require that programs either mobilize more public resources or increase the cost of family planning services to the individual, so that more users can be accommodated (Bulatao, 1998).

The Broader Effects of Family Planning Programs

Even before the 1994 International Conference on Population and Development in Cairo, concerns were expressed, both by women's groups and by policy makers, about the effects of family planning programs on the lives of individual women. Family planning programs have been criticized as exclusively concerned with reducing population growth. Studies in recent years have broadened research to include consideration of the effects of programs on the quality of life for women. In particular, the Women's Studies Project found that family planning programs provided the following benefits to women:

- "Most women and men are convinced that practicing family planning and having smaller families provide health and economic benefits."
- Family planning offers freedom from fear of unplanned pregnancy and can improve sexual life, partner relations, and family well-being.
- Where jobs are available, family planning users are more likely than non-users to take advantage of work opportunities.
- "Family planning helps women meet their practical needs and is necessary, but not sufficient, to help them meet their strategic needs" (Women's Studies Project, 1999).

The Women's Studies Project also found costs to women from such programs:

- "Contraceptive side effects—real or perceived—are a serious concern for many women, more so than providers realize.
- When partners or others are opposed, practicing family planning can increase women's vulnerability.
- When women have smaller families, they may lose the security of traditional roles and face new and sometimes difficult challenges, including the burden of multiple responsibilities at home and work" (Women's Studies Project, 1999).

In addition, the Women's Studies Project found that the exclusion of men from most family planning programs affected the ability of women to take advantage of their services, because men play a dominant role in family planning decisions in many regions.

Most family planning programs have emphasized only the positive benefits of family planning; they are now being urged to pay attention to at least some of these broader considerations.

A Broader Definition of Family Planning and Reproductive Health Programs

As discussed earlier, the 1994 ICPD brought about an international reevaluation of the conceptualization of family planning programs. These programs are now viewed as falling under the larger rubric of more general reproductive health services and interventions, some of which are directly health related, whereas others are related indirectly. It is useful, however, to first consider conventional family planning programs.

Organization and Structure of Family Planning Programs

While nearly all LMICs have established an infrastructure to deliver family planning services, their organization varies markedly. One example, Bangladesh, is discussed in **EXHIBIT 5-4**.

EXHIBIT 5-4 The Bangladesh Family Planning Program

Family planning in Bangladesh can be traced to the private Family Planning Association created in 1953, before Bangladesh achieved its independence from Pakistan. By 1960, Pakistan had begun public-sector programs, which Bangladesh continued after becoming a nation in 1971. The overall program has grown and changed over the years, but throughout there has been high-level political support and considerable funding from external donors. The Family Planning Association program has emphasized provision of services, outreach activities at the village level, and mass communication through a variety of media.

Within the early 1960s public-sector programs, family planning services were offered in government health clinics as part of regular health services. A system of using village aides to provide education was established but abandoned after only 18 months for a variety of reasons, including poor training of the aides, complaints that their services were directed only to family planning and not to other health problems, inadequate resources, and poor supervision. It was followed by renewed efforts run by a new Family Planning Board independent of the Ministry of Health.

Bangladesh's family planning efforts in the late 1960s met with little success, primarily because of poor-quality services provided by a program that had been instituted on a large scale, with little pilot testing and poor organization. The program emphasized use of the intrauterine device (IUD), which was met with resistance by many concerned about side effects and problems with its use. It did not help that the program was seen by many as having been imposed on Bangladesh, then East Pakistan, by a government whose political support was declining.

In the aftermath of the country's war for independence, although the health and social sectors of the government were particularly negatively affected, it was felt that family planning was urgently needed. A large and complex program was established. A separate Population Wing was created within the Ministry of Health and Population to run the program. Thus, health and family planning services were separated. At the local level, the primary healthcare staffs were predominantly male. In a society where little interaction is permitted between women and men who are not members of their families, male workers cannot provide maternal and child health services except for immunizations. This staffing was a legacy of early programs to combat smallpox, tuberculosis, and malaria; it was ill suited to the new focus. By contrast, local family planning workers were women, although their supervisors were men. These female workers went directly to households and offered family planning counseling and free supplies. They spent some time focusing on maternal and child health, although they were not well trained for this purpose.

Over the years, the Bangladeshi Family Planning program has continued to be revised, restructured, and expanded. In all cases, the elements of strong political support, strong financial support, and extensive administrative support have remained (Cleland et al., 1994; Streatfield & Kamal, 2013).

Cleland, J., Phillips, J. F., Amin, S., & Kamal, G. M. (1994). *The determinants of reproductive change in Bangladesh*. Washington, DC: World Bank.

According to Tsui, Wasserheit, and Haaga (1997), successful performance in the family planning realm is influenced by a focused commitment to achieving program objectives and access to adequate resources. At the national level, strong leadership, clearly formulated policies, explicit goals and objectives, and a clear agenda for meeting those goals can all contribute to the success of programs. In some countries, political commitment is evidenced by placing the family planning program under a national supervisory council or by establishing a separate ministry.

Programs also need ways of assessing progress toward meeting their objectives. Indicators such as contraceptive prevalence, proportion of unwanted births, maternal morbidity and mortality, pregnancy complications and their management, and actual fertility levels all provide information that, over time, can permit program evaluation. Therefore, one element of successful programs is the definition of result measures to be used and establishment of mechanisms for collecting the needed information. Caution is in order, however. For example, goals that are defined in terms of targets, such as the number of acceptors of particular methods in a given time, may lead workers to exert pressure on clients and reduce their options.

Family planning service programs have tended to focus on a narrow set of goals—reducing unwanted fertility by providing access to the means of fertility control. Several models have been used for the design of programs. In the vertical model, family planning administration and service delivery are carried out by staff for whom this is their single function. In a second model, a separate family planning administration unit is established, but field staff at each level of the healthcare system can deliver a variety of linked services. In practice, the linkage between family planning and maternal and child health services has been the most common. Under the new broader definition of reproductive health, it is expected that other types of services will be offered, so that the ways in which they

are linked, both administratively and in provision of care, will have to be addressed.

No matter which model is followed, the program design involves decisions about which services will be offered and at which level of the healthcare system. Tsui, Wasserheit, and Haaga (1997) have illustrated the various possibilities.[2] The levels of healthcare system that they suggest can deliver interventions for prevention and management of unintended pregnancies are summarized here:

Community

- Information, education, and communication programs
- Community-based distribution
- Social marketing of condoms and oral pills

Health Post

- Counseling/screening for contraception
- Counseling and referral for menstrual regulation or abortion
- Provision of injectable contraceptives
- IUD insertions
- Counseling and treatment of contraceptive side effects

Health Center

- Menstrual regulation/manual vacuum aspiration abortion
- Performing surgical contraception on set days
- Post-abortion counseling and contraception
- Counseling and treatment of contraceptive side effects

District Hospital

- Surgical contraception
- Abortions through 20 weeks, where indicated
- Post-abortion counseling and contraception

These authors' report concludes that the breadth and scope of the services to be delivered present formidable challenges in terms of design, execution, administration, and evaluation. Even if these challenges are met, a program can falter if adequate resources are not allocated to meet its needs for trained staff, equipment, and supplies. Additional demands will be placed on whatever system is in place if services related more generally to reproductive health are provided.

Research on design and implementation can help improve family planning and reproductive health programs. **EXHIBITS 5-5** and **5-6** describe the approach taken by Bangladesh in this regard.

Additional Reproductive Health Care Services

Some reproductive health services are closely linked to contraception; it is likely that, without much added expense, they can be integrated into conventional family planning programs relatively easily. These services may include pregnancy tests, Pap smears, and screening for sexually transmitted infections. STI screening has been carried out, in many cases, only in the context of separate programs to treat such infections. That type of intervention misses the general population of women who may not realize that they are infected or understand that they may pass their infection to their unborn or nursing children.

HIV/AIDS deserves special mention. In 1999, the United Nations convened a conference to assess the progress made on this front since the 1994 ICPD. Attendees concluded that the earlier conference had greatly underestimated the effects of HIV/AIDS on the populations of LMICs and called for specific programs and targets to reduce the spread of infection ("Conference Adopts Plan on Limiting Population," 1999). Directly related to the family planning realm is their call for greater access to methods such as female and male condoms that can reduce or prevent transmission of the virus. This and other approaches to preventing the spread of HIV/AIDS are discussed in the *Infectious Diseases* chapter.

Other desirable reproductive health interventions remain within the health realm but will require significant changes in staffing and significantly more financial resources to implement. These services include emergency obstetrics, general women's health services, abortion services where they are not already available, infertility services, and greatly expanded testing and counseling for HIV/AIDS. Some infertility services are already provided within the context of programs to reduce STIs, as these diseases are a major cause of infertility and premature sterility (Tsui et al., 1997). The others, however, are generally lacking. Without considerable expansion of the financial base for family planning and for the expanded reproductive health program, it is unlikely that these services can be provided in many LMICs.

2 Reprinted with permission from Tsui, A. O., Wasserheit, J. N., & Haaga, J. G. (Eds.). (1997). *Reproductive health in developing countries: Expanding dimensions, building solutions. Panel on Reproductive Health, Committee on Population, Commission on Behavioral and Social Sciences, and Education*. Washington, DC: National Academy Press. Copyright 1997 by the National Research Council.

EXHIBIT 5-5 Research to Improve the Family Planning Program in Bangladesh

How well do various types of family planning programs work? Few experiments in applied research have been conducted to determine whether a particular design is more effective in reaching the objectives of the program. The International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B) has carried out just these kinds of operations research experiments, which have served to improve the ways in which services are delivered within the Bangladeshi family planning program.

An early effort, in 1975, was intended to test the hypothesis that there was latent demand for family planning. ICDDR,B maintains a field station in Matlab, approximately 40 kilometers from the capital, Dhaka. A family planning program was introduced in roughly half of the area in which the center provided services, while people living in the remainder of the area had access only to standard government or private services. Local women, mostly illiterate widows, were hired to visit households approximately every 90 days to offer oral contraception to women. Later, condoms were added to the offerings. The hypothesis was that couples wanted to reduce their fertility, but would do so if only they were supplied with the necessary means. Initial acceptance was good; prevalence of use rose in the early stage of the program to almost 20% from its near-zero level prior to the program. Nevertheless, within less than nine months, it had dropped, so that the program area prevalence was only 6 percentage points higher than in the comparison area.

Lessons learned: This type of demand-oriented program is inadequate in a situation where women and couples have little social support for contraceptive use. Rather, a system that addresses the non-economic costs of use—whether social, psychological, or based on health concerns—is essential. In addition, problems arose in complying with a daily pill regimen, and condoms were not popular. Analysis of the experiment through interviews with people in the community also demonstrated the importance of the characteristics of the family planning worker. Women who had little status in the community and who were older than reproductive age themselves did not have sufficient credibility to help others withstand the social costs of use. Improving access was simply not enough.

A subsequent experiment begun in 1978 tested better follow-up for users, an expanded set of method choices, employment of better-educated and younger women, and new management strategies. These steps were undertaken to ensure that women received regular visits and that problems were addressed rapidly. A new dual leadership system was introduced, which included both technical (paramedical) and administrative supervision. The interval between visits was reduced to 14 days. Within a year, nearly one-third of women in the study area were contraceptive users, while there was little change in the comparison area. Increases in contraceptive use continued so that by 1990, nearly 60% of women in the targeted area were users, compared to only 25% in the comparison area. Clearly, taking advantage of latent demand for contraception required that the program address the social costs of contraceptive use. It also demonstrated the value of providing service in the home.

In 1983, a new experiment was begun outside of Matlab, to see if the lessons learned there could be applied within the government family planning program and without injection of major additional resources or changes in administration. This pilot project was carried out in two areas. Because of the success of this Extension Project, the government changed the national program to increase the number of female village workers, train them to provide injectable contraception within the home, and upgrade management to provide better support, both technical and supervisory, for local workers. In fact, one of the main lessons learned from these experiments was that careful supervision and support of workers were critical factors for success. Another lesson emphasized the importance of designing a program for local cultural circumstances—in this case, providing basic services in the home.

Since these experiments were conducted, other projects have tested variations of the Extension Project model to see how much it can be altered and still achieve the objectives of increased use of family planning and reduced fertility. Since the International Conference on Population and Development, the Extension Project (later renamed as the Operations Research Project) has initiated studies examining how the family planning program can be expanded to provide a wide array of reproductive and other health services under what is termed the Essential Services Package (Cleland et al., 1994).

Cleland, J., Phillips, J. F., Amin, S., & Kamal, G. M. (1994). *The determinants of reproductive change in Bangladesh*. Washington, DC: World Bank.

Reproductive Health Services Beyond Direct Health Care

Reproductive health interventions that go beyond the healthcare realm, while clearly valuable, are not linked in any obvious way to conventional family planning services. The ICPD's Cairo agenda focused on improving the status of women. It called for interventions such as income-generating activities for women and female education that improve the overall status of women as well as for sex education for youth, both male and female, to increase responsible sexual

EXHIBIT 5-6 Strategies Used by the Bangladesh Family Planning and Reproductive Health Programs

According to Cleland et al. (1994), at least four sets of strategies have been implemented:

1. Strategies to improve the coverage and quality of services:
   - Clinics, located within 5 miles of most couples, now provide free clinical contraception (IUDs, injectables, and implants) and treatment of side effects.
   - Sterilization is offered without charge at all subdistrict hospitals and is carried out by well-trained personnel.
   - Related health services for children and women are provided either in the home or in clinics.
   - Community-based distribution of low-cost nonclinical contraceptives (condoms and oral pills) is provided through pharmacies and is well publicized through various media.
2. Strategies to improve awareness and motivation:
   - Mass media are used to provide extensive relevant information; family planning and reproductive health are openly discussed in public media.
   - Focused programs (for example, with religious leaders) are carried out to build awareness and consensus.
3. Strategies to foster village-based and household services:
   - Outreach involves female workers who deliver services in the home. These services are now provided by both the government and nongovernmental organizations.
   - This strategy has been questioned in recent years by critics who say the time is past when women should be provided with services that encourage continued seclusion. They argue that the demand for family planning and reproductive health services is now so great that women will travel outside their homes to obtain these services and that this type of modernization is to be encouraged. In addition, issues of cost of maintaining the large cadre of home visitors is encouraging experimentation with less costly alternatives.
4. Strategies to foster community development and demand generation:
   - These strategies have not been carried out within the family planning program itself, but rather are directed toward improving the status of women. They include micro-credit and other programs sponsored by local organizations, such as Grameen Bank and the Bangladesh Rural Advancement Committee, and government strategies to increase education of girls. In fact, education of women has increased substantially in recent years. In many parts of Bangladesh, nearly all children—male and female—receive at least several years of primary schooling.

As a final note, the Bangladesh effort is characterized by the use of research to help determine the design of programs. Ongoing studies at the International Centre for Diarrhoeal Disease Research, Bangladesh are addressing how the new Essential Services Package (which includes services related to reproductive health, child survival, and curative care) can best be implemented within the existing fixed service provision sites, how to meet the health needs of adolescents, how to improve prevention and management of reproductive tract and sexually transmitted infections, and how to provide essential obstetric care (Cleland et al., 1994).

behavior. The 1999 follow-up conference reiterated and intensified these calls ("Conference Adopts Plan on Limiting Population," 1999).

Others have called for programs that decrease violence against women. Violence in women's intimate relationships can lead to death through homicide or through driving the woman to suicide. A less drastic outcome is loss of control by women over their sexuality and, therefore, their sexual health.

Another kind of violence against women is female genital mutilation, a practice that has been reported in more than 40 LMICs and has followed immigrants from these areas to high-income countries (Tsui et al., 1997). It is estimated that more than 200 million women have been affected globally by female genital mutilation and approximately 3 million girls are at risk for female genital cutting every year (UNICEF, 2016). Although any type of genital cutting carries a risk of infection, the implications of genital cutting for long-term reproductive health differ according to the severity of cutting, conditions of delivery, and sociodemographic factors. It is generally accepted that women who have undergone the most severe type of cutting, which includes removal of external genitalia and infibulation (stitching or narrowing of the vaginal opening), have an increased likelihood of delivery complications or obstetric morbidity (Slanger, Snow, & Okonofua, 2002). Even death may occur from excessive bleeding and infection resulting from use of unsterile equipment for performing the procedure by quacks. In some

countries, evidence suggests that a growing proportion of these procedures are being carried out by medical personnel (Yoder, Abderrahim, & Zhuzhuni, 2004).

Can These Goals Be Achieved? All parts of this new "beyond family planning" mandate are worthwhile. Nevertheless, it remains unclear how this expansion, in the absence of clearly designated additional funds to finance it, will affect the ability of family planning programs to reach their objective of promoting safe contraception (Cleland et al., 1994; Finkle & Ness, 1985). In addition, concerns have been raised that the expansion from family planning programs to reproductive health programs without additional funds will not only dilute what traditional family planning does reasonably well, but also fail to provide significant improvements in other areas (Bulatao, 1998; Mukaire, Kalikwani, Maggwa, & Kisubi, 1997; Twahir, Maggwa, & Askew, 1996). **EXHIBIT 5-7** discusses these issues in the context of Bangladesh.

In fact, the new agenda comes in an era when many high-income countries are reducing their

EXHIBIT 5-7 Challenges and Constraints for Family Planning and Fertility Reduction in Bangladesh

In the early 1990s, following the rapid decline in Bangladesh fertility from a TFR of 6.3 births per woman in 1975 to 3.4 births per woman in 1994, much optimism was voiced that fertility would continue to decline and that replacement-level fertility would be reached by 2005. However, fertility in Bangladesh plateaued at a TFR of approximately 3 children per woman from 1994 until 2005. Although this became a source of major concern, fertility once again started falling since 2005 and reached 2.3 births per woman (near replacement level) in 2015.

The initial dramatic fall in fertility took place within the context of relatively insignificant overt improvements in socioeconomic development, and was largely attributed to "supply-side initiatives"—that is, the impact of an intensive family planning program.

Various hypotheses have been put forth to explain the lack of progress in diminishing fertility in the 1995–2005 decade. Advocates of conventional family planning programs suggest that this trend may have been due to the change of focus (in the mid-1990s, following the expansion of the family planning agenda to include broad-based reproductive health services—see Exhibit 5-4) from household distribution of contraceptives by a vertical cadre of family planning workers to clinic-based family planning and reproductive health services provided by a unified integrated health and family planning service. They argue that this change (1) reduced the motivation of the family planning workers, who now had an expanded set of health-related duties that they were ill prepared to fulfill, and (2) made contraceptive provision just one of many government-sponsored healthcare activities, none of which had adequate resources committed to their fulfillment. Moreover, they question the optimistic view that family planning was so well established that women in rural Bangladesh no longer had social constraints vis-à-vis seeking family planning and reproductive health services outside the home.

Those who favor clinic-based services have pointed out that contraception prevalence rates have continued to rise in Bangladesh since 1995. They note that the change of method mix from predominantly oral contraceptives to more long-term contraceptive methods with lower failure rates required clinic-based initiatives.

Demand-side advocates argue that the exclusive focus on fine-tuning family planning programs to bring about fertility reduction is misplaced. They contend that further reductions in fertility will require broad-based initiatives to improve socioeconomic development (with a particular focus on improving women's status), which will help to reduce the benefits of additional children and increase their costs. They point to continued desired fertility of 2.3 children per woman—still higher than the replacement level, with quite significant regional variations that can be correlated with differences in women's education and employment, as well as continued preference for sons.

Due to the lack of progress in achieving further fertility declines, in 2006 the government of Bangladesh switched back from clinic-based services to household distribution. The current situation with regard to TFR is that it has declined from 3.0 births per woman in 2005 to 2.3 births per woman in 2015. The reasons behind this slow but steady decline continue to be debated.

In addition to the continuation of traditional supply-side initiatives such as household distribution of contraception, a highly successful social marketing program promoting contraceptive distribution and sales has been established all over the country. Bangladesh Demographic and Health Survey 2014 (BDHS) reports that 43.5% of the oral pill users and 62.0% of condom users relied on social marketing brands bought from commercial outlets (BDHS, 2016).

On the demand side, it is also important to note that in the 2005–2015 period, significant improvements occurred in female education, employment, and economic growth (Bangladesh has averaged 6% growth in its gross domestic product over the last decade). All of these factors can be said to increase the opportunity costs of childbearing and childrearing, leading to lower levels of desired fertility among Bangladeshi families.

The debate about demand-side versus supply-side policy initiatives as the best approach to fertility reduction continues. Most researchers and policy makers believe that both kinds of initiatives are important, but there is growing appreciation that the final stretch of fertility decline to replacement levels may be a long, hard battle all over the world.

aid contributions. International experts estimate $14 billion is needed annually to sustain coverage for existing users and to achieve the 2020 family planning goal of reaching 120 million additional contraceptive users in the poorest 69 countries (Singh & Darroch, 2012). In-country funding to purchase commodities and supplies has increased dramatically to sustain such programs in many countries (proportion of total FP expenditure)—from nothing in 2006 to 89% in 2012 in Paraguay; from 73% in 2009 to 99% in 2012 in Bangladesh; and from 5% in 2011 to 21% in 2012 in Rwanda (HIP, 2014). In contrast, funding commitments for international programs have experienced an overall decline. The largest donor—the U.S. government—has significantly curtailed its funding of family planning programs, particularly in the Donald Trump administration (BBC, 2017). Commitments from other donors including the Bill and Melinda Gates Foundation, the Canadian government, and several European governments have increased but might be inadequate to replace the shortfall due to the U.S. pull-out (Wildman, 2017). Increasing the political will and raising the funding for these new programs and for maintaining existing effective ones is perhaps the greatest challenge for reproductive health in the twenty-first century.

Impact of Reproductive Patterns on the Health of Children

Over the last several decades, an impressive body of evidence has accumulated suggesting that certain kinds of reproductive patterns are injurious to infant and child health (**TABLE 5-11**). These risk factors are usually discussed as if they were completely independent. In reality, a number of problematic issues arise with this approach to deleterious reproductive patterns. For example, many of the risk factors are integrally linked with one another and their independent

TABLE 5-11 Mechanisms by Which Reproductive Patterns Affect Child Health

Reproductive Pattern	Mechanism Through Which Child Health Is Affected
Firstborn children	First-time mothers have a higher frequency of health problems during pregnancy and childbirth; parents have less experience with child care; poorer intrauterine growth
Higher-order children	Possible cumulative effect of earlier maternal reproductive injury—"maternal depletion" syndrome—leading to poorer intrauterine growth
Large families	Competition for limited resources, with some children, possibly disproportionately girls, losing out; possible spread of infection
Children born to very young mothers	Inadequate development of maternal reproductive system and incomplete maternal growth; young mothers less likely to know about and use prenatal and delivery care or provide good child care
Children born to older mothers	Greater risk of birth trauma; greater risk of genetic abnormalities
Short interbirth intervals	Inadequate maternal recovery time (maternal depletion); competition among similar-age siblings for limited family resources; early termination of breastfeeding; low birth weight; increased exposure to infection from children of similar ages
Unwantedness	Neglect (conscious or unconscious); child born into a stressful situation
Maternal death or illness (e.g., chronic infection such as AIDS)	Early termination of breastfeeding; no maternal care; disease may be passed to child
Contraceptive use	Hormonal contraception may interrupt breastfeeding

effects are difficult to disentangle. Thus, first births and young age of mothers are separately cited as risk factors, but young mothers usually are having their first birth. Similarly, children of high parity come from large families and are likely to have older mothers, yet all three factors are often referred to individually.

Parity and Child Health

First births are known to be more dangerous for the child than are subsequent births. The excess risk relative to births of order 2–4 is limited to the first year of life, however—particularly to the neonatal period (the first 28 days of life), when the odds ratio for mortality is 1.7. There appears to be no survival disadvantage after the child reaches his or her first birthday (Hobcraft, 1987; Hobcraft, McDonald, & Rutstein, 1985).

The excess risk for firstborn children varies considerably across countries. It is not clear whether this risk reflects inadequate physiologic adjustment of first-time mothers to pregnancy (leading to lower intrauterine growth, shorter gestation, lower birth weight, a higher probability of birth trauma, higher risks of pregnancy-induced hypertension, higher prevalence of placental malaria in malaria-endemic areas, and so on) or whether it arises because of the lack of experience of first-time mothers in care seeking and care taking. The latter factor is, of course, amenable to policy prescriptions that encourage first-time mothers to seek prenatal and postnatal care (Haaga, 1989; National Research Council, 1989).

Higher-order births may suffer due to poor maternal health as a result of cumulative exposure to previous pregnancies (Hobcraft et al., 1985; Pebley & Stupp, 1987). Mothers may experience inadequate recovery of their energy stores after earlier pregnancies (maternal depletion hypothesis) or the long-term cumulative effects of prior delivery-related injuries. Thus, higher-order children (parity of 5 or greater) may be at greater risk of poor intrauterine growth, greater trauma during birth, and, more generally, poorer health than lower-parity (orders 2–4) children. Although these mechanisms are plausible, the empirical evidence supporting them is inconsistent and suggests that little additional risk can be attributed to higher-order births, once short birth intervals are taken into account (Gubhaju, 1986; Hobcraft et al., 1985; National Research Council, 1989).

In addition to physiologic deficiencies, higher-order children may suffer deleterious consequences of competition for limited family resources. In particular, they may get proportionately less food and less attention from their parents. This negative consequence of large families may not be limited solely to higher-order births. If family resources are limited and there is no preference for specific children, all children may suffer as a result of large family sizes. Some evidence suggests that certain children (particularly higher-order girls, and especially those with older sisters) suffer disproportionately from the impact of large family size in specific social settings (Muhuri & Menken, 1997).

Maternal Age

Hobcraft (1987) found that children born to teenage mothers had significantly higher risks of dying than children born to mothers age 25 to 34 years (Hobcraft et al., 1985). This excess mortality risk was 1.2 for the neonatal period, 1.4 for the post-neonatal period, 1.6 for toddlers, and 1.3 for children age 2 to 5 years. Nevertheless, considerable variability was observed among countries in terms of the excess mortality risk for children born to young mothers.

Plausible explanations reflecting both physiological and social causes have been offered for the health disadvantage experienced by children born to young mothers. Perhaps these children are disadvantaged because maternal reproductive systems are inadequately developed (Aitken & Walls, 1986) or because young mothers lack experience and knowledge about prenatal and postnatal care (Geronimus, 1987). Unfortunately, little solid empirical evidence from LMICs is available to prove (or disprove) any of these hypotheses. It has been difficult to study possible competition between fetal growth and maternal development as a factor underlying the excessive mortality of children born to young mothers in LMICs. Due to lack of reliable data on gynecologic age (i.e., age since menarche—a particularly important concern because of delayed age at menarche in LMICs [Foster, Menken, Chowdhury, & Trussell, 1986]) or chronological age of mothers younger than 20, Haaga (1989) concluded that there was only weak evidence for this proposed mechanism. Studies that have addressed social causes (i.e., poor knowledge and use of prenatal care) have used socioeconomic status as a crude proxy for use of prenatal care services. In multivariate analyses, this measure fails to help explain the high risk that accrues to infants and children born to young mothers.

Children born to older mothers may suffer because of poorer maternal health due to age-related declines in physiologic function and a higher risk of genetic abnormalities (Hansen, 1986). Little proof exists that this is a major risk factor in LMICs, however.

Short Birth Intervals

Short birth intervals, both prior and subsequent to the birth of a child, are probably the most consistent reproductive pattern identified as a risk factor for

excess child mortality. Hobcraft (1987) reports that the excess mortality risk of children born less than 24 months after the preceding birth compared to those born 24 months or more after the preceding birth is 1.8 in the first year of life, 1.3 for toddlers (ages 1 to 2 years), and 1.3 for children ages 2 to 5 years. In terms of subsequent birth intervals, on average, across 34 countries, children whose birth was followed by a subsequent birth within less than 24 months had 2.2 times the risk of dying of those children for whom the subsequent birth interval was longer (Hobcraft, 1987). As is the case for other demographic risk factors, considerable variation exists between countries in terms of risks related to short birth intervals.

A number of plausible explanations have been put forth for the relationship between short prior and subsequent birth intervals (less than 24 months) and a child's risk of poor health and increased mortality. First, due to maternal depletion (resulting from inadequate recovery time from the nutritional burdens of breastfeeding and prior pregnancy [Merchant & Martorell, 1987]), children born after a short birth interval may suffer poorer intrauterine growth, and possibly have a higher risk of preterm birth. To date, little empirical evidence has been collected that supports this mechanism (Ferraz, Gray, Fleming, & Maia, 1988; National Research Council, 1989; Pebley & DaVanzo, 1988; Winikoff & Sullivan, 1987).

Second, children born before a short birth interval may suffer from premature cessation of breastfeeding (which has been shown to be an important correlate of child survival in LMICs [Palloni & Millman, 1986]), as the mother shifts her attention to the more recent arrival. Given that studies that have controlled for the length of breastfeeding still show an association between short subsequent birth interval and high infant mortality (Pebley & Stupp, 1987), premature termination of breastfeeding does not entirely explain this effect.

Third, children born in close proximity to each other may suffer from competition for limited family resources of time and food. The evidence for this type of competition is unclear and sometimes contradictory (DaVanzo, Butz, & Habicht, 1983; Palloni, 1985).

Fourth, close birth spacing may increase the likelihood of transmission of infectious diseases such as diarrhea and measles, due to overcrowding and presence of children of similar ages (Aaby, Bukh, Lisse, & Smits, 1984).

Finally, despite adequate controls for observable confounding factors in multivariate analyses, part of the relationship between short birth intervals (either preceding or following the index birth) and increased child mortality may be due to unobserved factors such as short gestational length or parental characteristics. Babies born before or after very short birth intervals are known to be at high risk for short gestational durations, which independently have been shown to increase child mortality dramatically (Miller, 1989; Pebley & Stupp, 1987).

In terms of unobserved parental characteristics, it is possible that women who are likely to have short birth intervals are inherently at higher risk for poorer child health outcomes than their peers who have longer birth intervals. This would lead to a spurious inference that short birth intervals are causally related to higher child mortality (Pebley & Stupp, 1987; Potter, 1988; Rosenzweig & Schultz, 1983).

Unwanted Pregnancy and Birth

As discussed previously, unwanted children have much higher risks of morbidity and mortality. They may suffer from both conscious and unconscious neglect, due to smaller allocations of food, less parental time and attention, and less access to health care. In countries with a strong son preference (South and Southeast Asia), significant evidence points to higher mortality among female children relative to their male siblings (Das Gupta, 1987; D'Souza & Chen, 1980; Muhuri & Menken, 1997). Moreover, the recent rise in sex-selective abortion in China and Southeast Asia (which has led to a disproportionately high male-to-female sex ratio at birth) is evidence of the high risk of mortality for unwanted female fetuses (Larsen et al., 1998; Tsui et al., 1997). The pattern of gender discrimination is complex and nuanced, however, and may vary by societal setting (Muhuri & Menken, 1997).

Maternal Health

Maternal morbidity and mortality can have profoundly negative effects on child health, leading to high rates of morbidity and mortality. Population-based studies in South and Southeast Asia suggest that more than half of all perinatal deaths (i.e., deaths in the first week of a child's life) are associated with poor maternal health and pregnancy and delivery-related complications (Fauveau, Koenig, Chakraborty, & Chowdhury, 1988; Kusiako, Ronsmans, & Van der Paal, 2000; National Statistics Office [Philippines] & Macro International, 1994). These deleterious consequences may result from a combination of direct delivery-related consequences, such as premature labor, prolonged or obstructed labor, and abnormal fetal position (Kusiako et al., 2000); physiologic processes, such as cessation of breastfeeding following maternal morbidity and mortality

as well as maternal fetal transmission of a variety of infectious agents, including HIV, toxoplasmosis, cytomegalovirus (CMV), rubella, hepatitis B virus, herpes simplex, syphilis, malaria, and tuberculosis; and emotional impacts and lower levels of caregiving (National Research Council, 1989; Overall, 1987; Turner, Miller, & Moses, 1989; Weinbreck et al., 1988). This is particularly a major concern in sub-Saharan Africa, where significant numbers of mothers are infected with HIV and other STIs (National Research Council, 1989; Turner et al., 1989; Weinbreck et al., 1988).

Methodological Concerns

The previously mentioned mechanisms by which specific reproductive patterns affect infant and child health are certainly plausible and suggestive. Nevertheless, it is important to reiterate that the empirical evidence—in terms of the appropriateness of both data and statistical methods—supporting such mechanisms is variable and needs to be interpreted cautiously. Women have some control over their choices of reproductive patterns (i.e., whether to have children early or late, whether to have shorter or longer birth intervals, whether to have high parity births). Therefore, unobservable factors that are associated with both reproductive patterns and child health may be operating, and reverse causality may influence the results observed. As a consequence, our estimates of the impact of specific reproductive patterns on the risk of poor infant and child health may be overstated.

For example, if women who choose to be young mothers are prone to behavior patterns that devalue prenatal and postnatal care, delaying childbirth for these women will not produce the salutary effects that the earlier discussion suggests. Similarly, unobserved selection biases may operate such that a significant proportion of women who choose to have children at older ages, have higher-parity births, and have closely spaced births, are intrinsically in better health—most likely because of their higher socioeconomic status. If that is the case, then reducing higher-parity births, increasing interbirth intervals, and reducing births to older women (older than age 35) will not result in the degree of improvement that the current studies suggest.

With regard to reverse causality, one example is the often-cited relationship between the mortality of an index child and a short subsequent birth interval. The inference is that a child is at a higher risk of death if the next-younger sibling arrives after only a short interval, presumably because the pregnancy and the arrival of a newborn cause early cessation of breastfeeding for the older child and a shift of other maternal resources to the younger child. In reality, it is quite possible that the subsequent birth interval is short because the index child died, or was ill and weaned earlier because of his or her existing health problems. Thus, the direction of causation is not from the short subsequent birth interval to the death of the preceding child, but rather in the reverse direction—from the death of the preceding child to a short subsequent birth interval. Although in principle statistical methods can deal with this kind of potential bidirectionality (Rosenzweig & Schultz, 1983; Schultz, 1984), in practice relatively few published studies have employed such sophisticated methods of analysis.

In summary, because of the various methodological concerns, we should be careful not to overinterpret the evidence linking specific reproductive patterns and poor child health.

Summary of the Impact of Reproductive Patterns on Child Mortality

Despite these caveats about over-interpretation and overestimation, it is instructive to consider the impact of specific deleterious reproductive patterns on infant and child mortality. The National Research Council (1989) has simulated the impact of various reproductive patterns on child mortality rates using data from 18 LMICs reported by Hobcraft (1987). The simulations in **TABLE 5-12** refer to death rates that would be observed in individual families with particular reproductive patterns; they assume the mortality risks associated with the specific reproductive pattern are causative.

Table 5-12 shows that children of parity 2 and greater are at much higher risk of mortality if they are born to teenage mothers than if their mothers are ages 20 to 34 years. Moreover, both teenage and nonteenage mothers can significantly reduce the risks of child mortality by adopting better spacing patterns (i.e., birth intervals of 24 months or more). The best-case scenario for children of parity 2 and greater is for those born to mothers between the ages of 20 and 34, whose older sibling has survived, and whose birth interval is 24 months or more. Only 67 of 1,000 such children fail to reach their second birthday. This is less than half the mortality risk of their peers born to teenage mothers and whose birth intervals are less than 24 months. In the latter situation, 165 of 1,000 such children die before age 2 years.

Other simulations show the deleterious consequences of increasing family size and birth intervals on the probability of a child surviving to his or her fifth birthday. These calculations were carried out under low, moderate, and high baseline child mortality rates

TABLE 5-12 Estimated Risk of Dying (Deaths per 1,000 Births) Prior to Their Second Birthday for Second- and Higher-Order Births to Women with Different Reproductive Patterns

Age of Mother	Better Spacing Pattern	Poor Spacing Pattern
Teenage mothers	92	165
Mothers ages 20–34	67	120

Note: Better spacing pattern: Birth intervals both preceding and subsequent to this birth were 24 months or greater and the older sibling survived. Poor spacing pattern: Birth intervals both preceding and subsequent to this birth were less than 24 months and the older sibling survived.
Data from the National Research Council. (1989). *Contraception and reproduction: Health consequences for women and children in the developing world.* Washington, DC: National Academy Press; Hobcraft, J. N. (1987). Does family planning save lives? Paper presented at the International Conference on Better Health for Women and Children Through Family Planning, Nairobi, Kenya.

(to take into account variations in overall mortality among populations). The baseline child mortality rates represent the probability of a child surviving to his or her fifth birthday for children who have the lowest risk profile—that is, parity 2–3, preceding and subsequent birth intervals of 24 months or more, and the older sibling survived. Thus, in a population with a baseline child mortality rate of 150/1,000, out of 1,000 births of children who were parity 2–3, had long preceding and subsequent birth intervals (24 months or more), and whose older sibling survived, 150 would die before their fifth birthday.

Mortality rates for specific combinations of parities below and above the baseline, of short and long intervals, and of survival of older sibling were estimated by Hobcraft (1987). He then simulated the average number of children per 1,000 births who would die before their fifth birthday. Here we discuss only those cases where the older sibling survived. In terms of family size, small families were much better off than large families. Regardless of the baseline mortality rates or the closeness of birth spacing, the larger the family size, the higher the child mortality rates. In all cases, four-child families experienced fewer than half the deaths per 1,000 births of nine-child families. Similarly, families with long spacing experienced half or less the mortality per 1,000 births of families with consistently short spacing. Clearly, the most beneficial scenario for children is that of well-spaced births and small overall family sizes.

Impact of Reproductive Patterns on the Health of Women

Pregnancy is one of the major health risks for women in LMICs. Nearly 303,000 women die worldwide each year due to pregnancy-related causes, and the vast majority (99%) of these deaths occur in LMICs (WHO, 2015b). Although these numbers are alarming, it is important to recognize that 210 million pregnancies and approximately 140 million births occur annually in the world (Graham et al., 2016); thus, by and large, reproduction is relatively safe for women.

Maternal mortality risks are a fraction of infant mortality risks. For example, Bangladesh has still high infant and maternal mortality risks, but the latter is roughly a 1/20th of the former. The maternal mortality ratio is approximately 176 deaths per 100,000 births, while the infant mortality rate (which has fallen considerably in the last decade) is approximately 38 deaths per 1,000 births (Macro International, 2015; WHO, 2015b). Similarly, for Kenya, the maternal morality ratio, 510 deaths per 100,000 births, is roughly one-eighth of the infant mortality rate, approximately 39 deaths per 1,000 births (Macro International, 2015; WHO, 2015b).

Definitions

In any discussion of maternal mortality, a number of potentially confusing definitional issues arise. The first is the definition of a maternal death. A maternal death is usually defined as the death of a woman while pregnant or up to 42 days post delivery from any cause (except accidents). There has been some discussion as to whether this definition is overly restrictive (i.e., leading to an undercount of maternal deaths) and should be expanded to include female deaths up to 90 days post delivery. In reality, data from a number of well-conducted population-based studies using different post-delivery durations show that the majority of maternal deaths occur within 42 days post delivery, with approximately 40% occurring within 24 to 48 hours of delivery. Furthermore, extending the definition to up to 90 days would result in only a marginal increase (6%) in the number of deaths classified as related to maternal causes (Egypt Ministry of Health, 1994; Fauveau et al., 1988).

The second issue is the measure of maternal mortality risk that should be used in comparing and contrasting the situations in different populations both geographically and across time. Maternal mortality risks are conventionally described using three measures. It is important to understand and to think of these measures separately, because they are conceptually distinct.

The first measure is the *maternal mortality ratio*, which is defined as the ratio of the number of maternal deaths to the number of pregnancies. It is an indicator of the risk of death that a woman faces for each pregnancy she undergoes. Although conceptually the denominator for such a risk measure should include all pregnancies, operationally, because of the difficulty of counting miscarriages and induced abortions, the denominator used is live births.

The second measure is the *maternal mortality rate*, which is defined as the number of maternal deaths divided by the number of women of reproductive age (i.e., between ages 15 and 49). This composite measure is the product of the maternal mortality ratio (number of maternal deaths per births) and the birth rate in the reproductive age group (number of births for women between ages 15 and 49). The maternal mortality can be changed by altering the frequency of pregnancies or births in the population without changing the risk of maternal death per pregnancy/birth. Although the maternal mortality ratio and the maternal mortality rate are conceptually distinct, they are often confused in the public health literature, with rates referring to ratios, and vice versa. In this book, these two measures are carefully distinguished.

The third measure is the *lifetime risk of maternal mortality*. Also a composite measure, it not only takes into account the maternal mortality risk per pregnancy, but also factors in the cumulative exposure to pregnancy that an individual woman experiences. The average cumulative exposure to pregnancy is usually taken to be the total fertility rate for the population. It is an estimate of the number of births a woman in a particular society would have over her lifetime if she were to adhere to the current age-specific fertility rates in that population.

The lifetime maternal mortality risk for a woman in one of the LMICs in 2005 was estimated to be 1/75 (Population Reference Bureau, 2008; United Nations Population Fund [UNFPA], 2004, 2009). This estimate can be interpreted as follows: A woman in the LMICs who (1) has the same total number of pregnancies over her lifetime as the current fertility norm (estimated to be 2.7 births) and (2) experiences at each pregnancy the same independent risk of maternal death as the current maternal mortality ratio (approximately 500 maternal deaths per 100,000 births) would have 1 chance in 75 of dying from pregnancy-related causes. This lifetime risk captures both the risk of dying per pregnancy and the cumulative effect of exposure to multiple pregnancies.

Maternal Mortality Risks

A major constraint with respect to investigating the magnitude of maternal mortality risks and its determinants is the lack of available and reliable population-based data. Even in LMICs, many of which have high maternal mortality rates and ratios, maternal deaths are relatively rare. For example, in sub-Saharan Africa and South Asia, where maternal mortality ratios of 800 maternal deaths per 100,000 live births have been reported, one would need very large sample sizes to obtain reasonable estimates of maternal mortality risks and their accompanying determinants. A sample of 10,000 births (a very large sample by any standards) would be expected to yield only 80 maternal deaths. In contrast, infant mortality rates in these settings are typically 15 times as large, and the same sample would yield 1,050 infant deaths.

To date, relatively few large-scale population-based studies of maternal mortality have been conducted in LMICs. The major investigations were carried out in Bangladesh (Alauddin, 1986; Chen, Gesche, Ahmed, Chowdhury, & Mosley, 1974; Fauveau et al., 1988; Koenig et al., 1988; National Institute of Population Research and Training [NIPORT], 2003), Ethiopia (Kwast, Rochat, & Kidane-Mariam, 1986), Egypt (Egypt Ministry of Health, 1994; Fortney et al., 1985), and Jamaica (Walker, Ashley, McCaw, & Bernard, 1985). Estimates for other countries are derived from model-based assumptions and, therefore, are not as precise.

Huge disparities in maternal mortality are observed among various regions of the world. The disparity between LMICs and high-income countries is much greater for maternal mortality (20 times higher risk of maternal death per pregnancy in 2015) than for infant mortality (10 times higher risk of infant death per pregnancy). Lifetime risks of maternal mortality vary from 1/41 in African countries to 1/4,900 in in high-income countries (**TABLE 5-13**).

Table 5-13 shows the estimates of total number of pregnancies per woman (TFR), maternal mortality ratio (MMR), individual risk of dying per pregnancy (lifetime risk), and total number of maternal deaths in different regions of the world in 2005 and 2015. Although global maternal mortality was almost halved in 2015 relative to the 2005 maternal mortality, all the estimates are still much higher in LMICs than

TABLE 5-13 Total Fertility Rate, Maternal Mortality Ratios, and Lifetime Risks of Maternal Death by Region of the World, 2005 and 2015

					Maternal Deaths			
Region	Total Fertility Rate (births per woman)		Maternal Mortality Ratio (deaths per 100,000 live births)		Lifetime Risk		Deaths per Year	
	2005	2015	2005	2015	2005	2015	2005	2015
World total	2.6	2.5	400	216	1 in 92	1 in 180	536,000	303,000
High-income countries	1.7	1.7	9	12	1 in 7,300	1 in 4,900	960	1,700
Low- and middle-income countries	2.7	2.6	450	238	1 in 75	1 in 150	533,000	302,000
Africa	4.8	4.7	820	495	1 in 26	1 in 42	276,000	204,000
Asia	2.3	2.1	330	119	1 in 120	1 in 370	241,000	90,000
Latin America and Caribbean	2.3	2.1	130	67	1 in 290	1 in 670	15,000	7,300

Data from Population Reference Bureau, Family Planning Worldwide: 2015 Data Sheet. Washington, DC: The Population Reference Bureau. Adapted with permission. Accessed on November 8, 2016 from http://www.prb.org/Publications/Datasheets/2015/2015-world-population-data-sheet.aspx; Trends in maternal mortality: 1990 to 2015: estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division. Accessed on November 8, 2016 from http://www.who.int/reproductivehealth/publications/monitoring/maternal-mortality-2015/en/

in high-income countries. Even so, it is the risk of dying per pregnancy that accounts for the vast majority of the difference in lifetime risk of maternal mortality. Total fertility rates in LMICs were, on average, 1½ times as high as in the high-income countries (2.6 births per woman versus 1.7 births per woman). Maternal mortality ratios were still 20 times as high (238 maternal deaths per 100,000 births in LMICs versus 12 maternal deaths per 100,000 births in high-income countries), although they were about 50 times higher in 2005. Only one country (Sierra Leone) in 2015 had an MMR of more than 1,000 deaths per 100,000 live births, compared to 14 countries—Afghanistan, Angola, Burundi, Cameroon, Chad, Democratic Republic of the Congo, Guinea-Bissau, Liberia, Malawi, Niger, Nigeria, Rwanda, Sierra Leone, and Somalia—in 2005. However, still 19 countries had MMRs of more than 500 deaths per 100,000 live births; all of these countries were in sub-Saharan Africa: Burundi, Cameroon, Central African Republic, Chad, Cote d'Ivoire, Democratic Republic of the Congo, Eritrea, Gambia, Guinea, Guinea-Bissau, Kenya, Liberia, Malawi, Mali, Mauritania, Nigeria, Sierra Leone, Somalia, and South Sudan (WHO, 2015b).

Direct and Indirect Causes of Maternal Mortality and Morbidity

What are the sources of the mortality risk per pregnancy? The causes of maternal mortality are conventionally divided into direct causes—those that occur only during pregnancy and the immediate post-delivery period—and indirect causes—those derived from conditions that precede, but are aggravated by, pregnancy, such as anemia, diabetes, malaria, tuberculosis, cardiac disease, hepatitis, and increasingly AIDS (Ronsmans et al., 2006; WHO, 2005).

In LMICs, direct causes account for 73.0% of maternal mortality and include, in approximate order of importance, hemorrhage (27.1%), hypertensive disorders of pregnancy (eclampsia—14%), sepsis (10.7%), complications of unsafe abortion (7.9%), and embolism and other direct causes (1.8%) (Filippi, Chou, Ronsmans, Graham, & Say, 2016). Obstructed labor and complications due to delivery, each of which accounts for 2.8% or maternal deaths, are included in the "other causes" group (Filippi et al., 2016). The vast majority of these maternal deaths can be attributed to just three causes: hemorrhage, eclampsia, and sepsis.

Attribution of cause of death is complicated by the fact that in most cases unsafe abortion and obstructed or prolonged labor eventually cause death via the proximate causes of hemorrhage or sepsis.

There is variation in the order of importance of these causes in different studies from different parts of the world (Jamison, Mosley, Measham, & Bobadilla, 1993), partly due to real differences in the availability and use of obstetric care and partly due to differences in the quality of reporting. Thus, in countries that have poor access to obstetric care facilities, a relatively large proportion of deaths are attributed to hemorrhage, sepsis, and abortion. Differential reporting—particularly reluctance to attribute maternal deaths to abortion in countries where it is illegal—also artifactually inflates the proportion of deaths attributed to hemorrhage and sepsis. For example, Jamison and associates (1993) estimated that, for the period 1980–1985, these two causes were responsible for 56% of maternal deaths in Indonesia, 40% of such deaths in Egypt, and 18% of maternal deaths in the United States.

The remaining 27% of maternal deaths can be attributed to illnesses aggravated by pregnancy (Jamison et al., 1993; WHO, 1993b). For example, anemia hampers a woman's abilities to resist infection and to survive hemorrhage; it may increase the likelihood of her dying in childbirth by a factor of 4 (Chi, Agoestina, & Harbin, 1981). Hepatitis can cause hemorrhage or liver failure in pregnant women (Kwast & Stevens, 1987). Latent infections such as tuberculosis, malaria, or STIs can also be activated or exacerbated during pregnancy and cause potentially severe complications for both mother and child (Jamison et al., 1993).

In keeping with the high rates of maternal mortality in LMICs, there are also high rates of maternal morbidity in these countries. An estimated 30 to 50 morbidities (temporary and chronic) occur for every maternal death (Safe Motherhood Initiative, 2010). Between 30% and 40% of the approximately 180 million women who are pregnant annually in the world, or roughly 54 million women, report some kind of pregnancy-related morbidity (Koblinsky, Campbell, & Harlow, 1993; WHO, 1993b). Of these women, 10 to 15 million each year develop relatively long-term disabilities deriving from complications from obstetric fistula or prolapse, uterine scarring, severe anemia, pelvic inflammatory disease, reproductive tract infections, or infertility (Filippi et al., 2006; Ronsmans et al., 2006; Tsui et al., 1997).

These figures demonstrate significant variability from one country to another (Tsui et al., 1997). For example, Guatemalan women report one in five pregnancies as being complicated (Bailey, Szászdi, & Schieber, 1994), women in West Java report one in three pregnancies as being complicated (Alisjahbana et al., 1995), and two out of three pregnancies in Ghana had some complications (De Graft-Johnson, 1994). This variability in rates stems at least in part from differences in study design and data quality. To date, relatively few population-based surveys have been carried out. Moreover, maternal morbidity data are often based on self-reported symptoms—a data collection technique that has been shown to have relatively low reliability and validity (Stewart & Festin, 1995). A few data sets with reliable information have come from the United States and Canada (Koblinsky et al., 1993; WHO, 1994) and from specific validation studies (Stewart & Festin, 1995). These data show that, on an annual basis, 12% to 15% of women who are pregnant suffer life-threatening obstetric complications, equivalent to approximately 20 million women in LMICs (assuming approximately 150 million women giving birth annually). In contrast to the evidence on acute pregnancy-related morbidity, little is known about the long-term chronic morbidity sequelae of pregnancy-related complications, which may significantly affect women's lives.

Specific Causes of Pregnancy-Related Morbidity and Mortality

This section focuses on some of the more prominent causes of pregnancy-related morbidity and mortality in LMICs.

Obstructed and Prolonged Labor

Obstructed or prolonged labor leads to approximately 8,500 maternal deaths annually, with high proportions of the survivors developing obstetric fistulas and their newborns often experiencing long-term sequelae of anoxia and even death (Kusiako et al., 2000; Say et al., 2014). Predictive risk factors for this outcome are not particularly reliable (Fortney, 1995; Maine, 1991). Monitoring during labor using a partograph is the only effective way to detect such problems (Mathai, 2009).

Obstetric Fistula and Genital Prolapse

Both obstetric fistulas and genital prolapse have severe consequences for pregnant women. As a consequence, they represent a major global health problem.

An obstetric fistula is a passage or channel from the vaginal wall to either the rectum (recto-vaginal fistula) or the bladder (vesico-vaginal fistula). It is usually a result of a tear in the vaginal wall during complicated labor. Risk factors for obstetric fistulas include being a young mother, a stunted mother, or a mother who

has complicated labor and delivers in a nonhospital setting with the help of traditional birth attendants (Lawson, 1992; Tahzib, 1983, 1985). Population-level estimates of obstetric fistulas are difficult to find, but a recent meta-analysis reports that the population-based pooled prevalence of obstetric fistula is 0.29 case per 1,000 women of reproductive age in all regions, with the highest prevalence in sub-Saharan Africa (1.60 cases per 1,000 women) and South Asia (1.20 cases per 1,000 women) (Adler, Ronsmans, Calvert, & Filippi, 2013). The prevalence of obstetric fistula is estimated to be 1.69 per 1,000 ever-married women in Bangladesh (Waiz et al., 2003).

The consequences of fistulas are quite severe, especially for young primipara women. The baby is often stillborn, and the mother is incontinent of urine and/or feces. This condition is a source of enormous personal discomfort, whose consequences are exacerbated by social stigma. In many cases, it leads to divorce and social ostracism—these women are often barred from food preparation or even participating in prayer, due to lack of personal hygiene (Reed et al., 2000).

Genital prolapse occurs when the vagina and uterus descend below their normal positions. This condition is usually a result of damage to supporting muscles and ligaments during childbirth and is most often associated with high parity. It is particularly uncomfortable for women who are squatting, which is the normal position for doing many chores in LMICs. It can also lead to chronic backache, urinary problems, and pain during sexual intercourse. Subsequent pregnancies have a higher probability of fetal loss and further maternal morbidity.

Although good estimates are difficult to obtain, some reliable population studies suggest that 2% to 20% women younger than age 45 suffer from genital prolapse worldwide (Kuncharapu, Majeroni, & Johnson, 2010). The Giza study (Younis et al., 1993), which clinically validated reported prolapse, found that one-third of women suffered from genital prolapse and also documented a relationship between genital prolapse and risk of reproductive tract infections.

Anemia

Approximately 50% of pregnant women around the world are estimated to be anemic (i.e., to have hemoglobin levels less than 11 g/dL). Dietary iron deficiency is the primary cause of this condition, followed by malaria, other parasitic diseases (schistosomiasis, hookworm), folate deficiency, AIDS, and sickle cell disease (Tsui et al., 1997). In addition to its well-documented effects on pregnancy outcomes—prematurity, stillbirths, and spontaneous abortions (Levin, Pollitt, Galloway, & McGuire, 1993)—anemia, even at fairly mild levels, has been implicated as directly contributing to maternal deaths (Harrison & Rossiter, 1985; United Nations, 1991). In addition, some evidence suggests that anemia predisposes women to higher risks of complications during pregnancy, including urinary tract infections, pyelonephritis, and preeclampsia (Kitay & Harbort, 1975). Anemia is also associated with reduced productivity and quality of life for women (Bothwell & Charlton, 1981).

Data on the effectiveness of iron supplementation in reducing the prevalence of iron-deficiency anemia are not encouraging (Sloan, Jordan, & Winikoff, 1992). While a significant part of the failure to reduce anemia is due to inadequate efforts to provide iron supplementation, even in situations where properly conducted supplementation trials have been conducted, these programs do not appear to be very effective in reducing anemia levels (Sood et al., 1975). The relatively poor outcomes may reflect inadequacies of strategies that focus just on pregnant women. Long-term success probably will require the use of a multipronged strategy including iron supplementation schemes, efforts to raise household income, and efforts to reduce workload during pregnancy (Tsui et al., 1997).

Pregnancy-Related Hypertension

Both eclampsia and preeclampsia are significant causes of maternal morbidity. Unfortunately, these conditions are difficult to predict and prevent, although routine prenatal blood pressure measurements and urinalysis for proteinuria in the first prenatal visit continue to be recommended for this purpose (Rooney, 1992; Stone et al., 1994). Women with moderate hypertension and proteinuria require appropriate follow-up. Treatment options and their effectiveness vary, so no definitive conclusions or recommendations can be made. As a general rule, a combination of bed rest, antihypertensive agents, and anticonvulsant medications (especially magnesium sulfate for frank convulsions) may provide some relief (Eclampsia Trial Collaborative Group, 1995; Rooney, 1992).

Consequences of Pregnancy and Delivery Complications for Infants

Pregnancy- and delivery-related complications have important health consequences not only for mothers but also for infants, particularly stillbirths (intrauterine

fetal death after 28 weeks' gestation) and neonatal deaths (death of live-born babies in the neonatal period—that is, the first 28 days following birth/delivery), especially in the early neonatal period (i.e., the first 7 days of birth/delivery) (Kusiako et al., 2000; WHO, 2006; Zupan, 2005). Each year, 2.5 million stillbirths and 5.9 million under-5 child deaths occur globally, resulting in total 8.4 million deaths from the 28th week of pregnancy to 5 years of age among the children. Out of the 5.9 million under-5 child deaths (not including stillbirths), 2.7 million (45%) occur during the neonatal period (Liu et al., 2016). Notably, the proportion of under-5 child deaths that occur during the neonatal period has been increasing over time, due to the fact that the decline in neonatal deaths has been slower than the decline in post-neonatal and 1- to 4-year-old child deaths globally. Developing countries account for more than 99% of both stillbirths and neonatal deaths (Blencowe et al., 2016; Liu et al., 2015). Of the 2.8 million neonatal deaths, 73% (2.02 million deaths) occur in the first 7 days of life (Oza, Cousens, & Lawn, 2014).

In LMICs, neonatal mortality declined from 40 deaths per 1,000 live births in 1990 to 21 deaths per 1,000 live births in 2015. These countries have five times the neonatal mortality rate of high-income countries. The highest neonatal mortality rates are found in South Asia and sub-Saharan Africa (30 and 29 neonatal deaths per 1,000 live births, respectively), and the lowest rates in East Asia and Pacific (9 neonatal deaths per 1,000 live births) (You, Hug, Ejdemyr, & Beise, 2015). Early neonatal mortality rates (ENMR—deaths in the first 7 days of life) declined from 23 per 1,000 live births in 2000 to 14.7 per 1,000 live births in 2015 (Oza et al., 2014; WHO, 2006). The global rate of stillbirth in 2015 was 18.4 per 1,000 births, with the highest rate found in sub-Saharan Africa (28.7 per 1,000 births) and the lowest rate in high-income countries (3.4 per 1,000 births) (Blencowe et al., 2016). The early neonatal death rate in 2013 was 14.7 per 1,000 live births globally, with the ENMR ranging from 22.7 per 1,000 livebirths in sub-Saharan Africa to 2.6 per 1,000 live births in high-income countries (Oza et al., 2014). Most neonatal deaths occur in Asia and Africa, where most children are born (Liu et al., 2016). Within Asia, the most strongly affected region is South Asia (e.g., Afghanistan, Bangladesh, India, Nepal, and Pakistan), which accounts for more than 39% of global neonatal deaths (Liu et al., 2016).

In 2015, a total of 5.9 million under-5 child deaths occurred globally (not including stillbirths), of which 2.7 million deaths occurred in the neonatal period (first 28 days of life), 1.8 million deaths occurred between 29 days and 1 year of age, and 1.4 million deaths occurred between 1 and 5 years of age. The number of infant deaths (death between birth and 1 year of age) was 4.5 million (Liu et al., 2016). Thus, neonatal deaths accounted for approximately 60% (2.7 million/4.5 million) of all infant deaths and 45% (2.7 million/5.9 million) of all under-5 child deaths on a worldwide basis. As infant mortality rates drop, the proportion of infant deaths attributed to the neonatal period actually increases, because post-neonatal deaths (those after the first 28 days of life, which are most sensitive to environmental contamination and amenable to public health interventions) are the first to decrease (Black, Morris, & Bryce, 2003; Hill & Choi, 2006; WHO, 2006; Zupan, 2005).

Of the estimated 2.6 million stillbirths and 2.7 million neonatal deaths worldwide, the vast majority are associated with maternal health problems during pregnancy (e.g., preterm birth, maternal infections, and severe malformations) and around delivery (e.g., intrapartum complications or trauma during birth, infections such as tetanus). The contribution of each of these factors varies with the level of neonatal mortality. In areas where neonatal mortality is high, birth trauma and infections play larger roles; in contrast, in areas where neonatal mortality is low, preterm birth and severe malformations predominate (Beck et al., 2010; Fauveau, Wojtyniak, Mostafa, Sarder, & Chakraborty, 1990; Hill & Choi, 2006; Zupan, 2005). Most of the adverse outcomes leading to stillbirth and neonatal death are due to poor antenatal care (e.g., non-immunization against tetanus, poor prevention of HIV and other STIs) inadequately treated maternal conditions such as eclampsia, complications of labor and delivery, poor-quality neonatal care, and harmful traditional practices such as discarding the colostrum, nonsterile umbilical cord cutting, and failure to keep babies warm. Roughly half of the 2.6 million stillbirths are due to delivery complications and could largely be prevented by skilled care during delivery and in the first 24 hours post delivery (de Bernis et al., 2016).

One of the most important risk factors for neonatal mortality is low birth weight (LBW), with babies weighing less than 2,500 grams having 20 to 30 times the mortality risk of babies of normal weight. In addition to increasing the risk of neonatal mortality, low birth weight is associated with a substantial burden of long-term disability (e.g., impaired immune function and greater susceptibility to infection, long-term undernourishment and decreased muscle strength throughout the child's life, inhibited cognitive development, higher incidence of diabetes and heart disease, cerebral palsy, seizures, and severe learning disorders) for babies who survive (Jamison et al., 1993; Tsui et al.,

1997; UNICEF & WHO, 2004; WHO, 2006; Zupan, 2005). Women who have inadequate nutritional status (including short stature, poor prepregnancy weight, inadequate weight gain during pregnancy, and anemia) or infections during pregnancy are more likely to have LBW babies (UNICEF & WHO, 2004; WHO, 1993b; Zupan, 2005).

When assessing outcomes for LBW babies, it is necessary to distinguish between those who are born preterm (less than 37 weeks' gestation) but appropriate for their gestational age and those who are full term and inappropriately small for their gestational age due to intrauterine growth retardation (IUGR). Considering gestational age at birth and birth weight, babies can be classified into four groups:

1. Full term and appropriate weight for gestational age (AGA)
2. Full term and small for gestational age (SGA)
3. Preterm and AGA
4. Preterm and SGA

In LMICs, the majority of LBW babies are full term but suffer from IUGR (Group 2). In contrast, in the high-income world, most LBW babies are preterm and AGA (Group 3).

The risk of neonatal mortality substantially varies across these groups. In a recent study reported that compared term and AGA babies, term and SGA babies were at lowest risk of neonatal mortality (relative risk [RR]: 2.44; 95% confidence interval [CI]: 1.67–3.57), preterm and AGA babies were at higher risk of neonatal mortality (RR: 8.05; 95% CI: 3.88–16.72), and preterm and SGA babies had highest risk of neonatal mortality (RR: 15.42; 95% CI: 9.11–26.12) (Katz, Lee, Kozuki, & Black, 2015). In 2013, out of approximately 138 million babies born worldwide, 16%—22 million—were LBW; 96% of these infants were born in developing countries (UNICEF, 2013). The level of LBW in LMICs was more than double that in high-income countries (16.5% versus 7%), with the highest rates being found in South Asia (27%). Of 18 million LBW babies born in 2010 in LMICs, 59% (10.6 million) were term-SGA, 16% (2.9 million) were preterm-SGA, and 25% (4.5 million) were preterm-AGA (Lee et al., 2013).

Preterm birth is a major risk factor for neonatal mortality and morbidity, accounting for 35% of neonatal mortality other than those deaths related to congenital malformations (Liu et al., 2015). It is estimated that 11.1% of all births worldwide in 2010 were preterm, which translates into a total of 14.9 million preterm births (Blencowe et al., 2013). In light of the earlier discussion of LBW and IUGR, it is important to note that most preterm babies are, in fact, appropriately developed for their gestational age and have LBW (i.e., weigh less than 2,500 g) because of premature delivery. In 2010, the majority of these births occurred in Asia (53%), followed by Africa (20%), the Caribbean (11.2%), Latin America (8.4%), and Oceania (7.4%), with only 8.6% taking place in Europe and North America (Blencowe et al., 2012). A variety of factors are associated with preterm birth, including medical conditions of the mother or the fetus, genetics, environmental exposures, infertility treatments, behavioral and socioeconomic factors, and iatrogenic interventions (Goldenberg, Culhane, Iams, & Romero, 2008).

High-Risk Pregnancies

For the sake of simplicity, the calculations of lifetime risk presented earlier assumed that the risk of maternal death per pregnancy is the same across all women and across successive pregnancies within each woman. The reality is somewhat more complicated, with some types of pregnancies being riskier than others. Higher-risk conditions include pregnancies of first-time mothers, mothers with multiple previous pregnancies (five or more pregnancies), very young and older mothers, women already in poor health, and pregnancies that are terminated by unsafe abortions (National Research Council, 1989; Tsui et al., 1997). **TABLE 5-14** summarizes the hypothesized mechanisms by which different reproductive patterns affect maternal health.

First pregnancies have a higher risk of maternal mortality than subsequent pregnancies (up to five) both in high-income countries and in LMICs (Filippi et al., 2016). Population-based data from Bangladesh (Koenig et al., 1988), Ethiopia (Kwast et al., 1986), and the Gambia (Greenwood et al., 1987) suggest that first pregnancies may be as much as three times riskier than later pregnancies.

The impact of young maternal age on maternal mortality is more difficult to evaluate, because relatively few studies have disaggregated the confounding effects of young maternal age and first pregnancy. Among those that have controlled for confounding factors appropriately, the largest study to date looked at maternal mortality among 14,631 first births in rural Bangladesh (Koenig et al., 1988); it showed no age effect. Other smaller studies from the same area (Chen et al., 1974), Indonesia (Chi et al., 1981), and Jamaica (Walker et al., 1985) have shown a higher mortality risk for mothers younger than age 20 compared to mothers age 20 to 24. A more recent large study shows that the mortality risk pattern is similar

TABLE 5-14 Mechanisms by Which Reproductive Patterns Affect Maternal Health

Reproductive Pattern	Mechanism Through Which Maternal Health Is Affected
Number of pregnancies	Each pregnancy carries a risk of morbidity and mortality
High-Risk Pregnancies	
First-time mothers	Higher risk than pregnancies 2–4 for obstructed labor, pregnancy-induced hypertension, and other obstetric complications due to initial adaptation to pregnancy
Higher-order pregnancies	Higher risk for hemorrhage and uterine rupture, due to the cumulative toll of previous pregnancies and reproductive injuries
Pregnancy at very young maternal ages	Higher risk due to physiologically immature reproductive systems and reduced propensity for timely care seeking
Pregnancy at older maternal ages	Body in poorer condition for pregnancy and childbirth
Short interbirth intervals	Inadequate time to rebuild nutritional stores and regain energy levels
Unwanted pregnancies ending in unsafe abortions	Unsafe abortions increase exposure to injury, infection, hemorrhage, and death
Pregnancies for women already in poor health	Aggravated health condition

in all regions, with the lowest risk being in 20- to 24-year-old women. A "J" shaped curve, with markedly higher risk after age 30, is evident in all regions (Blanc, Winfrey, & Ross, 2013). A further problem that makes interpretation of data difficult is that maternal age younger than 20 is not disaggregated into single years of age; in fact, the highest risk may be for young mothers in the 15- to 17-year-old age group (Harrison & Rossiter, 1985).

The major causes of morbidity and mortality for young primigravida women include a high risk of pregnancy-induced hypertension (WHO, 1988), a high frequency of obstructed labor due to the pelvis being too small for the child's head to pass (Aitken & Walls, 1986), and a high incidence of placental malaria (McGregor, Wilson, & Billewicz, 1983).

A number of studies have shown that women with four or more pregnancies have 1.5 to 3 times the risk of maternal death as women of parities 2 and 3. In general, for any given parity, older women—particularly those beyond age 35—have higher risks of maternal death (Chi et al., 1981; Koenig et al., 1988; Walker et al., 1985). A major cause of maternal morbidity and mortality for older multiparous women is the higher risk of malpresentation (in which the fetus assumes a position other than the usual head-first delivery, as in a breech or transverse lie presentation). Malpresentation of the fetus may occur due to the flaccidity of the uterine wall from repeated stretching from successive pregnancies. It can lead to uterine rupture, hemorrhage associated with rupture, or infections resulting from unsuccessful attempts to deal with malpresentation. Another major cause of morbidity and mortality in older multiparous women is hemorrhage due to placental abnormalities such as placenta previa (where the placenta overlies the cervical opening of the uterus) and placenta abruptio (where the placenta separates prematurely from the uterus prior to delivery of the baby) (Faundes, Fanjul, Henriquez, Mora, & Tognola, 1974).

There has been much talk about the possibility of a maternal depletion syndrome whereby multiple short birth intervals result in women not having enough time to recover their energy and nutritional levels between pregnancies, which in turn may lead to higher risks of maternal mortality (Jelliffe, 1976; Winikoff, 1983). As yet, no convincing evidence validating this hypothesis has been

presented (Koenig et al., 1988; National Research Council, 1989; Ronsmans & Campbell, 1998). This could be due to the fact that intrinsically healthier women (in this case, women of higher socioeconomic status) may be more likely to have multiple short birth intervals, whereas less healthy women may take longer between subsequent births (Duffy & Menken, 1998).

Pregnancy is more dangerous for women who are already in poor health. It increases the likelihood that a woman will die if she has certain preexisting conditions, such as malaria, hepatitis, anemia, sickle cell disease, or rheumatic heart disease (Koblinsky et al., 1993; Morrow, Smetana, Sai, & Edgcomb, 1968; National Research Council, 1989; Tsui et al., 1997; WHO, 1993b).

Finally, unsafe abortions are a significant cause of maternal death in countries where abortion is not legal and regulated. Kwast, Rochat, and Kidane-Mariam (1986) reported that the primary cause of maternal mortality in Addis Ababa, Ethiopia, especially among primigravida, unmarried women, is complications from illegal abortions. In Bangladesh, abortion-related deaths declined dramatically from 18% in the 1980s to 5% in 2001, and then to only 1% in 2010 (Koenig et al., 1988; NIPORT, 2011). The major reason for this remarkable decline is wide availability of menstrual regulation (MR) service, although abortion is not legally approved in Bangladesh (Benson, Andersen, & Samandari, 2011). More recent global data indicate that in LMICs, a woman dies from an unsafe abortion every 8 minutes (WHO, 2007b). Approximately one in every four women having an unsafe abortion develops significant medical complications, and unsafe abortion causes approximately 8% of all maternal deaths and 20% of the overall burden of maternal death and long-term sexual and reproductive ill health.

Mechanisms to Reduce Maternal Morbidity and Mortality in Low- and Middle-Income Countries

Although significant progress was made toward Millennium Development Goal 5, which required reduction of maternal mortality by 75% between 1990 and 2015, the target was not achieved at the end (UNICEF & WHO, 2015). The global maternal mortality ratio has been reduced by roughly half since 1990 (from 385 to 216 per 100,000 live births) and the rate of improvement has accelerated since 2000. Only 6 countries achieved MDG 5 (Maternal health) out of 75 countdown countries, while only 4 (Cambodia, Eritrea, Nepal, and Rwanda) achieved both MDG 4 (Child Health) and 5 (UNICEF & WHO, 2015). Bangladesh, for example, was able to reduce maternal mortality significantly during the MDG era, from 569 maternal deaths per 100,000 live births in 1990 to 196 maternal deaths per 100,000 live births in 2016 (NIPORT, icddr,b & MEASURE Evaluation). Although the MDG 5 target was not fully achieved (the target was 143 maternal deaths per 100,000 live births), the progress made in Bangladesh was remarkable. Kenya, however, did not perform well in terms of MDG 5. The MMR was reduced from 678 in 1990 to only 510 in 2015 against the MDG target of 170.

The reasons behind the significant decline in the MMR in Bangladesh is attributed to factors both within and outside the health sector. Improved access to and use of health facilities are the key contributors of MMR reduction, along with a decrease in the proportion of births associated with high risk of maternal mortality, a decline in the poverty rate, and significant improvements in female education (El Arifeen et al., 2014).

In addition to reducing the risk per pregnancy, a strategy focused on reducing teenage fertility in Bangladesh (which remains stubbornly high, despite the country's great success in lowering overall TFR to near replacement levels) will most likely lead to modest but significant reductions in MMR. Nevertheless, social/political constraints in Bangladesh continue to be a barrier in increasing the average age at marriage (*Dhaka Tribune*, 2017, p. 3).

Weak health systems, continued high fertility, lack of supportive polices, inadequate financial investment, and poor availability of data to monitor change were the major barriers to achieving MDG 5 in many LMICs, where 99% of maternal deaths occur, and particularly in sub-Saharan Africa (Koblinsky et al., 2006; Ronsmans et al., 2006; UNICEF & WHO, 2015; WHO, 2008a). Researchers have suggested that a multisectoral approach will be required to decrease the number of maternal deaths through reduced exposure to pregnancy (i.e., by reducing fertility), optimization of access to emergency obstetric care, and improvement of general health status and treatment of pregnancy- and childbirth-related complications (Bartlett et al., 2005; Campbell, Graham, & Lancet Maternal Survival Series Steering Group, 2006; Dogba & Fournier, 2009; Koblinsky et al., 2006; Paxton, Maine, Freedman, Fry, & Lobis, 2005; WHO, 2005, 2007a).

The major emphasis in the last few decades in LMICs in reducing maternal morbidity and mortality has been on decreasing the total number of

pregnancies per woman. As documented earlier, the total number of pregnancies per woman has fallen quite sharply in many of these countries; this drop has contributed significantly to the decrease in the lifetime risk of maternal mortality by reducing cumulative exposure. For example, assume a constant MMR of 750 maternal deaths per 100,000 live births. In Kenya, if the total fertility rate had remained at its 1975 level of 8.2 births per woman instead of dropping to the 2015 level of 3.9 births per woman, the lifetime risk of maternal mortality would have been 1/16 instead of 1/35. In addition to the significant drop in TFR, some fairly modest improvement in mortality risk per pregnancy has occurred; the 2015 MMR was 510 deaths per 100,000 births (WHO, 2015b). Taken together, these changes have led to a decrease in the lifetime maternal mortality risk to its current level of 1/51, with most of the improvement in lifetime risk coming from reduction of the overall number of births and, therefore, women's exposure to the risk of maternal mortality.

In addition to the reduction in the TFR in LMICs, some progress has been made in reducing the numbers of high-parity births (**TABLE 5-15**), births to older women (**TABLE 5-16**), and births following short intervals (i.e., less than 24 months). By 2000, however, as fertility continued to decline in most of these countries, the proportion of births to older women increased again due to postponement of births (United Nations Population Division, 2009). Access to safe abortion has also increased (Alan Guttmacher Institute, 2009a, 2009b; Tsui et al., 1997); unfortunately, the number of unsafe abortions remained more or less steady from 1995 to 2008 at approximately 20 million per year, with 98% of these procedures occurring in developing countries (Alan Guttmacher Institute, 2009a, 2009b; Sedgh et al., 2012; WHO, 2007b). An often-overlooked benefit of fertility-reduction initiatives is their impact on reducing the frequency of unsafe abortions by providing access to effective contraception

TABLE 5-15 Change in the Distribution of Birth Order Over the Course of Fertility Decline and Percentage Decline in Total Fertility Rates for Selected Countries

	Percentage of All Births of Order 1		Percentage of All Births of Order 5+		Percent Decline in Total Fertility Rate
Country	1960s	1970s–1980	1960s	1970s–1980	
Singapore	23	44	33	2	65
Hong Kong	25	43	23	4	64
Barbados	22	40	35	10	54
Mauritius	18	36	36	11	52
Costa Rica	18	32	45	17	50
Chile	25	41	31	9	49
Trinidad and Tobago	19	32	37	19	43
Puerto Rico	27	32	27	10	42
Panama	21	29	35	22	42
Malaysia	12	26	41	22	42
Fiji	23	35	36	13	41

TABLE 5-16 Change in the Distribution of Births by Maternal Age in Selected Countries Over the Course of Fertility Decline and Percent Decline in Total Fertility Rates Between 1960 and 1980

	Percentage of All Births to Women Younger Than Age 20		Percentage of All Births to Women Age 35 and Older		Percent Decline in Total Fertility Rate
Country	1960s	1970s–1980	1960s	1970s–1980	
Singapore	8	4	14	5	65
Hong Kong	5	4	20	6	64
Barbados	21	25	15	6	54
Mauritius	13	14	15	7	52
Costa Rica	13	20	18	9	50
Chile	12	17	17	9	49
Trinidad and Tobago	17	19	11	8	43
Puerto Rico	18	18	11	7	42
Panama	18	20	11	9	42
Malaysia	11	7	14	13	42
Fiji	13	11	12	8	41

Reproduced from United Nations Demographic Yearbook, various years. Printed in Contraception and Reproduction: Health Consequences for Women and Children in the Developing World. Copyright 1989 by the National Academy of Sciences.

for those who want it (Alan Guttmacher Institute, 2009a, 2009b; Bulatao, 1998; National Research Council, 1989; WHO, 2007b).

Despite these positive efforts, a significant burden of maternal mortality persists in LMICs. Two reasons explain why this condition remains in spite of the increase in the proportion of pregnancies that are in the demographically low-risk category (i.e., parity 2–4 among mothers age 24 to 29 with birth intervals of 24 months or more). First, the absolute risk of maternal death is still very high in these settings, even in low-risk pregnancies. Second, some of the factors associated with the highest-risk pregnancies, such as first pregnancies, are unavoidable.

In this regard, it is worth considering the following scenario. Kenya—which has very high risks of maternal mortality per pregnancy—experiences little change in the mortality risk per pregnancy (now 510 maternal deaths per 100,000 live births), but fertility declines to replacement levels (i.e., each woman has just 2.1 births over her lifetime). Under these conditions, the lifetime risk of maternal mortality would still be 1/94—almost 52 times as high as the lifetime maternal mortality risk (1/4,900) of women in high-income countries (WHO, 2015b).

Even this scenario may be overly optimistic. Our calculation does not allow for any heterogeneity in maternal mortality risks across pregnancies. If first pregnancies are intrinsically more dangerous or if healthier women were the ones who had more pregnancies, simply reducing the number of births would not achieve the reduction in maternal mortality estimated previously.

Finally, a number of studies have shown that older women with more surviving sons have significantly lower mortality than their peers with fewer

sons (Rahman, 1999; Rahman, Foster, & Menken, 1992). This research suggests that although increased numbers of pregnancies may expose a woman to considerable risk of morbidity and mortality, women in certain social settings (where family support is crucial) are actually better off when they are older if they have had higher fertility. These considerations lead to the conclusion that, to reduce maternal deaths to a significant extent in LMICs, an emphasis on reducing overall fertility levels and the frequency of high-risk pregnancies is unlikely to be sufficient. Instead, the emphasis must be on reducing the mortality risk for each and every pregnancy.

Obstetric Care

As pointed out earlier, the majority of maternal and perinatal mortality and morbidity stems from complications related to the delivery process, which are difficult to predict and avoid prenatally (Maine, 1991; National Statistics Office [Philippines] & Macro International, 1993; Ronsmans et al., 2006; Thaddeus & Maine, 1994; WHO, 2005). Perhaps the most important benefit of prenatal care is in sensitizing the mother to warning signs of obstetric emergencies and the need to seek appropriate obstetric care when they occur (Tsui et al., 1997).

It is instructive to review briefly what qualifies as the basic acceptable package of obstetric care. WHO uses the following eight criteria to assess the adequacy of obstetric care facilities (Campbell et al., 2006; Dogba & Fournier, 2009; Paxton et al., 2005; WHO, 1995):

1. The ability to treat infection, both orally and intravenously
2. The ability to provide intravenous labor-inducing agents such as oxytocin
3. The facilities for the medical treatment of shock, anemia, and hypertensive disorders of pregnancy (provision of parenteral anticonvulsants)
4. The ability to provide manual procedures for removal of the placenta, including vacuum extraction
5. The ability to carry out the removal of retained products of conception and ectopic pregnancy and provide safe abortion services
6. The ability to carry out assisted vaginal delivery and monitor labor
7. The ability to carry out caesarean section along with appropriate facilities for anesthesia
8. The ability to provide blood transfusions

Facilities that can carry out the first six interventions are called basic EmOC (Emergency Obstetric Care) centers; those that can carry out all eight interventions are termed complete EmOC centers (Campbell et al., 2006; Paxton et al., 2005).

The historical record in the United States and Europe (Hogberg, Wall, & Brostrom, 1986; Loudon, 1991) and the more recent experience of specific LMICs such as Sri Lanka show that the implementation of even limited parts of the essential obstetric care package can result in major declines in maternal mortality. The maternal mortality ratio in Sri Lanka dropped dramatically, from 555 maternal deaths per 100,000 births in the mid-1950s to 239 maternal deaths in the 1960s and again to 95 maternal deaths in 1980 (WHO, 2005). These gains were largely due to the expansion of health centers appropriately equipped for essential obstetric care and the increase in births attended by trained personnel.

A 2008 report focusing on maternal health interventions in 68 countries that experience 97% of maternal and child deaths worldwide suggested that several interventions have been proven to improve maternal survival (Bhutta et al., 2008; Bryce et al., 2008; United Nations Children's Fund, 2008). Under the circumstances when complications are not predictable, skilled care at delivery is the key to reduce maternal mortality. For example, a package of healthcare services consisting of skilled care by doctors, nurses, and midwives during pregnancy and childbirth, including family planning and emergency medical services, and costing less than $1.50 per person, can reduce an estimated 80% of maternal mortality. Increased access to hospital and midwifery care, improved quality of care, and control of infectious diseases have been successfully combined to significantly lower maternal mortality in several low-and middle-income countries, including Bangladesh, Nepal, Thailand, Malaysia, Sri Lanka, Egypt, Honduras, and some states of India (WHO, 2005).

For a variety of reasons—including cultural taboos, lack of social mobility, lack of economic resources, lack of logistic resources, and lack of information—women in LMICs for the most part do not use appropriate obstetric care services, with the most vulnerable group in this regard being rural, less educated, and poorer women (Govindasamy, Stewart, Rutstein, Boerma, & Sommerfelt, 1993). In 1993, only 37% of births in LMICs took place in a health facility (WHO, 1993a). The remaining deliveries (some 55 to 60 million infants) were carried out with the help of traditional birth attendants, family members, or no assistance at all.

In this regard, it is interesting to note that prenatal care services are used significantly more than obstetric care services. In 39 of 43 countries covered by DHS surveys between 1985 and 1994, prenatal care coverage was significantly higher than the rates of delivery care from a trained provider (i.e., doctor, nurse, or midwife) (Macro International, 1994). These results need to be viewed with some caution, however, as prenatal care is self-reported and may include very low and episodic use of appropriate services.

There is considerable variation in the DHS-surveyed countries in use of both prenatal and maternal care services. South Asia (with the exception of Sri Lanka) has particularly low rates of use of prenatal care and maternal care services. More recent data show significant improvements in developing countries in terms of access to both antenatal and skilled care services during delivery, with the latter showing a greater increase in their utilization rate. Approximately 66% of pregnant women in LMICs currently receive at least one antenatal care visit, compared to at least four antenatal care visits for most women in high-income countries. Huge disparities in the distribution of antenatal care are apparent, however: In Ethiopia, for example, this figure is only 12%. Moreover, only 62% of pregnant women in LMICS currently receive assistance from skilled birth attendants during their deliveries. As in the case of antenatal care, large regional differences in access to these services exist, with utilization rates ranging from 34% in Eastern Africa to 93% in South America (Maternal Deaths, 2008; United Nations Children's Fund, 2008; WHO, 2008b).

Despite these improvements in access to skilled delivery care, there is currently a 50% deficit worldwide against the estimated demand for 700,000 midwives to ensure universal coverage with maternity care. More generally, there is a global shortage of 4.3 million health workers (WHO, 2008b).

Leaving aside the significant gaps in the availability of skilled maternal care/delivery services, even when such services are available they are not used to the degree they should be. The failure to use appropriate maternal/delivery care services can be viewed as resulting from a multipart process, which includes the following factors:

- Deficiencies in identifying life-threatening medical complications that would benefit from obstetric care services
- Constraints that prevent the use of obstetric services even when appropriate conditions for such use are identified (e.g., lack of financial resources, transportation difficulties, other logistic problems)
- Obstetric care facilities of such poor quality that they are ineffective in preventing obstetric complications and death, even when pregnant women come to such facilities (Thaddeus & Maine, 1994; Tsui et al., 1997)

Each of these problems is elaborated upon in the following subsections.

Identification of Serious Medical Complications That Will Benefit from Obstetric Care

It is important to reiterate that most women in LMICs deliver their babies at home, far away from even rudimentary obstetric care facilities, and assisted for the most part by family members and/or traditional birth attendants. The ability to recognize a potentially life-threatening complication (such as obstructed or prolonged labor, incipient hemorrhage, or other fetal distress), and the corresponding need for specialized obstetric care, depends on appropriate education and sensitization of both family members and traditional birth attendants, with the former being perhaps the more important constituency.

A number of cultural factors affect the recognition that obstetric complications can benefit from specific kinds of medical care. When a woman is bleeding, the need for obstetric intervention by trained medical personnel is better recognized. Other complications, however, are generally not viewed as benefiting from specialized obstetric care. In some cultures, complications are seen as determined by fate, with little that can be done to alter the course of events. In Indonesia, for example, malposition of the fetus is seen as the domain of traditional birth attendants, who deal with it with soothing massage whereby the pregnant woman maintains the inner calm considered necessary for correcting the baby's position (Ambaretnani, Hessler-Radelet, & Carlin, 1993). As a result, referral to obstetric care services is often significantly delayed. In other instances, due to concerns about privacy and modesty, the home is perceived as the more natural and fitting place for birth than any healthcare facility. For the Fulani and Hausa in Nigeria, pregnancy is seen as a shameful period with an unpredictable outcome; thus no preparations are made for referral for obstetric services (Public Opinion Polls, 1993; Tsui et al., 1997).

It is clear from these examples that cultural sensitivities need to be taken into account when constructing appropriate educational messages that emphasize the identification of particularly serious obstetric risks and the need to refer these women to health facilities with adequate obstetric care services. A purely technocratic approach, which focuses on modern systems of

logic and evidence, may not be very effective in bringing about changes in behavior. Much attention has been focused on training birth attendants, given that they are present in the majority of births in LMICs. Nevertheless, the evidence suggests that focusing on just training birth attendants leads to mixed results in terms of increased referrals for appropriate conditions. In urban/peri-urban areas, such training appears to be no impact, whereas there is some improvement in rural areas with low prevalence rates of use of maternal care services (Alisjahbana et al., 1995; Bailey, Dominik, Janowitz, & Araujo, 1991). The consensus from a number of sources is that husbands and possibly other family members (mothers-in-law, mothers, and sisters) are the key people who should be targeted, as they are the final decision makers with regard to whether the woman will go to a modern healthcare facility (Alisjahbana et al., 1995; Bailey et al., 1991; Center for Health Research Consultation and Education & MotherCare/John Snow, 1991; Howard-Grabman, Seoane, & Davenport, 1994).

Constraints Affecting Use of Obstetric Care Services Once Complications Are Identified

Several concerns affect the decision to use modern obstetric care services once a life-threatening complication is identified and the need for obstetric care is acknowledged. These issues include economic constraints (i.e., not being able to afford the costs of such care, including transportation costs), logistic constraints (taking time off to accompany the patient to often distant care facilities), and quality concerns (attitude and treatment of healthcare providers) (Sundari, 1992; Thaddeus & Maine, 1994).

The economic costs of care include not only the nominal costs charged by the healthcare facility for delivery-related services, but also transportation costs (which are not insignificant in many rural settings), costs of medications, and costs of housing. In many cultures, family members need to accompany the patient, which adds to transportation and housing costs. In addition to these direct costs, opportunity costs or lost wages for the patient and particularly family members who accompany the patient are key considerations. Relatively few population-based data have been gathered from which to estimate the various components of costs that are incurred for a maternal delivery in a healthcare facility, but some reports suggest that these costs may be a real barrier to the use of obstetric services. For example, data from three countries in Africa (Nigeria, Ghana, and Sierra Leone) show that there were declines in deliveries in seven referral sites from 1983 to 1989, paralleling increases in costs to patients for drugs and services (Prevention of Maternal Mortality Network, 1995). To date, the impact of user fees on the decision to seek out obstetric care for complicated cases (as opposed to normal deliveries) has not been adequately investigated. The existing evidence shows a mixed response to the imposition of user fees, with use of modern medical facilities for obstetric complications being reduced to different degrees in different countries (Ambaretnani et al., 1993; Prevention of Maternal Mortality Network, 1995).

Clearly, much more information needs to be collected to understand how demand and use of obstetric care services change with changes in costs of services. Special attention needs to be focused on obtaining data on the nonservice components of obstetric costs (i.e., travel costs, opportunity costs).

Aside from economic costs, transportation constraints are a major factor in the low use of maternal care services. In most LMICs, advanced obstetric care (including surgical services with appropriate transfusion capabilities) is available in only a few healthcare facilities, which are often located at a considerable distance from the patient. Transportation facilities are, for the most part, poorly developed and quite expensive. Thus, problems need to be anticipated in advance to allow for enough time for the patient to reach the care facility. The degree to which lack of transportation infrastructure is a major constraint on the use of modern obstetric services is not certain; the evidence on this issue is mixed. On the one hand, data from rural Bangladesh suggest that relatively modest investments in transportation can have a significant impact on use of obstetric services, leading to reductions in maternal mortality (Fauveau, Stewart, Khan, & Chakraborty, 1991; Maine, Akalin, Chakraborty, de Francisco, & Strong, 1996). On the other hand, three different experiments that aimed at ensuring transport did not by themselves increase the use of obstetric services (Alisjahbana et al., 1995; Poedje et al., 1993; Prevention of Maternal Mortality Network, 1995).

Transportation concerns assume the patient has to be brought to the care facility. A complementary approach, of course, is to bring the provider closer to the patient. Some limited success has been achieved with programs that place certified midwives in health posts closer to the pregnant patient population, but such staffing may be difficult to sustain logistically and financially (Fauveau et al., 1991; Tsui et al., 1997).

Finally, the perception of the welcoming nature of the referral site and its flexibility in accommodating accompanying family members are often ignored, but are particularly important constraints to

seeking modern obstetric care. Oftentimes, referral sites are perceived as impersonal and unfriendly and are passed over in favor of care from traditional birth attendants (Bailey et al., 1994; Eades, Brace, Osei, & LaGuardia, 1993).

Quality of the Obstetric Care in the Healthcare Facility

Relatively few systematic data on the quality of healthcare services in obstetric care facilities exist. The rare studies that have been conducted suggest that the majority of obstetric care facilities in LMICs fall far short of minimal acceptable standards of care. Important indicators of quality include waiting time from admission to treatment, trends in numbers and rates of maternal and perinatal deaths, and trends in case fatality rates for all complications including cesarean deliveries (O'Rourke, 1995; Prevention of Maternal Mortality Network, 1995).

Much of this problem can be traced to lack of adequate resources in terms of trained personnel (Dogba & Fournier, 2009), equipment, and bed capacity. For example, a UNICEF survey of three districts in India in 1993 found not only inadequate numbers of beds, but also huge disparities in bed allocation between different levels of the healthcare system. The majority of beds were allocated to referral sites, where a small minority of complicated births were managed. There were also major deficiencies in availability of essential drugs and appropriately trained surgical and anesthesia professionals.

In addition to supply constraints, major deficiencies in the management of services and in provider attitudes are often in evidence in obstetric facilities located in LMICs. Triage is not done systematically, such that very sick patients are often left waiting for much longer than medically desirable, while others with less severe problems are treated before them. Obstetric care is often provided in an ad hoc manner, with no consistent set of case management algorithms being followed. Nursing is often seriously below standard, and basic levels of hygiene are not adhered to, leading to considerable postoperative morbidity and mortality. There is also very little sensitivity to patient concerns, and little effort is expended in explaining complexities of care to patients, leading to widespread patient dissatisfaction (Dogba & Fournier, 2009).

In a nationally representative study of 718 maternal deaths in Egypt in 1992 (Egypt Ministry of Health, 1994), avoidable factors (i.e., those that could have been changed by either the health delivery system or the patient) were assessed by an expert panel. In approximately half of the cases, the primary avoidable factor identified was poor management and diagnoses by healthcare professionals. In the other half of cases, patient factors—particularly delay in seeking medical care or compliance with medical recommendations—were implicated as key issues in the mother's death. Notably, the health professionals most often cited for poor quality of care were not traditional birth attendants or general practitioners, but rather obstetricians with supposedly appropriate training. This sorry state of affairs is significantly related to the lack of consistent management guidelines for complicated obstetric cases.

Similar results have been reported from China (WHO, 1994), where a study of 1,173 maternal deaths in 1990 implicated deficiencies in the healthcare system as the major contributor (48%), followed by individual and family delays in using health care and transport problems. There appears to be a clear rural/urban divide in this situation, with rural areas having a much higher frequency of problems that are avoidable both from the point of view of the healthcare system and from the point of view of the patient.

Improvements in the quality of care require a number of initiatives to be undertaken simultaneously. Governments must appropriately fund healthcare referral facilities so that they have adequate supplies and equipment and are staffed by adequate numbers of appropriate specialists. There must be a clear chain of referral whereby trained birth attendants refer complicated obstetric cases to higher-level facilities, where specialist care is available. Efforts must be made to follow consistent management protocols that are clearly articulated, and both birth attendants and specialists must be trained to adhere to these protocols (Marshall & Buffington, 1991; Schieber, Mejia, Koritz, Gonsalez, & Kwast, 1995). In addition, a monitoring system must be established to provide regular audits of both process and outcome indicators such as waiting times and case fatalities, with the results being used in a continuous process of review and upgrading of the healthcare system (Egypt Ministry of Health, 1994).

If even some of these improvements in quality of care are made, they will not only improve outcomes for those women who reach the healthcare facility, but also increase the demand for such services by pregnant women (Dogba & Fournier, 2009; Mantz & Okong, 1994; O'Rourke, 1995).

Conclusion

This chapter has outlined the need for reproductive health and family planning to help individual women and men and larger populations reduce fertility and maintain reproductive health. Population growth rates are declining in much of the world because people are reducing their fertility. The primary factors effecting this reduction are early termination of the reproductive period through sterilization, use of contraception to reduce conception rates, and induced abortion. Because of increased desire for smaller families, both unwanted fertility and unmet needs for family planning and reproductive health services exist in most LMICs. To meet this challenge, improvements are needed in the quality of family planning services, especially in the areas of information exchange and method choice, interaction of reproductive health services and contraceptive provision, and financial sustainability.

Maternity care needs to be significantly expanded so that the adverse sequelae of pregnancy and childbearing can be reduced. While preventive services (including education of both men and women regarding health and sexuality, family planning, and prevention of STIs) need to be increased and targeted to those at greatest risk of adverse outcomes, special attention should be paid to increasing both physical and financial access to skilled birth attendants and high-quality emergency and non-emergency obstetric care.

At the societal level, programs need to be supported to improve the status of women, whether through improved access to education and employment or through changes in laws and culture to reduce violence and exposure to unwanted potentially high-risk births. Although this may be unfamiliar territory for public health professionals, the consequences for the health of women and children make it essential that this broader perspective become part of global health programs.

An overriding concern regarding continuing—let alone expanded—funding remains. Many high-income countries have reduced their aid contributions; many LMICs are experiencing financial and health crises. Part of the agenda for the future must be research to determine sustainable cost-effective programs that will address the reproductive health needs of the twenty-first century using public–private partnerships.

Discussion Questions

1. Using the proximate determinants framework (originally proposed by Bongaarts and subsequently revised by Stover), discuss the relative impacts of contraception, breastfeeding, abortion, sexual activity, and infecundity on total fertility rates in Latin America versus Africa. What are the policy implications of these findings? Similarly, explain why Japan, which has a relatively low contraceptive prevalence rate (56%), has such a low TFR (1.4).
2. Discuss fertility changes using the framework of wanted versus unwanted fertility and the forces that drive each of them. Why might unwanted fertility increase in tandem with increasing contraceptive prevalence? What are the policy approaches that stem from a consideration of this kind of framework?
3. The 1994 Cairo ICPD substantially enlarged the scope of family planning to include a broader conception of women's health and development. Discuss the pros and cons of this expansion in the context of limited financial and logistic resources, particularly in sub-Saharan Africa, with its very heavy HIV/AIDS burden.
4. Discuss the impact of the birth interval length, both prior and subsequent, on the health of children, taking into account methodological concerns about reverse causality. What are the implications for policy?
5. Consider the following statement: "Family planning has only a limited role to play in reducing the risk of maternal mortality." Do you agree with this statement? Elaborate on the policy implications of your analysis.
6. Discuss the reasons why, in the context of LMICs, despite initial dramatic success in lowering maternal mortality as per MDG goals, there has been a recent plateauing of these gains. What are the major constraints faced by countries and the policy implications that follow from this analysis?
7. Consider the following statement: "Specific health technological inputs are a necessary but not sufficient determinant of significant improvements in reproductive health." Using the examples of changes in women's status and sociopolitical resistance to raising age at marriage, discuss the validity of this proposition.

References

Aaby, P., Bukh, J., Lisse, I. M., & Smits, A. J. (1984). Overcrowding and intensive exposure as determinants of measles mortality. *American Journal of Epidemiology, 120*(1), 49–63.

Adler, A. J., Ronsmans, C., Calvert, C., & Filippi, V. (2013). Estimating the prevalence of obstetric fistula: a=A systematic review and meta-analysis. *BMC Pregnancy and Childbirth, 13*, 246.

Aitken, I. W., & Walls, B. (1986). Maternal height and cephalopelvic disproportion in Sierra Leone. *Tropical Doctor, 16*(3), 132–134.

Alan Guttmacher Institute. (2009a). *Abortion worldwide: A decade of uneven progress*. New York, NY: Author.

Alan Guttmacher Institute. (2009b). Facts on induced abortion worldwide. Retrieved from https://www.guttmacher.org/united-states/abortion?scope=comparative%20international

Alan Guttmacher Institute. (2016). Facts on induced abortion worldwide. Retrieved from https://www.guttmacher.org/united-states/abortion?scope=comparative%20international

Alauddin, M. (1986). Maternal mortality in rural Bangladesh: The Tangail District. *Studies in Family Planning, 17*(1), 13–21.

Alisjahbana, A., Williams, C., Dharmayanti, R., Hermawan, D., Kwast, B. E., & Koblinsky, M. (1995). An integrated village maternity service to improve referral patterns in a rural area in West-Java. *International Journal of Gynaecology and Obstetrics, 48*(suppl), S83–S94.

Ambaretnani, N. P., Hessler-Radelet, C., & Carlin, L. E. (1993). *Qualitative research for the social marketing component of the Perinatal Regionalization Project, Tanjungasari, Java* (MotherCare Working Paper No. 19, prepared for the U.S. Agency for International Development, Project No. 936-5966). Arlington, VA: John Snow.

Anonymous. (1996). Japan's TFR plunges to record low of 1.43. *JOICFP News, 267*, 1–2.

Arnold, F., Kishor, S., & Roy, T. (2002). Sex-selective abortions in India. *Population and Development Review, 28*(4), 759–785.

Bailey, P., Szászdi, J., & Schieber, B. (1994). *Analysis of the vital events reporting system of the Maternal and Neonatal Health Project, Quetzaltenango, Guatemala*. Arlington, VA: John Snow.

Bailey, P. E., Dominik, R. C., Janowitz, B., & Araujo, H. (1991). Assistência obstétrica e mortalidade perinatal em uma área rural do Nordeste Brasileiro. *Boletin de la Oficina Sanitaria Panamericana, 111*(4), 306–318.

Bangladesh Demographic and Health Survey (BDHS). (2016). *Bangladesh Demographic and Health Survey 2014*. Dhaka, Bangladesh/Calverton, MD: National Institute of Population Research and Training (NIPORT), Mitra and Associates, & Macro International.

Bangladesh Planning Commission. (2015). *Millennium Development Goals: Bangladesh progress report—2015*. Dhaka, Bangladesh: Author.

Bankole, A., & Westoff, C. F. (1995). *Childbearing attitudes and* intentions (DHS Comparative Studies No. 17). Calverton, MD: Macro International.

Bartlett, L. A., Mawji, S., Whitehead, S., Crouse, C., Dalil, S., Ionete, D., . . . Afghan Maternal Mortality Study Team. (2005). Where giving birth is a forecast of death: Maternal mortality in four districts of Afghanistan, 1999–2002. *Lancet, 365*(9462), 864–870.

BBC. (2017, April 4). US withdraws funding for United Nations Population Fund. *BBC News*. Retrieved from http://www.bbc.com

Beck, S., Wojdyla, D., Say, L., Betran, A. P., Merialdi, M., Requejo, J. H., . . . Van Look, P. F. (2010). The worldwide incidence of preterm birth: A systematic review of maternal mortality and morbidity. *Bulletin of the World Health Organization, 88*(1), 31–38.

Benson, J., Andersen, K., & Samandari, G. (2011). Reductions in abortion-related mortality following policy reform: Evidence from Romania, South Africa and Bangladesh. *Reproductive Health, 8*(1), 39.

Bhutta, Z. A., Ahmed, T., Black, R. E., Cousens, S., Dewey, K., Giugliani, E., . . . Maternal and Child Undernutrition Study, Group (2008). What works? Interventions for maternal and child undernutrition and survival. *Lancet, 371*(9610), 417–440.

Black, R. E., Morris, S. S., & Bryce, J. (2003). Where and why are 10 million children dying every year? *Lancet, 361*(9376), 2226–2234.

Blanc, A. K., Winfrey, W., & Ross, J. (2013). New findings for maternal mortality age patterns: Aggregated results for 38 countries. *PLoS One, 8*(4), e59864.

Bledsoe, C. H., & Cohen, B. (1993). *Social dynamics of adolescent fertility in sub-Saharan Africa. Working Group on the Social Dynamics of Adolescent Fertility in Sub-Saharan Africa, Committee on Population, National Research Council*. Washington, DC: National Academies Press.

Blencowe, H., Cousens, S., Chou, D., Oestergaard, M., Say, L., Moller, A., . . . Lawn, J. (2013). Born too soon: The global epidemiology of 15 million preterm births. *Reproductive Health, 10*(1), S2.

Blencowe, H., Cousens, S., Jassir, F. B., Say, L., Chou, D., Mathers, C., . . . Lancet Stillbirth Epidemiology Investigator, Group. (2016). National, regional, and worldwide estimates of stillbirth rates in 2015, with trends from 2000: A systematic analysis. *Lancet Global Health, 4*(2), e98–e108.

Blencowe, H., Cousens, S., Oestergaard, M. Z., Chou, D., Moller, A. B., Narwal, R., . . . Lawn, J. E. (2012). National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: A systematic analysis and implications. *Lancet, 379*(9832), 2162–2172.

Bongaarts, J. (1978). A framework for analyzing the proximate determinants of fertility. *Population and Development Review*, 105–132.

Bongaarts, J. (1983). The proximate determinants of natural marital fertility. In R. D. BulaLee & R. D. Lee (Eds.), *Determinants of fertility in developing countries* (Vol. 1, pp. 103–108). New York, NY: Academic Press.

Bongaarts, J. (1991). Do reproductive intentions matter? *Demographic and Health Surveys World Conference, 1*, 223–248.

Bongaarts, J. (1997). The role of family planning programmes in contemporary fertility transitions. In G. W. Jones, R. M. Douglas, J. C. Caldwell & R. M. D'Souza (Eds.), *The continuing demographic transition* (pp. 422–443). New York, NY/Oxford, UK: Oxford University Press.

Bongaarts, J., & Bruce, J. (1995). The causes of unmet need for contraception and the social content of services. *Studies in Family Planning, 26*(2), 57–75.

Bongaarts, J., & Johansson, E. (2002). Future trends in contraceptive prevalence and method mix in the developing world. *Studies in Family Planning, 33*(1), 24–36.

Bongaarts, J., & Potter, R. E. (1983). *Fertility, biology, and behavior: An analysis of the proximate determinants*. New York, NY: Academic Press.

Bongaarts, J., & Watkins, S. C. (1996). Social interactions and contemporary fertility transitions. *Population and Development Review, 22*(4), 639–682.

Bothwell, T. H., & Charlton, R. (1981). *Iron deficiency in women*. Washington, DC: International Nutrition Anemia Consultative Group.

Bradley, E., Croft, T., & Rutstein, S. O. (2011). *The impact of contraceptive failure on unintended births and induced abortions: Estimates and strategies for reduction*. Calverton, MD: ICF Macro.

Brown, S. S., & Eisenberg, L. (1995). *The best intentions: Unintended pregnancy and the well-being of children and families*: Washington, DC: National Academies Press.

Bryce, J., Coitinho, D., Darnton-Hill, I., Pelletier, D., Pinstrup-Andersen, P., & Maternal and Child Undernutrition Study, Group. (2008). Maternal and child undernutrition: Effective action at national level. *Lancet, 371*(9611), 510–526.

Bulatao, R. A., & Elwan, A. (1985). *Fertility and mortality transition: Patterns, projections, and interdependence* (World Bank Staff Working Paper No. 681). Washington, DC: World Bank.

Bulatao, R. A. (1993). *Effectiveness and evolution in family planning programs.* Paper presented at the International Union for the Scientific Study of Population, International Population Conference, Liege, Belgium.

Bulatao, R. A. (1998). *The value of family planning programs in developing countries*. Santa Monica, CA: RAND Corporation.

Bulatao, R. A., & Lee, R. (1983). *Determinants of fertility in developing countries* (Vol. I). New York, NY: Academic Press.

Caldwell, J. C. (1986). Routes to low mortality in poor countries. *Population and Development Review, 12*(2), 171–220.

Caldwell, J. C., Barkat, E. K., Caldwell, B., Pieris, I., & Caldwell, P. (1999). The Bangladesh fertility decline: An interpretation. *Population and Development Review, 25*(1), 67–84.

Campbell, O. M., Graham, W. J., & Lancet Maternal Survival Series Steering Group. (2006). Strategies for reducing maternal mortality: Tetting on with what works. *Lancet, 368*(9543), 1284–1299.

Casterline, J. B., Perez, A. E., & Biddlecom, A. E. (1996). *Factors underlying unmet need for family planning in the Philippines* (Research Division Working Paper No. 84). New York, NY: Population Council.

Center for Health Research Consultation and Education & MotherCare/John Snow. (1991). *Qualitative research on knowledge, attitudes, and practices related to women's reproductive health* (MotherCare Working Paper No. 9, prepared for the U.S. Agency for International Development, Project No. 936-5966). Arlington, VA: John Snow.

Center for Reproductive Rights. (2013). Fact sheet: The world's abortion laws map 2013 update. Retrieved from https://www.reproductiverights.org/sites/crr.civicactions.net/files/documents/AbortionMap_Factsheet_2013.pdf

Center for Reproductive Rights. (2014). Abortion worldwide: 20 years of reform. Retrieved from https://www.reproductiverights.org/document/abortion-worldwide-20-years-of-reform

Chen, L. C., Gesche, M. C., Ahmed, S., Chowdhury, A. I., & Mosley, W. H. (1974). Maternal mortality in rural Bangladesh. *Studies in Family Planning, 5*(11), 334–341.

Chi, I. C., Agoestina, T., & Harbin. (1981). Maternal mortality at twelve teaching hospitals in Indonesia: An epidemiologic analysis. *International Journal of Gynecology & Obstetrics, 19*(4), 259–266.

Cleland, J., Phillips, J. F., Amin, S., & Kamal, G. M. (1994). *The determinants of reproductive change in Bangladesh*. Washington, DC: World Bank.

Coale, A. J., & Watkins, S. C. (1986). *The decline of fertility in Europe*. Princeton, NJ: Princeton University Press.

Colombo, B., & Masarotto, G. (2000). Daily fecundability: First results from a new data base. *Demographic Research, 3*, 39.

Curtis, S. L., & Neitzel, K. (1996). *Contraceptive knowledge, use, and sources* (DHS Comparative Studies No. 19). Columbia, MD: Institute for Resource Development.

Darroch, J. E., & Singh, S. (2013). Trends in contraceptive need and use in developing countries in 2003, 2008, and 2012: An analysis of national surveys. *Lancet, 381*(9879), 1756–1762.

Das Gupta, M. (1987). Selective discrimination against female children in rural Punjab, India. *Population and Development Review, 13*(1), 77–100.

DaVanzo, J., Butz, W. P., & Habicht, J. P. (1983). How biological and behavioural influences on mortality in Malaysia vary during the first year of life. *Population Studies, 37*(3), 381–402.

David, H. P. (1992). Abortion in Europe, 1920–91: A public health perspective. *Studies in Family Planning, 23*(1), 1–22.

de Bernis, L., Kinney, M. V., Stones, W., Ten Hoope-Bender, P., Vivio, D., Leisher, S. H., . . . Lancet Ending Preventable Stillbirths Series Advisory, Group. (2016). Stillbirths: Ending preventable deaths by 2030. *Lancet, 387*(10019), 703–716.

De Graft-Johnson, J. (1994). *Maternal morbidity in Ghana*. Paper presented at the Annual Meeting of the Population Association of America, Miami, FL.

Desai, S. (1995). When are children from large families disadvantaged? Evidence from cross-national analyses. *Population Studies, 49*(2), 195–210.

Dhaka Tribune. (2017, February 28). Child Marriage Restraint Act 2017 to practise no restraint. p. 3.

Dixon-Mueller, R., & Germain, A. (1992). Stalking the elusive "unmet need" for family planning. *Studies in Family Planning, 23*(5), 330–335.

Dogba, M., & Fournier, P. (2009). Human resources and the quality of emergency obstetric care in developing countries: A systematic review of the literature. *Human Resources for Health, 7*(1), 7.

D'Souza, S., & Chen, L. C. (1980). Sex differentials in mortality in rural Bangladesh. *Population and Development Review, 6*(2), 257–270.

Duffy, L., & Menken, J. (1998). *Health, fertility, and socioeconomic status as predictors of survival and later health of women: A twenty year prospective study in rural Bangladesh* (Working Paper WP-98-11). Boulder, CO: Population Program, Institute of Behavioral Science, University of Colorado.

Eades, C. A., Brace, C., Osei, L., & LaGuardia, K. D. (1993). Traditional birth attendants and maternal mortality in Ghana. *Social Science & Medicine, 36*(11), 1503–1507.

Eclampsia Trial Collaborative Group. (1995). Which anticonvulsant for women with eclampsia? Evidence from the Collaborative Eclampsia Trial. *Lancet, 345*(8963), 1455–1463.

Egypt Ministry of Health. (1994). *National maternal mortality study: Egypt, 1992–1993*. Alexandria, Egypt: Ministry of Health, Child Survival Project.

El Arifeen, S., Hill, K., Ahsan, K. Z., Jamil, K., Nahar, Q., & Streatfield, P. K. (2014). Maternal mortality in Bangladesh: A countdown to 2015 country case study. *Lancet, 384*(9951), 1366–1374.

Faundes, A., Fanjul, B., Henriquez, G., Mora, G., & Tognola, C. (1974). Influencia de la edad y de la paridad sobre algunos parámetros de morbilidad materna y sobre la morbi-mortalidad fetal. *Revista Chilena de Obstetricia y Ginecología, 37*(1), 6–14.

Fauveau, V., Koenig, M. A., Chakraborty, J., & Chowdhury, A. I. (1988). Causes of maternal mortality in rural Bangladesh, 1976–85. *Bulletin of the World Health Organization, 66*(5), 643–651.

Fauveau, V., Stewart, K., Khan, S. A., & Chakraborty, J. (1991). Effect on mortality of community-based maternity-care programme in rural Bangladesh. *Lancet, 338*(8776), 1183–1186.

Fauveau, V., Wojtyniak, B., Mostafa, G., Sarder, A. M., & Chakraborty, J. (1990). Perinatal mortality in Matlab, Bangladesh: A community-based study. *International Journal of Epidemiology, 19*(3), 606–612.

Ferraz, E. M., Gray, R. H., Fleming, P. L., & Maia, M. (1988). Interpregnancy interval and low birth weight: Findings from a case-control study. *American Journal of Epidemiology, 128*(5), 1111–1116.

Filippi, V., Chou, D., Ronsmans, C., Graham, W., & Say, L. (2016). Levels and causes of maternal mortality and morbidity. In R. E. Black, R. Laxminarayan, M. Temmerman & N. Walker (Eds.), *Reproductive, maternal, newborn, and child health* (p. 51). Washigton DC: World Bank.

Filippi, V., Ronsmans, C., Campbell, O. M., Graham, W. J., Mills, A., Borghi, J., . . . Osrin, D. (2006). Maternal health in poor countries: The broader context and a call for action. *Lancet, 368*(9546), 1535–1541.

Finkle, J., & Ness, G. (1985). *Managing delivery systems: Identifying leverage points for improving family planning program performance*. Ann Arbor, MI: University of Michigan Department of Population Planning and International Health.

Fortney, J. A., Susanti, I., Gadalla, S., Saleh, S., Feldblum, P. J., & Potts, M. (1985). *Maternal mortality in Indonesia and Egypt*. Paper presented at the WHO Inter-regional Meeting on Prevention of Maternal Mortality.

Foster, A., Menken, J., Chowdhury, A., & Trussell, J. (1986). Female reproductive development: A hazards model analysis. *Social Biology, 33*(3–4), 183–198.

Foster, A., & Roy, N. (1996). *The dynamics of education and fertility: Evidence from a family planning experiment.* Working paper, University of Pennsylvania, Department of Economics.

Freedman, R. (1987). The contribution of social science research to population policy and family planning program effectiveness. *Studies in Family Planning, 18*(2), 57–82.

French, F. E., & Bierman, J. M. (1962). Probabilities of fetal mortality. *Public Health Reports, 77*, 835–847.

Frenzen, P. D., & Hogan, D. P. (1982). The impact of class, education, and health care on infant mortality in a developing society: The case of rural Thailand. *Demography, 19*(3), 391–408.

Gelbard, A., Haub, C., & Kent, M. M. (1999). World population beyond six billion. *Population Bulletin, 54*(1), 1–44.

Geronimus, A. T. (1987). On teenage childbearing and neonatal mortality in the United States. *Population and Development Review, 13*(2), 245–279.

Goldberg, C. Insurance for Viagra spurs coverage for birth control. (1999, June 30). *New York Times*. Retrieved from https://www.nytimes.com/1999/06/30/us/insurance-for-viagra-spurs-coverage-for-birth-control.html

Goldenberg, R. L., Culhane, J. F., Iams, J. D., & Romero, R. (2008). Epidemiology and causes of preterm birth. *Lancet, 371*(9606), 75–84.

Govindasamy, P., Stewart, K., Rutstein, S., Boerma, J., & Sommerfelt, A. (1993). *High-risk births and maternity care* (Demographic and Health Surveys Comparative Studies No. 8). Columbia, MD: Macro International.

Graham, W., Woodd, S., Byass, P., Filippi, V., Gon, G., Virgo, S., . . . Singh, S. (2016). Diversity and divergence: The dynamic burden of poor maternal health. *Lancet, 388*(10056), 2164–2175.

Gray, R. H., Wawer, M. J., Serwadda, D., Sewankambo, N., Li, C., Wabwire-Mangen, F., . . . McNairn, D. (1998). Population-based study of fertility in women with HIV-1 infection in Uganda. *Lancet, 351*(9096), 98–103.

Greenwood, A. M., Greenwood, B. M., Bradley, A. K., Williams, K., Shenton, F. C., Tulloch, S., . . . Oldfield, F. S. (1987). A prospective survey of the outcome of pregnancy in a rural area of the Gambia. *Bulletin of the World Health Organization, 65*(5), 635–643.

Gubhaju, B. B. (1986). Effect of birth spacing on infant and child mortality in rural Nepal. *Journal of Biosocial Science, 18*(4), 435–447.

Haaga, J. G. (1989). Mechanisms for the association of maternal age, parity, and birth spacing with infant health. In A. M. Parnell (Ed.), *Contraceptive use and controlled fertility: Health issues for women and children* (pp. 96–139). Washington, DC: National Academy Press.

Hansen, J. P. (1986). Older maternal age and pregnancy outcome: A review of the literature. *Obstetrical & Gynecological Survey, 41*(11), 726–742.

Harrison, K. A., & Rossiter, L. A. (1985). Child-bearing health and social priorities: A survey of 22774 consecutive hospital births in Zaria northern Nigeria. *British Journal of Obstetrics and Gynaecology, 5*(suppl), 1–119.

Harrison, P. F., & Rosenfield, A. (1996). *Contraceptive research and development: Looking to the future. Committee on Contraceptive Research and Development, Institute of Medicine*. Washington, DC: National Academies Press.

Hesketh, T., Lu, L., & Xing, Z. W. (2011). The consequences of son preference and sex-selective abortion in China and other Asian countries. *Canadian Medical Association Journal, 183*(12), 1374–1377.

High-Impact Practices in Family Planning (HIP). (2014). *Financing commodities and services: Mobilizing resources to sustain current and future family planning demand*. Washington, DC: USAID.

Hill, K., & Choi, Y. (2006). Neonatal mortality in the developing world. *Demographic Research, 14*, 429–452.

Hobcraft, J. (1987). *Does family planning save childrens lives?* Paper presented at the International Conference on Better Health for Women and Children Through Family Planning.

Hobcraft, J. N., McDonald, J. W., & Rutstein, S. O. (1985). Demographic determinants of infant and early child mortality: A comparative analysis. *Population Studies, 39*(3), 363–385.

Hogberg, U., Wall, S., & Brostrom, G. (1986). The impact of early medical technology on maternal mortality in late 19th century Sweden. *International Journal of Gynaecology and Obstetrics, 24*(4), 251–261.

Howard-Grabman, L., Seoane, L. G., & Davenport, C. A. (1994). *The Warmi Project: A participatory approach to improve maternal and neonatal health. An implementors manual*. Arlington, VA: MotherCare Project, John Snow.

International Consortium for Emergency Contraception (ICEC). (2010). Background. Retrieved from http://www.cecinfo.org

International Planned Parenthood Foundation. (1995). *Consensus statement on emergency contraception*. London, UK: Author.

Jamison, D. T., Mosley, W. H., Measham, A. R., & Bobadilla, J. L. (1993). *Disease control priorities in developing countries*. New York, NY: Oxford University Press.

Jelliffe, D. B. (1976). *Maternal nutrition and lactation.* Paper presented at the CIBA Foundation Symposium, Breastfeeding and the Mother, Amsterdam.

Jones, R. E. (1988). A biobehavioral model for breastfeeding effects on return to menses post-partum in Javanese women. *Social Biology, 35,* 307–323.

Katz, J., Lee, A. C., Kozuki, N., & Black, R. E. (2015). Mortality risk among term and preterm small for gestational age infants. *Nestlé Nutrition Institute Workshop Series, 81*, 29–35.

Kenney, G. M. (1993). *Assessing legal and regulatory reform in family planning: Manual on legal and regulatory reform* (OPTIONS Projects Policy Paper No. 1). Washington, DC: Futures Group.

Kharsany, A. B., & Karim, Q. A. (2016). HIV Infection and AIDS in sub-Saharan Africa: Current status, challenges and opportunities. *Open AIDS Journal, 10*, 34–48.

Kitay, D. Z., & Harbort, R. A. (1975). Iron and folic acid deficiency in pregnancy. *Clinics in Perinatology, 2*(2), 255–273.

Knodel, J., Chamratrithirong, A., & Debavalya, N. (1987). Societal change and the demand for children. In *Thailand's reproductive revolution: Rapid fertility decline in a Third World setting* (pp. 117–154). Madison, WI: University of Wisconsin Press.

Koblinsky, M. A., Campbell, O., & Harlow, S. D. (1993). Mother and more: A broader perspective on womens health. In M. A. Koblinsky, J. Timyan & J. Gay (Eds.), *The health of women: A global perspective* (pp. 33-62). Boulder CO: Westview Press.

Koblinsky, M., Matthews, Z., Hussein, J., Mavalankar, D., Mridha, M. K., Anwar, I., . . . Lancet Maternal Survival Series steering, Group. (2006). Going to scale with professional skilled care. *Lancet, 368*(9544), 1377–1386.

Koenig, M. A., Fauveau, V., Chowdhury, A. I., Chakraborty, J., & Khan, M. A. (1988). Maternal mortality in Matlab, Bangladesh: 1976–85. *Studies in Family Planning, 19*(2), 69–80.

Kuncharapu, I., Majeroni, B. A., & Johnson, D. W. (2010). Pelvic organ prolapse. *American Family Physician, 81*(9), 1111–1117.

Kusiako, T., Ronsmans, C., & Van der Paal, L. (2000). Perinatal mortality attributable to complications of childbirth in Matlab, Bangladesh. *Bulletin of the World Health Organization, 78*(5), 621–627.

Kwast, B. E., Rochat, R. W., & Kidane-Mariam, W. (1986). Maternal mortality in Addis Ababa, Ethiopia. *Studies in Family Planning, 17*(6 Pt 1), 288–301.

Kwast, B. E., & Stevens, J. A. (1987). Viral hepatitis as a major cause of maternal mortality in Addis Ababa, Ethiopia. *International Journal of Gynecology & Obstetrics, 25*(2), 99–106.

Labbok, M. H., Hight-Laukaran, V., Peterson, A. E., Fletcher, V., von Hertzen, H., & Van Look, P. F. (1997). Multicenter study of the Lactational Amenorrhea Method (LAM): I. Efficacy, duration, and implications for clinical application. *Contraception, 55*(6), 327–336.

Larsen, U. (2000). Primary and secondary infertility in sub-Saharan Africa. *International Journal of Epidemiology, 29*(2), 285–291.

Larsen, U., Chung, W., & Das Gupta, M. (1998). Fertility and son preference in Korea. *Population Studies, 52*(3), 317–325.

Lawson, J. (1992). Vaginal fistulae. *Journal of the Royal Society of Medicine, 85*(5), 254–256.

Lee, A. C., Katz, J., Blencowe, H., Cousens, S., Kozuki, N., Vogel, J. P., . . . CHERG SGA- Preterm Birth Working Group. (2013). National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010. *Lancet Global Health, 1*(1), e26–e36.

Levin, H., Pollitt, E., Galloway, R., & McGuire, J. (1993). Micronutrient deficiency disorders. In D. Jamison, W. H. Mosley, A. Measham, & J. L. Bobadilla (Eds.), *Disease control priorities in developing countries* (pp. 421–451). New York, NY: Oxford University Press.

Lewis, J. J., Ronsmans, C., Ezeh, A., & Gregson, S. (2004). The population impact of HIV on fertility in sub-Saharan Africa. *AIDS, 18*(suppl 2), S35–S43.

Lewis, P. Population control measures to aid women are stumbling. (1999, April 10). *New York Times*. Retrieved from https://www.nytimes.com

Lewis, P. Conference adopts plan on limiting population. (1999, July 3). *New York Times*. Retrieved from https://www.nytimes.com

Liu, L., Hill, K., Oza, S., Hogan, D., Chu, Y., Cousens, S., . . . Robert, B. E. (2016). Levels and causes of mortality under age five years. In R. E. Black, R. Laxminarayan, M. Temmerman, & N. Walker (Eds.), *Reproductive, maternal, newborn, and child health: Disease control priorities* (pp. 71-83) (3rd ed., Vol. 2). Washington, DC: World Bank Group.

Liu, L., Oza, S., Hogan, D., Perin, J., Rudan, I., Lawn, J. E., . . . Black, R. E. (2015). Global, regional, and national causes of child mortality in 2000–13, with projections to inform post-2015 priorities: An updated systematic analysis. *Lancet, 385*(9966), 430–440.

Lloyd, C. B. (1994). Investing in the next generation: The implications of high fertility at the level of the family. In R. Cassen (Ed.), *Population and development: Old debates, new conclusions* (pp. 181–202). New Brunswick, NJ: Transaction Publishers.

Loudon, I. (1991). On maternal and infant mortality, 1900–1960. *Social History of Medicine, 4*(1), 29–73.

Macro International. (1994). Selected statistics from DHS. *Demographic and Health Survey Newsletter, 6,* 2.

Macro International. (2014). Selected statistics from Demographic and Health Surveys. Retrieved from https://www.statcompiler.com/en/

Macro International. (2015). MEASURE DHS STAT compiler. Retrieved from https://www.statcompiler.com/en/

Maine, D. (1991). *Safe motherhood programs: Options and issues*. New York, NY: Columbia University, Center for Population and Family Health, School of Public Health.

Maine, D., Akalin, M. Z., Chakraborty, J., de Francisco, A., & Strong, M. (1996). Why did maternal mortality decline in Matlab? *Studies in Family Planning, 27*, 179–187.

Mantz, M. L., & Okong, P. (1994). *Evaluation report: Uganda life saving skills program for midwives, October–November, 1994* (MotherCare Project No. 5966-C-00-3038-00, prepared for the U.S. Agency for International Development). Arlington, VA: John Snow.

March of Dimes. (2016). *2016 premature birth report card*. White Plains, NY: Author.

Marshall, M. A., & Buffington, S. T. (1991). *Life-saving skills manual for midwives* (2nd ed.). Washington, DC: American College of Nurse–Midwives.

Martin, J. A., Hamilton, B. E., Sutton, P. D., Ventura, S. J., Menacker, F., & Munson, M. L. (2003). Births: Final data for 2002. *National Vital Statistics Reports, 52*(10), 1–114.

Mason, K. O., & Taj, A. M. (1987). Differences between women's and men's reproductive goals in developing countries. *Population and Development Review, 13*(4), 611–638.

Maternal deaths. (2008). Millennium Development Goals report 2008. Retrieved from http://www.un.org/millenniumgoals/2008highlevel/pdf/newsroom/mdg%20reports/MDG_Report_2008_ENGLISH.pdf

Mathai, M. (2009). The partograph for the prevention of obstructed labor. *Clinical Obstetrics and Gynecology, 52*(2), 256–269.

May, M. T., Gompels, M., Delpech, V., Porter, K., Orkin, C., Kegg, S., . . . Study, U. K. Collaborative HIV Cohort. (2014). Impact on life expectancy of HIV-1 positive individuals of CD4+ cell count and viral load response to antiretroviral therapy. *AIDS, 28*(8), 1193–1202.

McGregor, I. A., Wilson, M. E., & Billewicz, W. Z. (1983). Malaria infection of the placenta in the Gambia, West Africa: Its incidence and relationship to stillbirth, birthweight and placental weight. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 77*(2), 232–244.

McIntosh, C. A., & Finkle, J. L. (1995). The Cairo Conference on Population and Development: A new paradigm? *Population and Development Review, 21*, 223–260.

McNeilly, A. (1996). Breastfeeding and the suppression of fertility. *Food and Nutrition Bulletin, 17*, 340–345.

Menken, J. (1975). *Estimating fecundability.* Unpublished doctoral dissertation, Princeton University, Princeton, NJ.

Menken, J., Khan, M. N., & Williams, J. (1999). *The role of female education in the Bangladesh fertility decline*. Paper presented at the annual meeting of the Population Association of America.

Menken, J., & Kuhn, R. (1996). Demographic effects of breastfeeding: Fertility, mortality, and population growth. *Food and Nutrition Bulletin, 17*, 349–361.

Merchant, K., & Martorell, R. (1987). Frequent reproductive cycling: Does it lead to nutritional depletion of mothers? *Progress in Food & Nutrition Science, 12*(4), 339–369.

Miller, J. E. (1989). Is the relationship between birth intervals and perinatal mortality spurious? Evidence from Hungary and Sweden. *Population Studies, 43*(3), 479–495.

Moriki, Y., Hayashi, K., & Matsukura, R. (2015). Sexless marriages in Japan: Prevalence and reasons. In N. Ogawa & I. H. Shah (Eds.), *Low fertility and reproductive health in East Asia* (pp. 161–185). Dordrecht, Netherlands: Springer.

Morrow, R. H., Jr., Smetana, H. F., Sai, F. T., & Edgcomb, J. H. (1968). Unusual features of viral hepatitis in Accra, Ghana. *Annals of Internal Medicine, 68*(6), 1250–1264.

Mosher, W. D., Martinez, G. M., Chandra, A., Abma, J. C., & Willson, S. J. (2004). Use of contraception and use of family planning services in the United States: 1982–2002. *Advance Data, 350*(1), 1–36.

Muhuri, P. K., & Menken, J. (1997). Adverse effects of next birth, gender, and family composition on child survival in rural Bangladesh. *Population Studies, 51*(3), 279–294.

Mukaire, J., Kalikwani, F., Maggwa, B. N., & Kisubi, W. (1997). *Integration of STI and HIV/AIDS services with MCH-FP services: A case study of the Busoga Diocese Family Life Education Program Uganda*. Nairobi, Kenya: Operations Research Technical Assistance, Africa Project II, Population Council.

National Institute of Population Research and Training (NIPORT). (2003). *Bangladesh Maternal Health Services and Maternal Mortality Survey 2001*. Dhaka, Bangladesh/Calverton, MD: NIPORT, ORC Macro, Johns Hopkins University, & International Centre for Diarrhoeal Disease Research, Bangladesh.

National Institute of Population Research and Training (NIPORT). (2011). *Bangladesh Maternal Health Services and Maternal Mortality Survey 2010*. Dhaka, Bangladesh/Calverton, MD: NIPORT, ORC Macro, Johns Hopkins University, & International Centre for Diarrhoeal Disease Research, Bangladesh.

National Research Council. (1989). *Contraception and reproduction: Health consequences for women and children in the developing world. Committee on Population, Working Group on Healthy Consequences of Contraceptive Use and Controlled Fertility*. Washington DC: National Academies Press.

National Statistics Office (Philippines) & Macro International. (1993). *Philippines National Safe Motherhood Survey*. Calverton MD: National Statistics Office & Macro International.

National Statistics Office (Philippines) & Macro International. (1994). *Philippines National Safe Motherhood Survey, 1993*. Calverton, MD: National Statistics Office & Macro International.

NIPORT, icddr,b & MEASURE Evaluation (2017). *Bangladesh Maternal Mortality and Health Care Survey 2016: Preliminary Report*. Dhaka, Bangladesh and Chapel Hill, NC: NIPORT, icddr,b and MEASURE Evaluation.

Notestein, F. W. (1953). *Economic problems of population change.* Paper presented at the the Eighth International Conference of Agricultural Economics, London, UK.

Nybo Andersen, A. M., Hansen, K. D., Andersen, P. K., & Davey Smith, G. (2004). Advanced paternal age and risk of fetal death: A cohort study. *American Journal of Epidemiology, 160*(12), 1214–1222.

O'Rourke, K. (1995). The effect of hospital staff training on management of obstetrical patients referred by traditional birth attendants. *International Journal of Gynaecology and Obstetrics, 48*(suppl), S95–S102.

Overall, J. C. (1987). Viral infections of the fetus and neonate. In R. D. Feigin & J. D. Cherry (Eds.), *Textbook of pediatric infectious diseases* (2nd ed., pp. 966–1007). Philadelphia, PA: W. B. Saunders.

Oza, S., Cousens, S. N., & Lawn, J. E. (2014). Estimation of daily risk of neonatal death, including the day of birth, in 186 countries in 2013: A vital-registration and modelling-based study. *Lancet Global Health, 2*(11), e635–e644.

Palermo, T., Bleck, J., & Westley, E. (2014). Knowledge and use of emergency contraception: A multicountry analysis. *International Perspectives on Sexual and Reproductive Health, 40*(2), 79–86.

Palloni, A. (1985). Health conditions in Latin America and policies for mortality change. In J. Vaillin & A. Lopez (Eds.), *Health policy, social policy and mortality prospects* (pp. 465–492). Liege, Belgium: Ordina Editions.

Palloni, A., & Millman, S. (1986). Effects of inter-birth intervals and breastfeeding on infant and early childhood mortality. *Population Studies, 40*(2), 215–236.

Paxton, A., Maine, D., Freedman, L., Fry, D., & Lobis, S. (2005). The evidence for emergency obstetric care. *International Journal of Gynaecology and Obstetrics, 88*(2), 181–193.

Pebley, A. R., & DaVanzo, J. (1988, April). *Maternal depletion and child survival in Guatemala and Malaysia.* Paper presented at the annual meeting of the Population Association of America, New Orleans, LA.

Pebley, A. R., & Stupp, P. W. (1987). Reproductive patterns and child mortality in Guatemala. *Demography, 24*(1), 43–60.

Poedje, R., Setjalilakusuma, L., Abadi, A., Soegianto, B., Rihadi, S., Djaeli, A., & Budiarto, W. (1993). *1993. Final project report: East Java Safe Motherhood Study* (Report prepared for the U.S. Agency for International Development. MotherCare Project #936-5966). Arlington, VA: John Snow.

Population Reference Bureau. (2002). *Family planning worldwide 2008 data sheet*. Washington, DC: Author.

Population Reference Bureau. (2004). The unfinished agenda: Meeting the need for family planning in less developed countries. *PRB Policy Brief.* Retrieved from http://www.prb.org/Publications/Reports/2004/TheUnfinishedAgendaMeetingtheNeedforFamilyPlanninginLessDevelopedCountriesPDF306KB.aspx

Population Reference Bureau. (2008). *Family planning worldwide 2008 data sheet*. Washington, DC: Author.

Population Reference Bureau. (2016). *2016 world population data sheet*. Washington, DC: Author.

Potter, J. E. (1988). Does family planning reduce infant mortality? *Population and Development review, 14*(1), 179–187.

Prevention of Maternal Mortality Network. (1995). Situational analyses of emergency obstetric care: Examples from eleven operations research projects in West Africa. *Social Science and Medicine, 40*(suppl), 657–667.

Pritchett, L. H. (1994a). Desired fertility and the impact of population policies. *Population and Development Review, 20*(1), 1–55.

Pritchett, L. H. (1994b). The impact of population policies: Reply. *Population and Development Review, 20*(3), 621–630.

Public Opinion Polls. (1993). *MotherCare Nigeria maternal healthcare project qualitative research* (MotherCare Working Paper No. 17B, prepared for the U.S. Agency for International Development, Project No. 936-5966). Arlington, VA: John Snow.

Rafalimanana, H., & Westoff, C. F. (2001). *Gap between preferred and actual birth intervals in sub-Saharan Africa: Implications for fertility and child health.* DHS Analytical Studies No. 2. Calverton, MD: ORC Macro.

Rahman, M. O. (1999). Family matters: The impact of kin on the mortality of the elderly in rural Bangladesh. *Population Studies, 53*(2), 227–235.

Rahman, O., Foster, A., & Menken, J. (1992). Older widow mortality in rural Bangladesh. *Social Science & Medicine, 34*(1), 89–96.

Reed, H. E., Koblinsky, M. A., & Mosley, W. H. (2000). *The consequences of maternal morbidity and maternal mortality. Committee on Population, Commission on Behavioral and Social Sciences and Education, National Research Council.* Washington, DC: National Academies Press.

Ronsmans, C., & Campbell, O. (1998). Short birth intervals don't kill women: Evidence from Matlab, Bangladesh. *Studies in Family Planning, 29*(3), 282–290.

Ronsmans, C., Graham, W. J., & The Lancet Maternal Survival Series steering group. (2006). Maternal mortality: Who, when, where, and why. *Lancet, 368*(9542), 1189–1200.

Rooney, C. (1992). *Antenatal care and maternal health: How effective is it? A review of the evidence* (WHO/MSM/92.4). Geneva, Switzerland: World Health Organization.

Rosenzweig, M. R., & Schultz, T. P. (1983). Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *Journal of Political Economy, 91*(5), 723–746.

Ross, A., Van der Paal, L., Lubega, R., Mayanja, B. N., Shafer, L. A., & Whitworth, J. (2004). HIV-1 disease progression and fertility: The incidence of recognized pregnancy and pregnancy outcome in Uganda. *AIDS, 18*(5), 799–804.

Ross, J. A., & Frankenberg, E. (1993). *Findings from two decades of family planning research*. New York, NY: Population Council.

Ross, J., Keesbury, J., & Hardee, K. (2015). Trends in the contraceptive method mix in low- and middle-income countries: Analysis using a new "average deviation" measure. *Global Health: Science and Practice, 3*(1), 34–55.

Ross, J. A., & Mauldin, W. P. (1996). Family planning programs: efforts and results, 1972–94. *Studies in Family Planning, 27*(3), 137–147.

Ross, J. A., & Winfrey, W. L. (2002). Unmet need for contraception in the developing world and the former Soviet Union: An updated estimate. *International Family Planning Perspectives, 28*(3), 137–147.

Rutstein, S. O. (1998). *Change in the desired number of children: A cross-country cohort analysis of levels and correlates of change* (No. DHS Analytical Report No. 9). Calverton, MD: Macro International.

Safe Motherhood Initiative. (2010). Retrieved from http://www.safemotherhood.org

Salber, E. J., Feinleib, M., & Macmahon, B. (1965). The duration of postpartum amenorrhea. *American Journal of Epidemiology, 82*(3), 347–358.

Say, L., Chou, D., Gemmill, A., Tuncalp, O., Moller, A. B., Daniels, J., . . . Alkema, L. (2014). Global causes of maternal death: A WHO systematic analysis. *Lancet Global Health, 2*(6), e323–e333.

Schieber, B. A., Mejia, M., Koritz, S., Gonsalez, C., & Kwast, B. (1995). *Medical audit of early neonatal deaths. INCAP/Quetzaltenango Maternal and Neonatal Health project.* (Technical Working Paper No. 1, prepared for the U.S. Agency for International Development, MotherCare Project No. 936-5966). Arlington, VA: John Snow.

Schultz, T. P. (1984). Studying the impact of household economic and community variables on child mortality. *Population and Development Review, 10*(suppl), 215–235.

Sedgh, G., Bearak, J., Singh, S., Bankole, A., Popinchalk, A., Ganatra, B., . . . Alkema, L. (2016). Abortion incidence between 1990 and 2014: Global, regional, and subregional levels and trends. *Lancet, 388*(10041), 258–267.

Sedgh, G., Hussain, R., Bankole, A., & Singh, S. (2007). *Women with an unmet need for contraception in developing countries and their reasons for not using a method* (No. 37). New York, NY: Guttmacher Institute.

Sedgh, G., Singh, S., & Hussain, R. (2014). Intended and unintended pregnancies worldwide in 2012 and recent trends. *Studies in Family Planning, 45*(3), 301–314.

Sedgh, G., Singh, S., Shah, I. H., Ahman, E., Henshaw, S. K., & Bankole, A. (2012). Induced abortion: Incidence and trends worldwide from 1995 to 2008. *Lancet, 379*(9816), 625–632.

Sheps, M. C., & Menken, J. (1973). *Mathematical models of conception and birth*. Chicago, IL: University of Chicago Press.

Singh, S., & Darroch, J. E. (2012). *Adding it up: Costs and benefits of contraceptive services.* New York, NY: Guttmacher Institute & UNFPA.

Slama, R., Werwatz, A., Boutou, O., Ducot, B., Spira, A., & Härdle, W. (2003). Does male age affect the risk of spontaneous abortion? An approach using semiparametric regression. *American Journal of Epidemiology, 157*(9), 815–824.

Slanger, T. E., Snow, R. C., & Okonofua, F. E. (2002). The impact of female genital cutting on first delivery in southwest Nigeria. *Studies in Family Planning, 33*(2), 173–184.

Sloan, N. L., Jordan, E. A., & Winikoff, B. (1992). *Does iron supplementation make a difference?* (MotherCare Working Paper No. 15, prepared for the U.S. Agency for International Development, MotherCare Project No. 936-5966). Arlington, VA: John Snow.

Sood, S. K., Ramachandran, K., Mathur, M., Gupta, K., Ramalingaswamy, V., Swarnabai, C., C, . . . Baker, S. J. (1975). WHO sponsored collaborative studies on nutritional anemia in India. Part I: The effects of supplemental oral iron administration to pregnant women. *QJM, 44*(2), 241–258.

Stewart, M. K., & Festin, M. (1995). Validation study of women's reporting and recall of major obstetric complications treated at the Philippine General Hospital. *International Journal of Gynecology & Obstetrics, 48*(suppl), S53–S66.

Stone, J. L., Lockwood, C. J., Berkowitz, G. S., Alvarez, M., Lapinski, R., & Berkowitz, R. L. (1994). Risk factors for severe preeclampsia. *Obstetrics and Gynecology, 83*(3), 357–361.

Stover, J. (1998). Revising the proximate determinants of fertility framework: What have we learned in the past 20 years? *Studies in Family Planning, 29*(3), 255–267.

Streatfield, P. K., & Kamal, N. (2013). Population and family planning in Bangladesh. *Journal of Pakistan Medical Association, 63*(4 suppl 3), S73–S81.

Sundari, T. K. (1992). The untold story: How the health care systems in developing countries contribute to maternal mortality. *International Journal of Health Services, 22*(3), 513–528.

Tahzib, F. (1983). Epidemiological determinants of vesicovaginal fistulas. *BJOG: An International Journal of Obstetrics & Gynaecology, 90*(5), 387–391.

Tahzib, F. (1985). Vesicovaginal fistula in Nigerian children. *Lancet, 326*(8467), 1291–1293.

Thaddeus, S., & Maine, D. (1994). Too far to walk: Maternal mortality in context. *Social Science & Medicine, 38*(8), 1091–1110.

Toroitich-Ruto, C. (2001). The evolution of the family planning programme and its role in influencing fertility change in Kenya. *Journal of Biosocial Science, 33*(2), 245–260.

Trottier, D. A., Potter, L. S., Taylor, B. A., & Glover, L. H. (1994). User characteristics and oral contraceptive compliance in Egypt. *Studies in Family Planning, 25*(5), 284–292.

Trusell, J., & Cleland, K. (2010). The emergency contraception website. Retrieved from http://ec.princeton.edu/questions/dedicated.html

Trussell, J., Ellertson, C., & Stewart, F. (1996). The effectiveness of the Yuzpe regimen of emergency contraception. *Family Planning Perspectives, 28*(2), 58–64, 87.

Tsui, A. O., Wasserheit, J. N., & Haaga, J. (1997). *Reproductive health in developing countries: Expanding dimensions, building solutions. Panel on Reproductive Health, Committee on Population, Commission on Behavioral and Social Sciences and Education, National Research Council*. Washington DC: National Academy Press.

Turner, C., Miller, G., & Moses, L. (1989). *AIDS: Sexual behaviour and intravenous drug use. Committee on AIDS Research and the Behavioral, Social, and Statistical Sciences, Commission on Behavioral and Social Sciences and Education, National Research Council.* Washington, DC: National Academy Press.

Twahir, A., Maggwa, B. N., & Askew, I. (1996). *Integration of STI and HIV/AIDS services with MCH-FP services: A case study of the Mkomani clinic society in Mombasa, Kenya*. Nairobi, Kenya: Operations Research Technical Assistance, Africa Project II, Population Council.

UNAIDS Fact Sheets. (2009). AIDSinfo database: Country fact sheets. Retrieved from http://www.unaids.org/en/regionscountries/countries/

UNICEF. (2013). *UNICEF data monitoring the situation of children and women. LBW: Country, regional and global estimates*. New York, NY: Author.

UNICEF. (2016). Female genital mutilation/cutting: A global concern. Retrieved from https://www.unicef.org/media/files/FGMC_2016_brochure_final_UNICEF_SPREAD.pdf

UNICEF & World Health Organization (WHO). (2004). *Low birthweight: Country, regional and global estimates*. New York, NY: UNICEF.

UNICEF & World Health Organization (WHO). (2015). *Countdown to 2015: A decade of tracking progress for maternal, newborn and child survival: The 2015 report* (No. 0140-6736). Geneva, Switzerland: World Health Organization.

United Nations. (1991). *Controlling iron deficiency* (Nutrition State-of-the-Art Series, Nutrition Policy Discussion Paper No. 9). Geneva, Switzerland: Author.

United Nations. (1994). *Programme of action of the 1994 International Conference on Population and Development*. Paper presented at the (A /CONF.171/13).

United Nations. (2016). Progress towards the Sustainable Development Goals: Report of the Secretary-General. Retrieved from https://sustainabledevelopment.un.org/sdg3

United Nations Children's Fund. (2008). Countdown to 2015: Tracking progress in maternal, newborn and child survival: The 2008 report. Retrieved from http://www.countdown2015mnch.org/index.php?option=com_content&view=article&id=68&itemid=61

United Nations Commission on Population and Development. (2004). Flow of financial resources for assisting in the implementation of the Programme of Action of the International Conference on Population and Development: A 10-year review. Retrieved from https://documents-dds-ny.un.org/doc/UNDOC/GEN/N04/206/10/PDF/N0420610.pdf?OpenElement

United Nations General Assembly. (2000). United Nations Millennium Declaration (A/RES/55/2). Retrieved from http://www.un.org/en/development/devagenda/millennium.shtml

United Nations General Assembly. (2001). Road map towards the implementation of the United Nations Millennium Summit: Report of the Secretary-General (A/56/326). Retrieved from http://www.un.org/documents/ga/docs/56/a56326.pdf

United Nations Population Division. (2002). *HIV/AIDS and fertility in sub-Saharan Africa: A review of the research literature (ESA/P/WP.174).* Retrieved from http://www.un.org/esa/population/publications/fertilitysection/HIVAIDSPaperFertSect.pdf

United Nations Population Division. (2009). *World population prospects: The 2008 revision population database.* Retrieved from http://www.un.org/esa/population/publications/wpp2008/

United Nations Population Division. (2015a). *Trends in contraceptive use worldwide 2015 (ST/ESA/SER.A/349).* Retrieved from http://www.un.org/en/development/desa/population/publications/pdf/family/trendsContraceptiveUse2015Report.pdf

United Nations Population Division. (2015b). *World population prospects: The 2015 revision. Volume I: Comprehensive tables (ST/ESA/SER.A/379).* Retrieved from https://esa

.un.org/unpd/wpp/publications/Files/WPP2015_Volume-I_Comprehensive-Tables.pdf

United Nations Population Fund (UNFPA). (2004). *State of world population 2004: The Cairo consensus at ten: Population, reproductive health and the global effort to end poverty*. New York, NY: Author.

United Nations Population Fund (UNFPA). (2009). State of the world's population 2009: Facing a changing world: women, population and climate change. *State of World Population.* Retrieved from http://www.unfpa.org/swp/2009/en/overview.shtml

U.S. Census Bureau. (2004). Global population profile 2002: The AIDS pandemic in the 21st century. Retrieved from http://www.census.gov/ipc/prod/wp02/wp02-2.pdf

Waiz, N., Faroqul, M., Begum, A., Sultana, N., Sarker, S., & Faisel, A. (2003). *Situation analysis of obstetric fistula in Bangladesh*. Dhaka, Bangladesh: Engender Health.

Walker, G. J., Ashley, D. E., McCaw, A. M., & Bernard, G. W. (1985). *Maternal mortality in Jamaica: A confidential enquiry into all maternal deaths in Jamaica 1981–1983.* Paper presented at the WHO Inter-regional Meeting on Prevention of Maternal Mortality, Geneva, Switzerland.

Weinbreck, P. V., Loustaud, F., Denis, B., Vidal, M., Mounier, M., & DeLumley, L. (1988). Postnatal transmission of HIV infection. *Lancet, 1*(8583), 482.

Westley, S. B. (1995). *Evidence mounts for sex-selective abortion in Asia* (Asia-Pacific Population and Policy No. 34). Honolulu, HI: East-West Population Center.

Westoff, C. F. (2001). *Unmet need at the end of the century*. Calverton, MD: ORC Macro.

Wilcox, A. J., Weinberg, C. R., O'Connor, J. F., Baird, D. D., Schlatterer, J. P., Canfield, R. E., . . . Nisula, B. C. (1988). Incidence of early loss of pregnancy. *New England Journal of Medicine, 319*(4), 189–194.

Wildman, S. (2017, May 24). Trump's budget eliminates US funding for global family planning and famine relief. *Vox*. Retrieved from https://www.vox.com/policy-and-politics/2017/5/24/15681216/trump-budget-cuts-funding-global-family-planning-famine-relief

Winikoff, B. (1983). The effects of birth spacing on child and maternal health. *Studies in Family Planning, 18*(3), 128–143.

Winikoff, B., Elias, C., & Beattie, K. (1994). Special issues of IUD use in resource-poor settings. In C. W. Bardin & D. R. Mishell (Eds.), *Proceedings of the Fourth International Conference on IUDs* (pp. 230–238). New York, NY: Population Council.

Winikoff, B., & Sullivan, M. (1987). Assessing the role of family planning in reducing maternal mortality. *Studies in Family Planning, 18*(3), 128–143.

Women's Studies Project. (1999). *The impact of family planning on women's lives*. Research Triangle Park, NC: Family Health International.

Wood, J. W. (1994). *Dynamics of human reproduction: Biology, biometry, demography.* Piscataway, NJ: Aldine Transaction.

Wood, J. W., Lai, D., Johnson, P. L., Campbell, K. L., & Maslar, I. A. (1985). Lactation and birth spacing in highland New Guinea. *Journal of Biosocial Science, 9*(suppl), 159–173.

World Health Organization (WHO). (1988). Geographic variation in the incidence of hypertension in pregnancy. *American Journal of Obstetrics and Gynecology, 158*(1), 80–83.

World Health Organization (WHO). (1993a). *Coverage of maternity care: A tabulation of available information* (WHO/FHE/MSM/ 93.7). Geneva, Switzerland: Author.

World Health Organization (WHO). (1993b). *Making maternity care more accessible.* Geneva, Switzerland: Author.

World Health Organization (WHO). (1994). *Mother–baby package: Implementing safe motherhood in countries*. Geneva, Switzerland: Author.

World Health Organization (WHO). (1995). Essential or emergency obstetric care. *Safe Motherhood Newsletter, 18*(2), 1–2.

World Health Organization (WHO). (2005). *The world health report 2005: Make every mother and child count*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2006). *Neonatal and perinatal mortality: Country, regional and global estimates*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2007a). Maternal mortality in 2005: Estimates developed by WHO, UNICEF, UNFPA, and the World Bank. Retrieved from http://www.who.int/reproductive-health/publications/maternal_mortality_2005/index.html

World Health Organization (WHO). (2007b). *Unsafe abortion: Global and regional estimates of the incidence of unsafe abortion and associated mortality in 2003*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2008a). *Maternal mortality factsheet* (Document WHO/MPS/08.12). Geneva, Switzerland: Author.

World Health Organization (WHO). (2008b). *Skilled birth attendants factsheet* (Document WHO/MPS/08.11). Geneva, Switzerland: Author.

World Health Organization (WHO). (2011). *Unsafe abortion: Global and regional estimates of the incidence of unsafe abortion and associated mortality in 2008*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2015a). *Every woman, every child, every adolescent: Achievements and prospects: The final report of the independent Expert Review Group on Information and Accountability for Women's and Children's health* (No. 9241509287). Gdneva, Switzerland: Author.

World Health Organization (WHO). (2015b). *Trends in maternal mortality: 1990–2015: Estimates from WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division* (No. 9241565144). Geneva, Switzerland: Author.

World Health Organization (WHO). (2016). *World health statistics 2016: Monitoring health for the Sustainable Development Goals*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2017). Preventing unsafe abortion. Retrieved from http://www.who.int/mediacentre/factsheets/fs388/en/

Yoder, P. S., Abderrahim, N., & Zhuzhuni, A. (2004). *Female genital cutting in the Demographic and Health Surveys: A critical and comparative analysis*. Calverton, MD: ORC Macro.

You, D., Hug, L., Ejdemyr, S., & Beise, J. (2015). *Levels and trends in child mortality: Report 2015. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation*. New York, NY: UNICEF.

Younis, N., Khattab, H., Zurayk, H., el-Mouelhy, M., Amin, M. F., & Farag, A. M. (1993). A community study of gynecological and related morbidities in rural Egypt. *Studies in Family Planning, 24*(3), 175–186.

Zupan, J. (2005). Perinatal mortality in developing countries. *New England Journal of Medicine, 352*(20), 2047–2048.

CHAPTER 6

Infectious Diseases

Aubree Gordon and Arthur L. Reingold

This chapter describes the epidemiologic features of the infectious diseases of greatest public health significance in low- and middle-income countries (LMICs) and details available strategies to prevent and control them. Because these diseases cannot be covered in depth in a single chapter, the emphasis here is on their unique epidemiologic features and the relevant technological challenges, resource limitations, and cultural barriers that have shaped current approaches to their prevention and control. Conceptually, these approaches include preventing exposure to the infectious agent; making otherwise susceptible individuals or populations immune to the infectious agent; treating infected individuals or populations to prevent illness and transmission of the agent to others; and improving the timeliness and appropriateness of care for symptomatic individuals so as to minimize morbidity and mortality and, in some instances, to reduce the likelihood of transmission to others. Examples of successful programs using one or more of these various conceptual approaches are discussed, as are the challenges and obstacles that confront LMICs and their partners as they seek to reduce the burden of disease caused by infectious agents.

Overview

Collectively, infectious diseases have undoubtedly been the single most important contributors to human morbidity and mortality throughout history. Over the past 150 years, the mortality attributable to them has declined substantially in high-income countries, and chronic diseases such as cardiovascular disease, cancer, stroke, chronic obstructive pulmonary disease, and diabetes mellitus have assumed prominence as the leading causes of death in these countries. Although uncertainty exists about the relative importance of various social, economic, environmental, and public health factors in this epidemiologic transition, most of the reductions in mortality attributable to infectious diseases clearly preceded any advances in clinical medicine and public health that plausibly could have had an impact on the infectious diseases of public health significance of the time (e.g., tuberculosis, rheumatic fever, scarlet fever, typhoid fever, and cholera). At present, only influenza and pneumonia rank among the top 10 causes of mortality in the United States (Heron, 2016).

The global burden of disease and the epidemiologic transition are discussed in detail elsewhere in this text (see the *Measures of Health and Disease in Populations* chapter). Although current projections suggest that acute infectious diseases will decrease substantially in their absolute and relative importance as causes of death and disability in LMICs in the decades to come, today they remain issues of great public health significance. Acute infectious diseases are, collectively, the leading cause of death among children, accounting for more than half of all child deaths in LMICs (see the *Measures of Health and Disease in Populations* chapter), as well as being important causes of morbidity and mortality in people at all stages of life. In many countries, the human immunodeficiency virus (HIV)/

acquired immunodeficiency syndrome (AIDS) pandemic initially reversed decades of progress in reducing mortality due to acute infectious diseases, with a resulting decrease in life expectancy (Lozano et al., 2013; UNAIDS, 2013). Further, it is now well accepted that chronic infection contributes in important—if poorly understood—ways to the pathogenesis of a number of chronic diseases, including cervical cancer (in which human papillomavirus [HPV] plays a role), hepatic cancer and cirrhosis (in which hepatitis B virus [HBV] and probably hepatitis C virus [HCV] play a role), gastric cancer and peptic ulcer disease (in which *Helicobacter pylori* plays a role), and possibly cardiovascular disease (in which *Chlamydia pneumoniae* and perhaps other infectious agents may play roles). Hepatic cancer and cirrhosis due to HBV were the first vaccine-preventable chronic diseases, and it is expected that widespread vaccination against HPV will reduce by two-thirds the number of cervical cancer cases worldwide and reduce the incidence of anal and oral cancers. Thus, for all the reasons just mentioned, infectious diseases and their prevention and control will remain of major public health importance for LMICs for the foreseeable future.

Underlying almost every infectious disease of public health importance in LMICs is the significant role played by poverty and its associated problems. For example, both obvious and more subtle forms of malnutrition and micronutrient deficiencies are associated with an increased risk of severe morbidity and mortality from a wide range of infectious diseases (see the *Nutrition* chapter, which highlights the association between infectious disease and malnutrition). At the same time, lack of education, poor access to clean drinking water, inability to dispose properly of human waste, household crowding, and lack of access to medical care—all manifestations of poverty—make substantial contributions to their incidence.

However, LMICs and the people living in them cannot be lumped together into a single group insofar as their risk of various infectious diseases is concerned. Important geographic differences exist in the incidence and public health significance of various infectious diseases, due to differences in climate, the distribution of insect vectors and intermediate hosts, and variations in other environmental, social, and cultural factors. In addition, all LMICs are not equally resource poor—they vary enormously in the resources that are available to provide clinical services (e.g., oral rehydration therapy), mount public health programs (e.g., vaccination programs), and reduce environmental sources of infection (e.g., provide clean drinking water and adequate sanitation or control vector populations). Also, all LMICs include within them culturally, economically, and sometimes geographically diverse subpopulations with very different needs and resources, particularly with regard to infectious diseases. Given the enormous diversity of LMICs and the infectious diseases that confront them, it will be possible in this chapter to discuss these diseases only selectively, using representative examples when appropriate.

Control of Infectious Diseases

The twentieth century saw an ever-increasing number of programs to prevent morbidity and mortality from specific infectious diseases in LMICs. Strategies that have been employed include various combinations of vector control (e.g., for malaria, dengue, yellow fever, and onchocerciasis [river blindness]); vaccination (e.g., for smallpox, measles, polio, neonatal tetanus, diphtheria, pertussis, tetanus, hepatitis B, meningococcal meningitis, and yellow fever); mass chemotherapy (e.g., for hookworm, onchocerciasis, dracunculiasis [Guinea worm]); improved sanitation and access to clean water (e.g., for diarrheal diseases); improved care seeking and caregiving (e.g., for diarrheal diseases and acute respiratory infections); and behavior change (e.g., for HIV and other sexually transmitted infections [STI]s, diarrheal diseases, and dracunculiasis), among others. The successful eradication of smallpox in the late 1970s through a combination of enhanced case finding, containment, and vaccination gave considerable impetus to attempts to control other infectious diseases (**EXHIBIT 6-1**).

As the world health community has established goals for reducing morbidity and mortality from other infectious diseases, a variety of terms describing different levels of control have come into use. Organizations such as the World Health Organization (WHO) and its governing body, the World Health Assembly (WHA), have been careful to define their prevention goals vis-à-vis various diseases; these are set forth in appropriate sections of this chapter. A useful set of definitions of such terms was put forward at the Dahlem workshop on the Eradication of Infectious Diseases (Dowdle & Hopkins, 1998):

- *Control:* Reduction of disease incidence, prevalence, morbidity or mortality to a locally acceptable level as a result of deliberate efforts; continued intervention measures are required to maintain the reduction.
- *Elimination of disease:* Reduction to zero of the incidence of a specified disease in a defined

EXHIBIT 6-1 Smallpox Eradication

Most individuals who work in the area of global health consider the eradication of smallpox to have been the single most important contribution of public health efforts in the twentieth century, and possibly the most significant accomplishment in the field of human health in recorded history. Although the ultimate eradication of smallpox through the use of vaccine was foreseen by Edward Jenner and President Thomas Jefferson at the beginning of the nineteenth century, it took more than 150 years for it to become a reality. Eradication of smallpox was possible because of several important features of the disease itself; technological advances in vaccine preparation and administration; development and application of a new approach to using the vaccine selectively rather than in mass campaigns; and a combination of international will and cooperation, strong leadership, and the focused effort of large numbers of health workers in multiple countries.

One feature of smallpox that made it a candidate for eradication was its relatively inefficient transmission from person to person. In addition, individuals with smallpox were generally bedridden before the appearance of the rash—the stage of the illness when person-to-person transmission was most likely to occur. Moreover, subclinical cases did not occur in unvaccinated individuals, and vaccinated individuals who developed mild smallpox did not efficiently transmit the virus. Smallpox was also characterized by a lack of a carrier state and the existence of a single serotype of the virus. Other notable features of the disease included a marked seasonal fluctuation in cases and the lack of a nonhuman reservoir for the virus in nature.

A key advance in vaccine development and delivery that was crucial to smallpox eradication was the development of a heat-stable, freeze-dried vaccine. In addition, two improved methods of delivering the vaccine were introduced—a bifurcated needle that was inexpensive, easy to use, and economical in its use of a small volume of vaccine, and jet injector guns that allowed a team to vaccinate more than 1,000 persons per hour.

Nevertheless, despite the availability of a heat-stable smallpox vaccine and the means of vaccinating large numbers of persons, mass vaccination campaigns intended to render entire populations immune to smallpox were unsuccessful in eliminating the disease, even in countries that achieved vaccine coverage in the range of 80% to 95%. Smallpox virus continued to circulate among the remaining unvaccinated individuals, who were extremely difficult to identify and vaccinate.

A delay in the arrival of sufficient vaccine to mount a mass campaign against smallpox in Nigeria in 1966 led to the development of an alternative approach to preventing the spread of smallpox in an area: Energetic case detection followed by isolation of all infected individuals and intense vaccination efforts focused on the area and population immediately surrounding a case. This approach, dubbed a surveillance containment strategy, proved to be remarkably successful in eradicating smallpox once imaginative and locally acceptable approaches to detecting all suspected cases, isolating individuals with smallpox, and vaccinating those around them were implemented. Surveillance containment ultimately replaced mass vaccination campaigns, and global efforts to complete eradication of smallpox relying on this approach gained momentum.

In 1967, as many as 10 to 15 million cases of smallpox occurred in 33 countries with endemic smallpox and 14 other countries with imported cases of smallpox. These countries, which had a total population of more than 1.2 billion persons at the time, included many of the poorest countries in the world and those presenting the greatest logistical barriers to mounting an effective eradication program. As a result of the efforts of dedicated public health workers in these countries and a small cadre of public health professionals from unaffected countries, the last person with smallpox not caused by a laboratory accident had onset of a rash on October 26, 1977, in Somalia. On December 9, 1979, the World Health Organization's Global Commission for the Certification of Smallpox Eradication concluded that "smallpox eradication has been achieved throughout the world," a conclusion accepted by the World Health Assembly in May 1980.

geographic area as a result of deliberate efforts; continued intervention measures are required.

- *Elimination of infection:* Reduction to zero of the incidence of infection caused by a specific agent in a defined geographic area as a result of deliberate efforts; continued measures to prevent reestablishment of transmission are required.
- *Eradication:* Permanent reduction to zero of the worldwide incidence of infection caused by a specific agent as a result of deliberate efforts; intervention measures are no longer needed.
- *Extinction:* The specific infectious agent no longer exists in nature or in the laboratory.

In the near term, extinction is possible only for smallpox, although concerns about its use as an agent of bioterrorism have prevented the long-awaited destruction of the last known stocks of the virus. Eradication of other diseases, such as polio, measles, rubella, and dracunculiasis, is considered theoretically possible using existing control methods and is being actively pursued. Unfortunately, for most of the infectious

TABLE 6-1 Levels of Control Considered Achievable for Selected Infectious Diseases in the Foreseeable Future Using Currently Available Methods

Extinction	Eradication	Elimination of Infection	Elimination of Disease	Control
Smallpox	Polio	Onchocerciasis	Rabies	Malaria
	Measles Rubella		Trachoma	Neonatal tetanus
	Dracunculiasis (Guinea worm)			Cholera
				Tuberculosis
				Schistosomiasis
				Diarrheal disease
				Acute respiratory infections
				Acquired immunodeficiency syndrome
				Sexually transmitted infections
				Leprosy

diseases responsible for the majority of morbidity and mortality in LMICs at the beginning of the twenty-first century, only their control is considered achievable in the foreseeable future (**TABLE 6-1**).

Childhood Vaccine-Preventable Diseases: The Expanded Program on Immunization

Overview

Based on the success of the vaccination program mounted to control and then eradicate smallpox, WHO and various partner agencies launched the Expanded Program on Immunization (EPI, or PEV in French) in 1974. At that time, it was estimated that fewer than 5% of infants and children in LMICs were receiving relatively inexpensive and highly effective vaccines, despite the fact that these vaccines had been licensed and available for a number of years. Many obstacles to vaccinating these children existed, including the lack of demand for vaccines on the part of the community; the small number of sources of vaccines of adequate quality; the lack of the infrastructure needed to purchase, store, and distribute vaccines, some of which were temperature sensitive; a deficiency in the number of trained personnel to administer vaccination programs; and insufficient funds to purchase vaccines and vaccination supplies and equipment. Further, most countries lacked health information and surveillance systems to assess the burden of disease caused by various vaccine-preventable diseases or to evaluate the impact of a vaccination program. Remarkable progress since 1974 in correcting these problems has led to the expected eradication through immunization of polio; dramatic reductions in morbidity and mortality from measles and neonatal tetanus worldwide; and likely, but more difficult to demonstrate, reductions of a similar magnitude in morbidity and mortality from diphtheria and pertussis.

For the first 20 years or so of its existence, the EPI focused on diseases for which safe, effective, and inexpensive vaccines were available and could be given

entirely during the first year of life. These included oral polio vaccine (OPV, a trivalent live vaccine against poliomyelitis); measles vaccine (a live vaccine); a three-component killed vaccine against diphtheria, pertussis, and tetanus (DPT); and bacillus Calmette-Guérin (BCG, a live vaccine against tuberculosis). Vaccination with tetanus toxoid (TT) was included for women of childbearing age to prevent neonatal tetanus in their newborn babies, even though the group being targeted for vaccination did not consist of infants in the first year of life. Subsequently, it was recommended that vaccine against yellow fever and Japanese B encephalitis be added in countries where these two diseases pose a threat. Vaccines against hepatitis B and *Haemophilus influenzae* b were then included after large quantities of these relatively inexpensive vaccines became available. Newer vaccines, including rubella vaccine, rotavirus vaccine, pneumococcal conjugate vaccine, and HPV vaccine, have also recently been added (**FIGURE 6-1**). (**TABLE 6-2** provides the current WHO recommendations concerning infant and childhood immunizations.) Hepatitis B, *H. influenzae* b, *Streptococcus pneumoniae,* HPV, yellow fever, and tuberculosis are discussed elsewhere in this chapter, so they will not be considered here.

Poliomyelitis

Etiologic Agent, Clinical Features, and Characteristics of the Currently Available Vaccines

Poliomyelitis (polio) may be caused by any of the three known serotypes (1, 2, and 3) of poliovirus. This virus is efficiently transmitted through the fecal–oral route. Ingestion of the virus leads to asymptomatic or mild, self-limited infection and shedding of the virus from the throat and gastrointestinal tract in the vast majority of exposed persons. However, an estimated 1 in 100 to 1 in 850 infected persons develops symptomatic

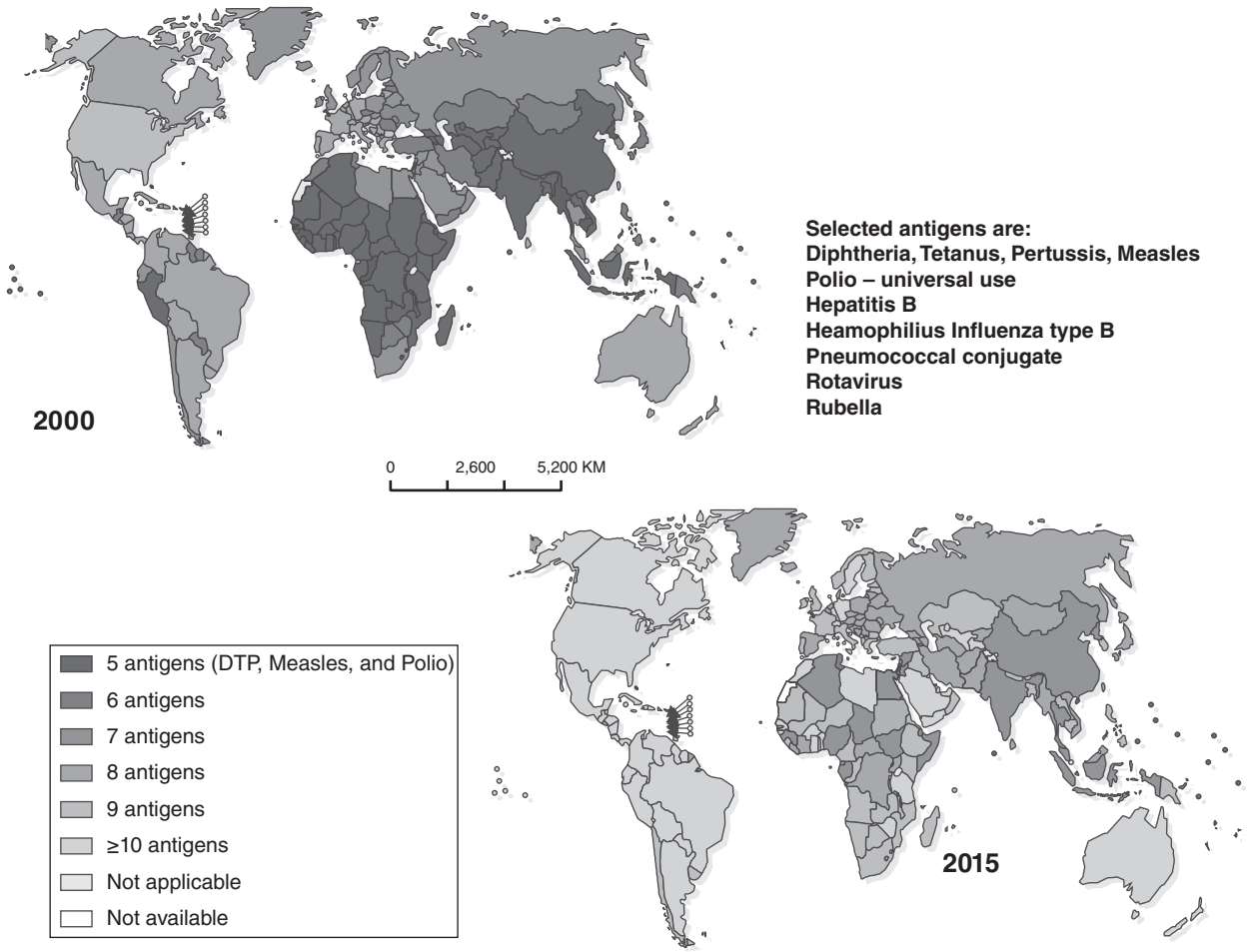


FIGURE 6-1 Number of vaccines/antigens introduced in national immunization schedules, 2000 compared to July 2015.

Reproduced from World Health Organization. (2016). Progress towards global immunization goals – 2015 summary presentation of key indicators [PowerPoint slides], slide 11.

polio, with or without paralysis. Of those who develop paralysis, which primarily affects one or both legs, approximately 10% die acutely, 10% to 15% are left permanently unable to walk, and 10% to 15% are left lame (unable to walk normally).

Treatment for polio is entirely supportive in nature. Before widespread vaccination was undertaken, polio was the leading cause of lameness in LMICs. Many of its victims in these countries remain a visible sign of the ravages of the disease and will

TABLE 6-2 Recommended Routine Immunizations for Children: Summary of WHO Position Papers*

Disease/Vaccine		Age of First Dose	Dose in Primary Series	Booster Dose
Recommended for All Children				
Bacillus Calmette-Guérin (BCG)		As soon as possible after birth	1	
Diphtheria, pertussis, and tetanus (DPT)		6 weeks	3	12–23 months (DPT containing vaccine); 4–7 years (Td); and 9–15 yrs (Td)
Haemophilus influenzae b	(Option 1)	6 weeks (minimum)–59 months (maximum)	3	
	(Option 2)		2–3	At least 6 months (minimum) after last dose
Hepatitis B	(Option 1)	As soon as possible after birth (< 24 hours)	3	
	(Option 2)		4	
Human papillomavirus (HPV)		As soon as possible from 9 years of age (females only)	2 3 (if ≥ 15 years old)	
Pneumococcal (conjugate)	(Option 1)	6 weeks (minimum)	3	
	(Option 2)		2	9–15 months
Polio	bOPV + IPV	6 weeks	4 (IPV dose to be given with bOPV dose from 14 weeks)	
	IPV/bOPV sequential	8 weeks (IPV first)	1–2 IPV 2 bOPV	
	IPV	8 weeks	3	
Measles		9 or 12 months	2	
Rubella		9 or 12 months with measles containing vaccine	1	
Rotavirus	Rotarix	6 weeks (min) with DTP1	2	
	Rota Teq		3	

Recommended for Children Residing in Certain Regions				
Japanese B encephalitis	(Inactivated Vero cell derived)	6 months	2 (generally)	
	(Live attenuated)	8 months	1	
	(Live recombinant)	9 months	1	
Yellow fever		9–12 months, with measles	1	
Tickborne encephalitis		≥ 1 year: FSME-Immun and Encepur ≥ 3 years: TBE_ Moscow and EnceVir	3	At least 1 every 3 years

* Not including vaccines recommended only for children in selected high-risk populations or immunization programs with certain characteristics.
Modified from WHO. Summary of WHO Position Papers - Recommended routine immunizations for children. March 2017. Retrieved from: http://www.who.int/immunization/policy/Immunization_routine_table2.pdf?ua=1

need assistance long after acute cases of polio have been eradicated.

Killed injectable polio vaccine (IPV) and live OPV became available in the 1950s and 1960s, respectively. Both types of vaccines are safe and highly effective, and each has been used successfully to eliminate disease caused by wild-type poliovirus in high-income countries. Although there are extremely rare cases of polio caused by the vaccine strain of the virus when OPV is given, for many years WHO and other supporters of the EPI considered OPV to be preferable to IPV for routine use in LMICs. Reasons for preferring OPV included its extremely low cost (approximately $0.02 per dose); its ease of administration; its ability to induce intestinal immunity that inhibits shedding of wild-type poliovirus; and its transmission to household and other close contacts through the fecal–oral route, thereby providing repeated exposures to the vaccine and boosting immunity to polio through such contacts. However, for unclear reasons, the efficacy of OPV in LMICs has consistently been found to be lower than that in high-income countries—approximately 85% versus approximately 95% for the primary series (Patriarca, Wright, & John, 1991).

Descriptive Epidemiologic Features and Risk Factors

Before the widespread use of polio vaccine, polio was endemic in virtually all LMICs, and most children were asymptomatically infected during the first few years of life. Symptomatic acute polio was similarly seen primarily in infants and young children. Based on surveys of the prevalence of lameness in school-aged children, the annual incidence of symptomatic polio in LMICs was estimated to be in the range of 20 to 40 cases per 100,000 total population (LaForce, Lichnevski, Keja, & Henderson, 1980).

As oral polio vaccine came into widespread use and vaccine coverage increased, endemic infections with wild-type poliovirus decreased. However, vaccine coverage levels in the range of 40% to 80%, combined with a vaccine efficacy of approximately 85%, led to an accumulation of susceptible individuals and subsequent outbreaks in many countries with "good" EPI programs (Sutter et al., 1991). In the early 1980s, more than 50,000 cases of polio were being reported annually to WHO (Otten et al., 1992). However, as described in the next subsection, wild-type polio has now been eliminated in much of the world as a result of intensified surveillance and immunization efforts; as of 2017, it remained a problem in only three countries.

Current Approaches to Prevention and Control

In 1988, the WHA set a goal of interrupting polio transmission worldwide by 2000 and certifying that polio had been eradicated by 2005. The polio eradication effort put into place at that time was based

on a combination of ongoing routine immunization of infants, annual national immunization days, and house-to-house mop-up campaigns to vaccinate those persons who were missed by these other approaches (Hull, Ward, Hull, Milstien, & de Quadros, 1994). In addition, sensitive surveillance for and laboratory testing of specimens from individuals with acute flaccid paralysis were put into place to help identify cases of polio, thereby allowing targeting of additional vaccination efforts and monitoring of the impact of the eradication program.

The last confirmed case of paralytic polio caused by wild-type poliovirus in the Western Hemisphere was detected in Peru in 1991, and three WHO regions (Americas, Western Pacific, and European) had been certified as polio free by 2002 (Centers for Disease Control and Prevention [CDC], 2002). In 2015, wild poliovirus type 2 was officially declared eradicated, and no case of paralytic polio due to wild poliovirus type 3 has been detected anywhere since 2012. In 2016, a total of only 37 cases of polio caused by wild poliovirus type 1 occurred in three countries—Pakistan, Afghanistan, and Nigeria. In 2017, however, there was an outbreak of polio caused by vaccine-derived polio virus in Syria, likely due to the poor sanitation and inadequate vaccine coverage that has resulted from the destruction of the public health system in this war-torn country.

The Global Polio Eradication Initiative (GPEI) was launched in 1988, and by 2012 global polio incidence was reduced by 99%. Unfortunately, the emergence of circulating, vaccine-derived polio viruses (cVDPV)—which consist of genetically unstable OPV strain viruses that revert to the profile of the virulent parent strain—has complicated elimination, necessitating the withdrawal of OPV to complete eradication. In 2016, all countries still using trivalent oral polio vaccine successfully switched to bivalent (types 1 and 3) oral polio vaccine and committed to introducing at least one dose of inactivated polio vaccine into their routine immunization systems (Hampton et al., 2016). Ultimately, use of oral polio vaccine will cease altogether, and it will be replaced by inactivated polio vaccine.

Measles

Etiologic Agent, Clinical Features, and Vaccine Characteristics

Measles is caused by the measles virus. Although some genotypic variation occurs, all measles virus strains are considered to be of a single type. The measles virus is spread via the respiratory route and is transmitted extremely efficiently. It is highly infectious, and in the absence of vaccine-induced immunity, almost every child can be expected to develop measles if the virus is circulating in the community.

Measles is characterized initially by fever, cough, runny nose, and malaise, making it indistinguishable from many other viral respiratory infections for the first several days, during which the child is highly infectious. A characteristic rash then appears. Although most cases are self-limited, complications commonly include pneumonia, diarrhea, and ear infections. Less common complications include encephalitis and blindness. Measles is not amenable to antibiotic therapy, but treatment with vitamin A can reduce the case fatality ratio (Hussey & Klein, 1990).

Before the measles vaccine came into widespread use, measles was consistently one of the leading causes of death among children worldwide, accounting for an estimated 20% to 30% of such fatalities (Walsh, 1983). Although the estimates of the case fatality ratio for measles have varied from 1% to greater than 30%, depending in part on whether the studies were conducted in the community or were hospital based, it is clear that the most potent predictors of mortality among children with measles are young age; malnutrition, particularly vitamin A deficiency (Markowitz et al., 1989); and HIV infection. Furthermore, measles frequently leaves a child weakened and at increased risk of illness and death from other causes for a year or more after the acute episode.

The measles vaccine consists of a live, attenuated strain of the measles virus. It is safe, inexpensive, and highly effective when given to a child after circulating measles antibody acquired from the mother has disappeared. Because the maternal antibodies tend to disappear somewhat later and exposure to measles virus is substantially less common in high-income countries, the first dose of measles vaccine is typically given at 15 months of age in such countries. In LMICs, the intensity of exposure to measles and the poorly understood, more rapid decline in maternal antibodies combine to put infants at much greater risk of acquiring measles at a young age, when complications and death are more likely. As a result, the first dose of measles vaccine is typically given to infants at 9 months of age in LMICs.

Descriptive Epidemiologic Features and Risk Factors

In the absence of vaccination, every child in an area where measles virus is circulating would be expected to contract measles. In the early 1980s, an estimated 3 million children died annually of measles and its sequelae. By 2016, however, the number of measles deaths had declined to an estimated 89,780 per year (Dabbagh et al., 2017).

The age at which an unvaccinated child develops measles is a function of when maternal antibodies disappear (generally at 6 to 12 months of age) and the intensity of exposure to measles virus in the community. Thus, in crowded urban areas, most unvaccinated children will develop measles between 6 months and 5 years of age; in contrast, in more sparsely populated, rural areas, such children acquire measles at an older age (Walsh, 1983). Family size, travel patterns, and types and locations of social interactions (e.g., marketplaces) also influence the local epidemiologic features of measles. HIV infection appears to increase the risk of acquiring measles in infancy, presumably by decreasing the level of circulating maternal antibody in the infant.

Current Approaches to Prevention and Control

Like smallpox, measles can, in theory, be eradicated because the causative virus does not infect other species or live in the environment. However, because measles is more widespread and much more infectious than smallpox, and because substantial transmission of the virus occurs among infants prior to the age of routine vaccination, eradication of measles will be more difficult to achieve.

Rather than establish a goal of measles eradication, the WHA resolved in 2012 to reduce measles-related mortality by 95% by the end of 2015 and to achieve regional measles elimination in at least five WHO regions by the end of 2020. Regional Verification Committees in the American, European, and Western Pacific Regions of WHO had verified elimination of measles in 61 countries by the end of 2015. As with polio, improved routine immunization of infants and periodic mass vaccination campaigns targeting infants and children from 9 months to 5 (or even up to 14) years of age have been undertaken as complementary strategies, together with a move to ensure all children receive two doses of measles vaccine, improved surveillance, and improved case management. As a result, measles cases and deaths were reduced by 84% between 2000 and 2016 (Dabbagh et al., 2017). Despite these impressive gains, eradication of measles is unlikely in the immediate future due to immunization coverage gaps and weaknesses in health systems.

Diphtheria

Etiologic Agent, Clinical Features, and Vaccine Characteristics

Diphtheria is caused by the bacterium *Corynebacterium diphtheriae*. It is spread primarily via the respiratory route, although in LMICs the organism is also a cause of ulcerative skin lesions and can be transmitted from such lesions. The incidence of diphtheria and the morbidity and mortality attributable to it in such countries are largely unknown, but the disease is not believed to pose a major public health threat.

Diphtheria is a disease of the upper respiratory tract, manifested by fever, sore throat, an inflamed pharynx (and possibly nose and larynx), and a grayish membrane covering the inflamed mucosa. With involvement of the larynx, the airways can be blocked and death can result. The disease is toxin mediated, and use of both antibiotics and diphtheria antitoxin (which is rarely, if ever, available in LMICs) is beneficial in its treatment.

The diphtheria component of the DPT vaccine is composed of inactivated diphtheria toxin adsorbed to aluminum salts. Two or more doses of DPT result in protection against diphtheria in 90% to 100% of those vaccinated.

Descriptive Epidemiologic Features and Risk Factors

Although diphtheria is easier to diagnose than pertussis, it is likely underreported in many LMICs. Thus the 4,530 cases of diphtheria reported to WHO in 2015 almost certainly do not reflect the actual burden of disease (WHO, 2017a).

Most cases of diphtheria occur in young children, primarily among those living in impoverished, crowded conditions. Lack of vaccination is undoubtedly the most important risk factor for developing or dying from diphtheria.

Current Approaches to Prevention and Control

The current approach to reducing morbidity and mortality from diphtheria is to improve levels of vaccine coverage achieved through ongoing infant immunization programs (i.e., EPI) in various countries. Strategies for improving vaccine coverage include expanding the times when vaccinations are offered in clinics, reducing waiting times, and reducing the number of missed opportunities to vaccinate unvaccinated infants, among other options. Because vaccination does not lead to elimination of carriage of *C. diphtheriae* from the nasopharynx, ongoing control of diphtheria requires achieving and maintaining high levels of vaccine coverage among the target population.

Pertussis

Etiologic Agent, Clinical Features, and Vaccine Characteristics

Pertussis (whooping cough) is caused by the bacterium *Bordetella pertussis.* Like measles, pertussis is spread via the respiratory route and is highly contagious, particularly within a household and in crowded institutional settings. In classic cases of pertussis, nonspecific respiratory tract symptoms are followed by severe and protracted bouts of coughing that typically end with an inspiratory whooping sound. These bouts of coughing can persist for many weeks and be quite debilitating, even when complications such as pneumonia and neurologic damage do not develop. Antibiotic treatment has little or no impact on the natural course of the disease once symptoms have begun, but probably shortens the time during which an individual is infectious.

Pertussis is believed to be the cause of substantial morbidity and mortality in the absence of high levels of coverage with pertussis vaccine, but it is difficult to assess the proportion of respiratory infections and related deaths due to this cause. In part, this difficulty arises from the fact that many infants with pertussis never have the characteristic whoop seen in older children. Similarly, a growing body of evidence suggests that in high-income countries, and presumably in LMICs as well, *B. pertussis* is the cause of many cough illnesses in young adults that are never recognized as pertussis. Finally, the laboratory diagnosis of pertussis has been plagued by the insensitivity, nonspecificity, technical difficulty, and cost of the various diagnostic tests, making research studies difficult and routine surveillance extremely inaccurate.

Pertussis vaccines have, for many years, consisted of killed whole bacteria adsorbed to aluminum salts to make them more immunogenic. In most instances, pertussis vaccine is given to infants and young children as a part of the DPT vaccine. When a full series of three doses is given, the efficacy of the pertussis component of DPT is in the range of 80%. Because whole-cell pertussis vaccine contains many bacterial products, it is the most reactogenic component of DPT; in turn, the DPT vaccine is the most reactogenic of those included in the EPI. Thus, pain and tenderness at the injection site, with or without fever, are common after a DPT injection.

More serious but rare complications of whole-cell pertussis vaccine (e.g., encephalopathy) were at the center of a protracted debate in many high-income countries over its safety. Concern over these rare complications and attendant declines in vaccine acceptance led to the development and licensure of acellular pertussis vaccines that are far less reactogenic and equally efficacious, but also far more expensive than whole-cell pertussis vaccine. Because of this large difference in cost, the EPI continues to use whole-cell pertussis vaccine in almost all LMICs. Given recent evidence that the clinical protection induced by acellular pertussis vaccines wanes substantially after only three or four years, it is unlikely that LMICs will switch to using this vaccine.

Descriptive Epidemiologic Features and Risk Factors

The number of pertussis cases reported to WHO annually in the early 1980s was in the range of 1.5 to 2 million; by 1997, this number had decreased to fewer than 200,000 (WHO, 1998a). In the mid-1980s, it was estimated that more than 600,000 children died of pertussis annually. By 2008, the estimated annual death toll from pertussis had declined to approximately 89,000 (WHO, 2017b). As noted earlier, however, any estimate of the burden of disease caused by pertussis is likely to underestimate the actual toll due to the difficulty of making the diagnosis.

Limited studies suggest that most cases of pertussis in LMICs occur in infancy and early childhood and that the highest case fatality ratio is seen in infants. It is likely that poverty and the resultant crowding increase the risk of pertussis and that malnutrition increases the likelihood of dying among those who develop the disease, although lack of vaccination is undoubtedly the most important risk factor for developing or dying from pertussis at any age.

Current Approaches to Prevention and Control

As is the case for diphtheria, the current approach to reducing morbidity and mortality from pertussis is to improve levels of vaccine coverage through the EPI and strategies for improved vaccine coverage. Because vaccination does not lead to elimination of carriage of *B. pertussis,* ongoing control of pertussis requires high levels of vaccine coverage among the target population.

Tetanus

Etiologic Agent, Clinical Features, and Vaccine Characteristics

Tetanus is caused by the toxin produced by the anaerobic bacterium *Clostridium tetanus.* This bacterium is

commonly found in the gastrointestinal tract of many domesticated animals (e.g., cows, sheep, and goats). When deposited in the soil, *C. tetanus* cells form spores that are highly resistant to heat and desiccation and remain viable for years or even decades. When these spores are introduced into a wound or other suitable environment, they sporulate and the bacterial cells reproduce, forming and releasing a highly potent neurotoxin as they grow. Signs and symptoms produced by the neurotoxin include painful stiffening and spasms of the muscles, including those of the jaw (hence the common name "lockjaw"), and, particularly in newborn infants, a resultant inability to suck or otherwise feed.

Treatment of tetanus is largely supportive, even in the rare circumstance when tetanus antitoxin is available. It consists largely of giving muscle relaxants and fluids intravenously. Even with such treatment, the case fatality ratio is very high, particularly in infants with neonatal tetanus, of whom 80% to 90% will die.

The vaccine against tetanus, either in a single preparation or in DPT, is a toxoid—an inactivated form of tetanus toxin, adsorbed to aluminum salts. Tetanus toxoid (TT) is extremely safe and produces few reactions. The three doses of DPT given to infants and the two doses of TT given to women of childbearing age or pregnant women through the EPI produce immunity in 90% to 100% of those vaccinated.

Descriptive Epidemiologic Features and Risk Factors

While tetanus in adults is an avoidable and tragic illness, tetanus in newborn infants is a public health problem. Studies in the 1970s and 1980s showed that as many as 60 newborn babies per 1,000 (6%) developed neonatal tetanus in some low-income countries, primarily due to contamination of the umbilical stump, and that almost all these babies died (Stanfield & Galazka, 1984). Deaths due to neonatal tetanus accounted for one-fourth to three-fourths of all neonatal deaths and as much as one-fourth of infant mortality in these countries. Neonatal tetanus caused more than 500,000 deaths worldwide in 1993 (WHO, 1994), but that number had declined to an estimated 61,000 deaths by 2011 (WHO, 2017c).

Neonatal tetanus is more common in rural areas, particularly those where animal husbandry practices lead to substantial fecal contamination of the soil, and it tends to be more common during the rainy season (Schofield, 1986). Nevertheless, the key risk factors for neonatal tetanus clearly relate to the site where a delivery occurs, the level of training of the person assisting with the delivery, the manner in which the umbilical cord is cut, and the way the umbilical stump is treated. Because of the enormous diversity of cultural practices regarding how the cord is cut and what is placed on the umbilical stump (including mud, animal dung, clarified butter, and other nonsterile materials), the rate of and risk factors for neonatal tetanus vary substantially in different regions of the world (Stanfield & Galazka, 1984).

Current Approaches to Prevention and Control

Because the *C. tetanus* spores that cause neonatal tetanus are ubiquitous in soil and can persist there indefinitely, the organism itself cannot be eradicated. Neonatal tetanus, however, can be controlled or eliminated as a public health problem (defined as fewer than 1 case per 1,000 live births in each health district) by ensuring that a high proportion of women giving birth have received two doses of TT or that the delivery and subsequent cord care practices minimize the chance that *C. tetanus* spores will be introduced. Studies such as those conducted in Egypt (**FIGURE 6-2**) demonstrated clearly that immunization of women of reproductive age or of pregnant women has a dramatic impact on the risk of neonatal tetanus (CDC, 1996). Other studies suggest that training and equipping birth attendants so they can perform a clean delivery (the "three cleans"—delivery with clean hands, delivery on a clean surface, and use of clean instruments and dressings to cut and dress the umbilical cord) may be somewhat less effective in reducing the risk of neonatal tetanus but has a greater impact on the risk of neonatal mortality from all causes combined. These two approaches can be, and have been, used together. In recent years, the proportion of women giving birth in various geographic regions who have been adequately immunized with TT has increased substantially, but neonatal tetanus remains a problem in many LMICs where many deliveries still occur at home.

Obstacles to Prevention and Control

Despite Herculean efforts on the part of WHO, UNICEF, and others, it has been challenging to achieve and maintain high levels of vaccine coverage in many countries, particularly in rural areas within LMICs. By 2015, global coverage was at 86%. Current estimates suggest that worldwide, 19.5 million children remain unvaccinated, of whom an estimated 1.5 million will die of childhood vaccine-preventable diseases (WHO, 2018a).

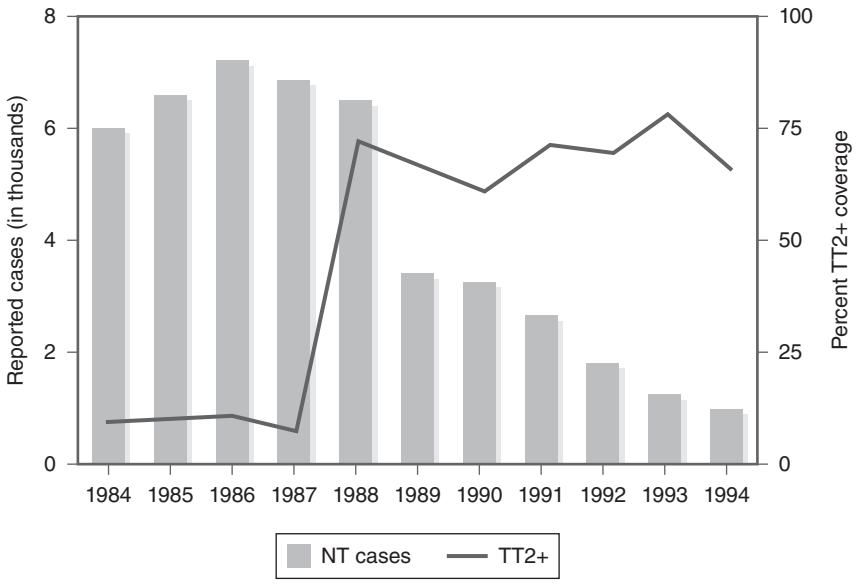


FIGURE 6-2 Number of reported cases of neonatal tetanus (NT) and percentage of pregnant women receiving at least two doses of tetanus toxoid (TT2+), by year, in Egypt, 1984–1994.

Centers for Disease Control and Prevention. (1996). Progress toward elimination of neonatal tetanus: Egypt, 1988–1994. *Morbidity and Mortality Weekly Report, 45,* pp. 89–92.

For a number of years, substantial controversy persisted regarding the relative merits of vertical approaches to vaccinating children (e.g., mass campaigns) and horizontal approaches (i.e., improving access to and use of primary care services that provide routine immunizations). Although the debate over vertical versus horizontal programs continues in some health areas, it has largely been replaced with respect to vaccination of infants and children by a broad consensus that the two approaches can be complementary, rather than conflicting. Thus, attempts to strengthen routine infant (and pregnant woman) immunization programs in LMICs around the world have proceeded in parallel with mass campaigns intended to hasten the eradication of polio and the elimination of measles. In 2012, the Global Vaccine Action Plan was implemented, with the goal of preventing millions of deaths by providing equal access to vaccines worldwide.

Further reduction or elimination of the childhood vaccine-preventable diseases discussed in this section is contingent on the availability of sustained funding of and technical support for immunization programs, stimulation of increased demand for vaccination on the part of parents and communities, expanded access to immunization services, and effective surveillance for these diseases. At the same time, there is a need to continue to develop, test, and make available new and improved vaccines; to expand local production of existing vaccines in LMICs; to ensure the potency and safety of the vaccines produced; and to ensure the availability and proper use (and disposal) of sterile injection equipment. The Global Alliance for Vaccines and Immunizations (GAVI; see **EXHIBIT 6-2**), together with the Bill and Melinda Gates Foundation and other partners, is working to ensure the earliest possible incorporation of other vaccines into the EPI. As existing and newly licensed infant and childhood vaccines are added to the group of vaccines recommended by WHO, countries with limited resources will need to establish their own priorities for which vaccines to add.

Based on the model of smallpox eradication, one argument cited to support the drive for worldwide polio eradication was that it would be possible to discontinue routine polio vaccination at some point after eradication was achieved. However, outbreaks of polio caused by circulating vaccine-derived polioviruses in Hispaniola, the Philippines, and Madagascar, together with the widespread presence of wild-type polio virus in numerous laboratories and vaccine production facilities around the world, have challenged previous notions about the need for ongoing polio vaccination after eradication (Fine, Oblapenko, & Sutter, 2004; Sutter, Cáceres, & Mas Lago, 2004). Thus, current plans call for the use of inactivated polio vaccine worldwide for at least some years following eradication of polio.

The inexplicable and unfortunate increase in all-cause mortality seen in infants (primarily female infants) given experimental high-titer measles vaccine at a young age has clearly set back efforts to develop a more potent measles vaccine (Aaby et al., 1994). Therefore, for the foreseeable future, efforts to eliminate measles will, of necessity, rely on the currently available vaccine. Important unanswered questions

EXHIBIT 6-2 The Global Alliance for Vaccines and Immunizations

An estimated 1.5 million infants and children still die each year from diseases that can be prevented by currently available vaccines. Several million more infants and children die each year from tuberculosis, malaria, and AIDS—infectious diseases for which new vaccines are urgently needed.

Numerous obstacles, many of which are external to the public health and healthcare delivery systems, make it difficult to achieve and sustain high levels of coverage with existing vaccines in LMICs. These obstacles include limited financial resources, insufficient numbers of trained healthcare workers, poor roads and other barriers to reaching remote parts of some countries, civil wars, and natural disasters. Other barriers hinder the development, testing, licensure, and ultimate availability of new vaccines, including the cost of research and testing (and hence the eventual cost of new vaccines, particularly those requiring technological sophistication, such as conjugated vaccines) and liability concerns on the part of potential manufacturers.

Recognizing these problems and seeking to increase the speed with which current and new vaccines reach the world's children, a group of international agencies—including WHO, UNICEF, the World Bank, and others—launched the Global Alliance for Vaccines and Immunizations (GAVI) in 2000. The mission of GAVI is "to fulfill the right of every child to be protected against vaccine-preventable diseases of public health concern." Its four-pronged approach includes accelerating the introduction and use of vaccines, strengthening the health systems that deliver immunizations, improving financing for immunizations, and shaping the vaccine market to make appropriate and affordable immunizations available to poor countries. To date, GAVI has been instrumental in helping make *H. influenzae* b conjugate vaccine, rubella-containing vaccine, pentavalent (DPT–Hib–hepatitis B) vaccine, and pneumococcal conjugate vaccine available to millions of children in poor countries, and it is currently working to speed the introduction of vaccines against rotavirus and HPV, among others (Loharikar et al., 2016).

relate to how best to use the currently available vaccine and which vaccination strategies or combinations of strategies will be most effective in interrupting transmission of this highly infectious agent, particularly in Africa. There is a clear need to sustain high levels of routine immunization against measles at or around 9 months of age and a second dose at some time after the first birthday. The relative importance of periodic mass campaigns remains to be determined. It is, at best, uncertain whether a new measles vaccine that is immunogenic at a younger age (in the face of circulating maternal antibodies) and that is safe can be developed and tested, or whether it is even needed to eliminate measles.

Historically, rubella vaccine was not included in the EPI package of vaccines, even though it has been available for 30 years and is both highly effective and very safe. However, because of growing recognition that congenital rubella syndrome is a problem in LMICs, rubella-containing vaccine had been included in the national immunization schedules in 149 countries as of 2016 (Loharikar et al., 2016).

In recent years, there has been growing interest in giving additional vaccines (i.e., other than tetanus toxoid) to pregnant women, in an effort to protect them (e.g., against influenza) and to protect their newborn babies prior to the time they can be vaccinated against pertussis or influenza (Omer, 2017). At the same time, there is ongoing work to develop other vaccines that might be administered to pregnant women to protect their newborn babies, such as a vaccine against group B *Streptococcus*.

Enteric Infections and Acute Respiratory Infections

Although it might seem odd to discuss enteric infections and acute respiratory infections (ARIs) together in one section, these seemingly disparate conditions have much in common. Each accounts for a substantial amount of childhood morbidity and mortality, as well as for a large proportion of outpatient visits and hospitalizations. Infants and young children almost uniformly experience multiple episodes of both types of illness, regardless of where they live. The identified risk factors for enteric infections and ARIs (e.g., poverty, crowding, lack of parental education, malnutrition, low birth weight, and lack of breastfeeding) overlap substantially, and most are difficult to change in the absence of major social change. Further, both enteric infections and ARIs are caused by a multitude of distinct microbial agents, for most of which no vaccine currently exists or is likely to be available in the near future. As a result, the overall approach to minimizing morbidity and mortality from both enteric infections and ARIs has been virtually identical—acceptance of the fact that

such infections and illnesses will occur, combined with attempts to ensure that prompt and appropriate care is sought and given.

Enteric Infections

Overview

Enteric infections encompass those viral, bacterial, and parasitic infections of the gastrointestinal tract that are, with the exception of typhoid fever, generally manifested as diarrhea, either alone or in combination with fever, vomiting, and abdominal pain. Although most episodes of diarrheal disease are mild and self-limited, the loss of fluids and salts accompanying severe diarrhea can be life threatening. Also, not all episodes of diarrheal disease are self-limited: Studies suggest that anywhere from 3% to 23% of diarrheal illnesses in infants and young children persist for longer than two weeks (Black, 1993). Both self-limited and persistent diarrhea can have a substantial negative impact on the growth of a child, through malabsorption of nutrients and reduced intake due to vomiting, loss of appetite, and undesirable changes in feeding practices in response to diarrheal illness. Thus, repeated episodes of diarrhea and persistent diarrhea often lead to malnutrition, which can in turn increase the likelihood of diarrhea persisting and producing a fatal outcome (El Samani, Willett, & Ware, 1988). It has been estimated that diarrheal disease has a more profound impact on the growth of children worldwide than any other infectious disease (see the *Nutrition* chapter).

Cholera, while in a sense just one of many causes of watery diarrhea, is in many ways a disease unto itself. It can produce the most dramatic fluid losses of any enteric infection and, in the absence of appropriate replacement of fluid and salts, can cause death within 24 to 48 hours of its onset. Cholera epidemics and pandemics can produce enormous numbers of cases and large numbers of deaths, resulting in profound social disruption. As a result, cholera has been accorded a special status by public health officials and agencies.

Many episodes of diarrheal disease in children in LMICs are accompanied by bloody stools or frank dysentery (abdominal cramps; painful, strained defecation; and frequent stools containing blood and mucus), which is usually the result of an invasive infection that produces local tissue damage and inflammation in the intestinal mucosa. Although the fluid loss that accompanies such episodes is generally not profound, life-threatening local intestinal and systemic complications can result. Damage to the intestinal mucosa can also lead to substantial losses of protein, resulting in growth retardation. The clinical management of such episodes poses a number of distinctive challenges (see "Current Approaches to Prevention and Control" later in this section).

On average, children younger than the age of 5 years in LMICs experience two to three episodes of diarrhea each year (Kosek, Bern, & Guerrant, 2003). The burden of disease attributable to diarrheal diseases worldwide is enormous, despite the impressive accomplishments of diarrheal disease control programs. The Global Burden of Disease (GBD) project estimated that more than 2.3 billion episodes of diarrhea occur annually and that diarrheal disease causes approximately 1.3 million deaths each year, 499,000 of which occur in children younger than 5 years of age (GBD 2015 Disease, 2016). Thus, diarrheal disease accounts for roughly 8.6% of all deaths among children in this age group (GBD 2015 Mortality, 2016).

Typhoid fever, which results from an enteric infection and, therefore, shares many individual- and community-level risk factors with diarrheal disease, is not accompanied by diarrhea. Although it can be life threatening, this disease has its most profound public health impact through its debilitating effects on school-age children, causing substantial morbidity and absenteeism from school and work (Medina & Yrarrazaval, 1983).

Etiologic Agents

As noted earlier, diarrheal disease can be caused by a wide variety of viral, bacterial, and parasitic infections. In cases of endemic diarrheal disease, one or more etiologic agents can be identified in 70% to 80% of patients when state-of-the-art laboratory testing is performed. However, many of these agents can also be found in the stools of children who do not have diarrhea, and multiple infectious agents may be present in the same child with diarrhea. Thus, the presence of a given causative agent in a stool sample may not mean that it is the cause of that episode of diarrhea. Recent studies have attempted to address this issue through various methods, including the use of quantitative detection methods and testing samples from control children without diarrhea to assess the causal role of various microbial agents (Liu et al., 2016; Platts-Mills et al., 2015).

These complications notwithstanding, the most important etiologic agents in young children in LMICs are rotavirus, enterotoxigenic *Escherichia coli* (ETEC), enteropathogenic *Escherichia coli* (EPEC), *Shigella* spp., *Campylobacter* spp., *Cryptosporidium* spp., and norovirus (Lanata et al., 2013; Liu et al., 2016; Platts-Mills et al., 2015). Rotavirus is the leading cause of

nonbloody diarrhea in infants and accounts for an estimated 38% of hospital admissions for diarrhea among children younger than 5 years of age (Lanata et al., 2013), whereas *Shigella* species and *Campylobacter* species appear to be the leading causes of bloody diarrhea (Platts-Mills et al., 2015). Amebiasis (infection with *Entamoeba histolytica*), although frequently diagnosed, appears to be an infrequent cause of bloody diarrhea in young children in LMICs. *Vibrio cholerae* is the cause of a substantial proportion of cases of nonbloody diarrhea in endemic areas, such as Bangladesh and India (Cholera Working Group, 1993; Nair et al., 1994); it is also the cause of large numbers of cases when epidemics of cholera occur in other regions, as demonstrated by recent or ongoing large outbreaks in Haiti, Mozambique, and Yemen (Luquero et al., 2016). Typhoid fever is caused by *Salmonella typhi*.

Descriptive Epidemiologic Features and Risk Factors

Children experience the highest risk of diarrheal illness between 6 and 11 months of age; the risk declines steadily thereafter (**FIGURE 6-3**) (Kosek et al., 2003). This age pattern is largely explained by the established risk factors for diarrheal disease and the likely sources of exposure to the causative agents.

The risk of diarrheal disease in young infants is determined in part by the feeding and hand-washing practices of the mother or other childcare providers. Breastfeeding and lack of exposure to contaminated food, water, and other environmental sources are protective factors. As infants grow and become mobile, however, they begin to encounter numerous potential sources of infection with the agents of diarrheal disease, including contaminated water and weaning foods, as well as human and animal wastes that have not been disposed of properly. Some evidence also indicates that uncontrolled fly populations can contribute to the risk of diarrheal illnesses—particularly those illnesses caused by etiologic agents requiring a small infectious dose (e.g., shigellosis). Overall, it is estimated that almost 90% of deaths due to diarrheal disease are attributable to unsafe water, inadequate sanitation, and poor hygiene (Black, Morris, & Bryce, 2003).

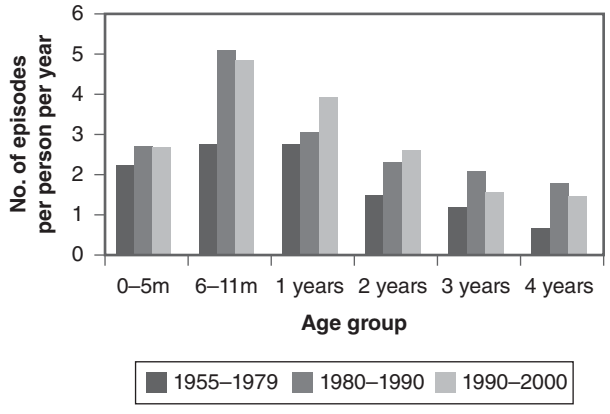


FIGURE 6-3 Median age-specific incidences for diarrheal episodes per child per year from three reviews of prospective studies in developing areas, 1955–2000.

Reproduced from Kosek, M., Bern, C., & Guerrant, R. L. (2003). The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bulletin of the World Health Organization, 81,* pp. 199.

As noted earlier, diarrheal disease and malnutrition are intricately intertwined in infants and young children. Although it is not clear that malnutrition is associated with an increased incidence of diarrheal disease, strong evidence indicates that malnutrition increases the likelihood that a child with diarrhea will die or develop persistent diarrhea; diarrhea, in turn, has a negative impact on nutritional status and growth. Furthermore, strong evidence supports the contention that deficiencies of vitamin A and zinc are associated with increased morbidity and mortality from diarrheal disease.

Unlike diarrheal disease, typhoid fever has long been thought to be a problem primarily in school-age children. However, infections with *S. typhi* clearly occur in infants and preschool-age children, and such infections in this younger age group may be substantially underreported (Sinha et al., 1999). More than 20 million new cases of typhoid fever are believed to occur worldwide each year, with the highest incidence seen in South and Southeast Asia (Mogasale et al., 2014; WHO, 2015a). *S. typhi* infection is primarily acquired from contaminated food and water.

Current Approaches to Prevention and Control

A number of approaches to the primary prevention of diarrheal illness have been studied and found to be successful at lowering the rate of diarrheal illness in the community. These measures include water, sanitation, and hygiene interventions; promoting breastfeeding and proper weaning practices; providing vitamin A and zinc supplementation; vaccines for rotavirus and cholera; and various combinations of these interventions (Bhutta et al., 2013; Clasen et al., 2015). Although historical approaches focused on structural changes to improve communities' water supplies, it has become clear that interventions targeting water quality (i.e., those that protect water at the point of use) are more effective in preventing diarrhea than those focused on improving water supply.

Although these approaches have been shown to be effective in well-funded research studies, the feasibility

of fostering and sustaining such improvements across large areas and populations is directly linked to the availability of the financial resources and the political will to do so. Many children still lack access to these interventions (Bhutta et al., 2013). In addition, such interventions can generally succeed only when the community is invested in making them work and when they are designed and introduced within a culturally acceptable framework (see the *Culture, Behavior, and Health* chapter).

In parallel with the testing of such primary prevention approaches to reduce the risk of diarrheal disease, since the early 1980s enormous efforts have been directed toward ensuring that infants and children who develop diarrhea suffer a minimum of morbidity and mortality. These efforts have focused on prevention and early treatment of dehydration through proper case management in the home, maintenance of adequate nutritional intake to minimize the impact on growth, and appropriate treatment of infants and children who are brought to health facilities. Mothers and other caretakers of children are educated to use fluids available in the home to prevent dehydration, to use oral rehydration solutions (ORS) or cereal-based alternative solutions to prevent and treat dehydration, and to continue to breastfeed babies when they have diarrhea. ORS packets are now manufactured in many countries and widely available.

At the same time, mothers and healthcare providers have been discouraged from using the wide range of largely ineffective, often expensive, and sometimes dangerous antimicrobial and antidiarrheal agents that are available. Care providers should reserve these antimicrobial agents for the treatment of cholera and of dysentery suspected to be caused by shigellosis and limit the use of intravenous fluids to those who are severely dehydrated. Although promoting and sustaining such changes in diverse countries, often in the face of sometimes powerful cultural barriers, requires a major effort, studies from Egypt and elsewhere suggest that morbidity and mortality from diarrheal disease can be reduced substantially by such efforts (El-Rafie et al., 1990).

Perhaps in part because of the emphasis given to primary prevention and proper case management for diarrheal disease, the role of vaccines in reducing morbidity and mortality from enteric infections has been quite limited until recently. Vaccines against cholera and typhoid that represent substantial improvements over older vaccines are now available (Sur, Lopez, et al., 2009; Sur, Ochial, et al., 2009), as are newly licensed vaccines against rotavirus that have been shown to be efficacious in Africa, Latin America, and elsewhere (Madhi et al., 2010; Richardson et al., 2010). Although a few countries in which typhoid fever is a major public health problem in school-age children have begun to administer typhoid vaccines to members of this age group, a longstanding debate has focused on the proper role of vaccines against typhoid, cholera, and other enteric infections and the circumstances under which they should be given routinely in LMICs (Keusch & Cash, 1997). The availability of these new vaccines is leading to renewed discussion of how best to use typhoid and cholera vaccines (Levine, 2009; Sridhar, 2009), and rotavirus vaccines had been introduced in the EPI of more than 80 countries by 2016.

Obstacles to Prevention and Control and Directions for Future Research

The greatest obstacle to reducing further the toll taken by diarrheal disease in LMICs is the difficulty and expense of ensuring that everyone has regular access to clean, safe drinking water and adequate sanitation. To do so requires both structural and behavioral changes. In addition, cultural and legal adjustments are needed to create a more permissive and supportive environment for breastfeeding. Effective interventions are available that could result in a substantial reduction in diarrhea-associated morbidity and mortality, but scaling up existing interventions and ensuring equitable access remain significant challenges.

Acute Respiratory Infections

Overview

Acute respiratory infections comprise infections of various parts of the respiratory tract, ranging from mild viral and bacterial infections of the upper respiratory tract (e.g., the common cold, viral and group A streptococcal pharyngitis, and middle ear infections) to life-threatening infections of the lower respiratory tract (e.g., bronchiolitis and pneumonia caused by a variety of bacterial and viral agents). Although upper respiratory tract infections globally cause substantial minor morbidity and economic loss through lost time at work, they rarely result in severe morbidity or in mortality. Interestingly, studies suggest that the incidence of upper respiratory tract infections, although varying with age and season, is remarkably similar in free-living populations throughout the world. Because they do not pose a major public health problem and because no effective interventions against them exist, upper respiratory tract infections will not be discussed further in this chapter, except for group A streptococcal pharyngitis, which can lead to rheumatic fever and is discussed later in this section.

Lower respiratory tract infections, in contrast, are the cause of substantial morbidity and mortality, particularly among infants and young children living in low- and middle-income settings, even when those settings are within industrialized countries. In LMICs, particularly those with good childhood immunization and oral rehydration therapy programs, lower respiratory tract infections in general, and pneumonias in particular, are typically one of the leading causes of death among infants and children younger than 5 years of age (GBD 2015 Child, 2016). An estimated 703,000 children younger than 5 years of age died in 2015 from pneumonia and other lower respiratory tract infections, most of them in LMICs (GBD 2015 Mortality, 2016). However, the accuracy of such estimates is questionable for several reasons: Establishing a diagnosis of lower respiratory tract infection or pneumonia is difficult, particularly in settings where chest radiography is not available; many infants and young children die outside of hospitals; and other illnesses, such as malaria, make verbal autopsies (i.e., postmortem interviews of next-of-kin to determine the most likely cause of death) an unreliable means of establishing a definitive diagnosis.

Although pneumonia and other lower respiratory tract infections also cause substantial morbidity and mortality in older children and adults, this section focuses on these infections in infants and young children. Pneumonia in adults is discussed briefly later, as is influenza, an important cause of lower respiratory tract infections in children and adults.

Etiologic Agents, Clinical Features, and Vaccine Characteristics

Lower respiratory tract infections can be caused by a variety of viral and bacterial agents, either singly or in combination. Numerous studies conducted in various countries around the world show similar results concerning the etiologic agents responsible for these infections in infants and young children. Excluding respiratory tract infections caused by agents included in the EPI vaccines, the most important viral causes of lower respiratory tract infections are influenza, parainfluenza, respiratory syncytial virus (RSV), and adenovirus (Pavia, 2011). The most important bacterial causes of pneumonia, as determined by lung aspirate studies (in which a needle is passed through the chest wall into the affected lung parenchyma, thereby avoiding contamination of samples by flora in the upper airway), are *Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Staphylococcus aureus* (Shann, 1986). Many infants and young children have evidence of dual infections (e.g., a virus and a bacterium). Moreover, in as many as one-third of the cases, no etiologic agent can be found, despite the use of state-of-the-art laboratory techniques.

Descriptive Epidemiologic Features and Risk Factors

Infants and young children living in low- and middle-income environments consistently have been found to experience high incidence rates of pneumonia. However, over the last two decades, there has been a significant reduction in pneumonia-related deaths in young children, likely due to a combination of improved access to health care, antibiotic therapy, and preventive vaccines (GBD, 2013; GBD 2015 Mortality, 2016). Despite this reduction, in 2015, 12.1% of deaths in children younger than 5 years of age were due to pneumonia and other forms of lower respiratory tract infections (GBD 2015 Mortality, 2016).

The single most important predictor of a child's risk of developing pneumonia or other lower respiratory tract infection is age. The cumulative incidence of lower respiratory tract infections is highest among young infants, but drops rapidly with increasing age, reaching markedly lower levels among children by the time they reach 2 or 3 years of age (Selwyn, 1990). Another important predictor of an increased risk of morbidity and mortality from lower respiratory tract infections is low birth weight. Other risk factors for either morbidity or mortality include exposure to indoor air pollution (from cooking, heating, and cigarette smoke), not breastfeeding, and malnutrition, including vitamin A deficiency. These factors are closely intertwined with poverty, inadequate access to health care, and with one another, so their independent effects can be difficult to disentangle (Berman, 1991). HIV infection is another important risk factor (Madhi, Peterson, Madhi, Khoosal, & Klugman, 2000).

Rheumatic fever following group A streptococcal pharyngitis occurs in a setting of poverty and household crowding. School-age children are primarily affected acutely, but the damage done to heart valves by this disease is usually permanent, producing lifelong disability. Population-based data concerning the incidence of acute rheumatic fever and the prevalence of rheumatic heart disease in LMICs are rarely available.

Current Approaches to Prevention and Control

In the 1980s, WHO, together with various partners, began a multifaceted research program intended to develop an approach to reducing the substantial

morbidity and mortality due to lower respiratory tract infections in infants and young children (WHO, 1981). This research program and the ARI control program that was subsequently developed were premised on the following observations:

- Upper respiratory tract infections, although frequent, are almost always benign and require only supportive care at home.
- At the time of the program's inception, many lower respiratory tract infections in infants and young children were not preventable with existing vaccines.
- The major known risk factors for morbidity and mortality from lower respiratory tract infections (e.g., age, low birth weight, malnutrition, and indoor air pollution) are impossible or difficult to change.
- Most morbidity and mortality from lower respiratory tract infections occur in locales where access to medical care is limited and where there are few, if any, diagnostic facilities (i.e., the ability to perform chest x-rays, microbiologic cultures, and other tests).

Given these circumstances, it was decided that an ARI control program based on triage performed by minimally trained village health workers according to readily observable clinical signs might be feasible, inexpensive, and effective at reducing at least mortality. As a result, a large body of multidisciplinary research relating to the various aspects of such a control program was commissioned and completed. Particularly important was research examining which readily observable clinical manifestations (e.g., cough, fever, respiratory rate, and chest indrawing), singly or in combination, best distinguished infants or young children with various levels of severity of respiratory tract infection (initially classified as mild, moderate, and severe, but later as no pneumonia, pneumonia, and severe pneumonia).

Based on this research, intervention programs were developed. These efforts were intended to train village health workers or their equivalents in how to assess and classify into one of these categories an infant or young child with signs of a respiratory tract infection. Based on their assessment, the village health workers were to recommend supportive care at home in cases of mild ARI (no pneumonia), provide an oral antimicrobial drug (either ampicillin or cotrimoxazole) and education about home care and follow-up in cases of moderate ARI (pneumonia), or refer the child immediately to the nearest hospital for assessment and inpatient care in cases of severe ARI (severe pneumonia). Well-designed intervention trials were carried out in a variety of countries to assess the efficacy of this approach in reducing mortality due to lower respiratory tract infections. As a meta-analysis of these trials showed, almost all of them successful in reducing mortality due to lower respiratory tract infections and all-cause mortality in infants and in children age 1 to 4 years (Sazawal & Black, 1992). One caveat should be noted: These trials typically assessed an intervention that included regular household visits by the village health workers in search of infants and children with signs of a respiratory tract infection.

Based on these favorable results, WHO promoted and supported the implementation of ARI control programs largely based on the model of having village health workers (or their equivalent) assess infants and young children with suspected ARI, classify them by severity of illness, and treat or refer them. The impact of such programs on mortality, as distinct from the impact seen in the intervention trials, has not been adequately assessed. On the one hand, these ARI control programs typically do not include a proactive, outreach component that ensures early case detection (i.e., regular household visits), as was present in the intervention trials, but instead rely solely on maternal recognition of illness and appropriate, timely care seeking; thus, they are not likely to have as large an impact as reported in the earlier trials. On the other hand, programs that include some form of regular household visits are likely to be difficult to sustain over the long term.

For a variety of reasons, the emphasis given to ARI control programs (and vertical disease-specific programs in general) by WHO and others has diminished in recent years, and attention has shifted to ensuring that any sick infant or child, regardless of his or her signs and symptoms, receives appropriate evaluation and care. This approach, referred to as the Integrated Management of Childhood Illness (IMCI), is described in **EXHIBIT 6-3** (WHO, 2012a). **TABLE 6-3** summarizes the interventions included in the IMCI program (WHO, 2012a). A recent systematic review and meta-analysis found that IMCI may reduce overall mortality in infants and children (relative risk [RR]: 0.85; 95% confidence interval [CI]: 0.78–0.93), but did not find consistent improvements in quality of care or vaccination coverage (Gera, Shah, Garner, Richardson, & Sachdev, 2016).

Prevention of rheumatic fever depends on the recognition of streptococcal pharyngitis and treatment with an appropriate antimicrobial agent (e.g., penicillin or erythromycin). In populations with high rates of acute rheumatic fever, school-based and other programs to detect and treat streptococcal pharyngitis have been suggested as a means to combat this disease,

EXHIBIT 6-3 Integrated Management of Childhood Illness

Throughout the 1970s and 1980s, concerted efforts were made to develop case management strategies for each of the infectious diseases that collectively accounted for the majority of morbidity and mortality among infants and young children in LMICs—measles, malaria, diarrheal disease, and acute respiratory infections. Although each of these disease-specific case management strategies has been shown to be effective at reducing severe morbidity and mortality when properly implemented, implementing multiple distinct disease-specific programs that all target the same healthcare providers can lead to overlap, inefficiency, and competition for the attention of an overworked healthcare worker. As a result, various programs within WHO and UNICEF collaborated in the development of the Integrated Management of Childhood Illness (IMCI), now referred to as the Integrated Management of Newborn and Childhood Illness (IMNCI) in some countries. This strategy attempts to assemble into a single, more efficient program the approaches of the various disease-specific control programs (Gove, 1997).

The IMCI program comprises three main components: improving case-management skills of healthcare workers, improving overall health systems, and improving family and community health practices, especially as they relate to the illnesses that collectively account for the majority of severe morbidity, mortality, and healthcare provider visits among infants and young children (see Table 6-3) (WHO, 2016b).

TABLE 6-3 Child Health Interventions Included in Integrated Management of Childhood Illness

Case Management Interventions	Preventive Interventions
Pneumonia	Immunization during sick child visits (to reduce missed opportunities)
Diarrhea	Nutrition counseling
Dehydration	Breastfeeding support (including the assessment and corrections of breastfeeding technique)
Dysentery	
Persistent diarrhea	
Meningitis, sepsis	
Malaria, measles, malnutrition, anemia	
Ear infection	

Data from World Health Organization (WHO). (2012). Recommendations for management of common childhood conditions.

although their effectiveness has rarely been assessed (Bach et al., 1996; Lennon et al., 2017).

Obstacles to Prevention and Control and Directions for Future Research

Primary prevention of lower respiratory tract infections, while in part requiring improvements in living conditions and socioeconomic status, is increasingly possible due to the growing availability of affordable vaccines against some of the major etiologic agents. The safe, highly effective conjugate vaccine against *H. influenzae* type b, which led to the virtual disappearance of invasive infections caused by this organism in the United States and other high- and middle-income countries, is now recommended by WHO for inclusion in the EPI of all countries. In turn, funding from GAVI resulted in the introduction of this vaccine in all GAVI-eligible countries by 2014. A number of studies conducted in diverse populations have documented a substantial reduction in meningitis and/or pneumonia following introduction of Hib conjugate vaccine (Hammitt et al., 2016; Mulholland et al., 1997; Pilishvili et al., 2013; Scott et al., 2013).

Similarly, conjugate pneumococcal vaccines including the most important serotypes of *S. pneumoniae* have been tested and shown to have excellent efficacy and safety. A conjugate pneumococcal vaccine including the seven serotypes that collectively accounted for approximately 80% of invasive infections in infants in the United States was approved for use in that country in 2000 and subsequently in other wealthy countries; its widespread use had a profound impact on the rate of invasive pneumococcal infections among infants in high-income countries. Conjugate pneumococcal vaccines with an expanded number of serotypes (e.g., 11 or 13) are now available, and funding from GAVI has made them affordable for 57 of the world's poorest countries to introduce. When (or, indeed, if) vaccines against the various viruses (other than

influenza) that are important causes of lower respiratory tract infections in infants and small children (e.g., RSV) will become available is uncertain.

Several barriers exist to reducing morbidity and mortality due to lower respiratory tract infections through means other than vaccination and improvement of living conditions. The first set of barriers relates to ensuring that the parents or guardians of a sick child will seek and have access to appropriate care in a timely fashion. A detailed discussion of care-seeking practices and obstacles to obtaining medical care is beyond the scope of this chapter, but suffice it to say that the obstacles are multifaceted and difficult to overcome. Many challenges also arise in ensuring that those ill infants and children who are brought to medical care facilities receive timely and appropriate care, including adequate assessment, treatment, and follow-up. Finally, concerns have focused on the likelihood that increased use of currently effective and inexpensive antimicrobial agents—particularly use of inappropriate or inadequate regimens—may lead to the development of resistant strains of bacteria (particularly *S. pneumoniae*), which may then not respond to currently available inexpensive, oral treatment regimens. Indeed, all WHO regions have reported decreased susceptibility of *S. pneumoniae* to penicillin, with some reports showing that more than 50% of samples have reduced susceptibility (WHO, 2014a).

Bacterial Meningitis

Overview

Meningitis is a nonspecific term that encompasses inflammation of the meninges (the membranous lining that covers the brain and spinal cord), which can be caused by a wide variety of infectious and noninfectious agents. Such inflammation, regardless of its cause, tends to produce a similar clinical picture—headache, stiff neck, fever, and variable other features. There is substantial overlap between the manifestations of meningitis, which is sometimes referred to as spinal meningitis or cerebrospinal meningitis, and those of many other infectious diseases. Although meningitis can be caused by a wide variety of viruses and other infectious agents (e.g., mycobacteria and parasites), the disease caused by certain bacteria poses a substantial public health threat. Therefore, this section is confined to a discussion of bacterial meningitis, excluding tuberculous meningitis, which is discussed briefly later in the section on tuberculosis.

From a public health perspective, it is important to subdivide bacterial meningitis into its endemic and epidemic forms. Endemic bacterial meningitis, while differing in a number of subtle ways with regard to its descriptive epidemiologic features and the distribution of etiologic agents in LMICs versus high-income countries, poses a similar set of challenges in these two different settings. Epidemic bacterial meningitis, for reasons that remain unexplained, has almost entirely ceased to be a public health problem in industrialized countries since World War II, although small clusters of cases or hyperendemic disease can still be a vexing problem. In contrast, in a number of LMICs, particularly those in the "meningitis belt" of sub-Saharan Africa, periodic epidemics of bacterial meningitis continue to occur on a scale never documented in high-income countries. These epidemics can be of such a magnitude and geographic scope as to be properly called public health disasters.

In high-income countries, suspected bacterial meningitis is considered a medical emergency, requiring appropriate clinical specimens for diagnostic testing be obtained and parenteral antimicrobial therapy in a hospital to be initiated immediately. Even under these ideal conditions, case fatality ratios for bacterial meningitis range from 3% to 25%, with the ratio depending primarily on the specific etiologic agent and the age of the patient. Further, many patients who survive the acute episode will be left with one or more serious sequelae, including deafness, blindness, mental retardation, and seizure disorders. Although the clinical outcomes of hospitalized cases of bacterial meningitis do not, in general, differ between LMICs and high-income countries, the resources available for treating such patients are much more limited in LMICs. Furthermore, epidemics involving tens of thousands of such cases cannot be dealt with easily by countries with extremely constrained health budgets and facilities.

Endemic Meningitis

Etiologic Agents

Studies of the etiology of endemic bacterial meningitis in LMICs have been hampered by the need for a reasonably well-equipped and staffed microbiology laboratory to conduct such studies. Even so, a number of hospital-based studies have been performed in areas where or in time periods when epidemic meningitis has not been present. These studies are in general agreement that the leading causes of endemic bacterial meningitis in these settings are *S. pneumoniae, H. influenzae,* and *Neisseria meningitidis*; before the introduction of *H. influenzae* b and pneumococcal conjugate vaccines, these same pathogens were responsible for most cases of bacterial meningitis in

high-income countries. Reviews have documented the important role of *H. influenzae* b and *S. pneumoniae* in bacterial meningitis (and other syndromes) in LMICs (O'Brien et al., 2009; Watt et al., 2009). Other organisms that account for a reasonable proportion of cases in high-income countries, such as group B *Streptococcus* and *Listeria monocytogenes*, appear to be infrequent causes of meningitis in LMICs, although this apparent difference may be artifactual. At the same time, meningitis due to *Salmonella* species appears to be more common in LMICs than in high-income countries.

Descriptive Epidemiologic Features and Risk Factors

The overall cumulative incidence of endemic bacterial meningitis in LMICs appears to be four or five times that in high-income countries, although the data are limited. Somewhere between 1 in 60 and 1 in 300 children die of bacterial meningitis before the age of 5 years in nonepidemic areas (Greenwood, 1987). Endemic bacterial meningitis is primarily a problem in infants and young children, although age-specific incidence rates vary with the etiologic agent. Notably, meningitis caused by *H. influenzae* b occurs almost exclusively during the first 12 to 24 months of life. Although the highest incidence rates of meningitis due to *S. pneumoniae* and endemic meningitis due to *N. meningitidis* occur in the first 12 to 24 months of life, cases can also occur in older children and adults. Endemic bacterial meningitis probably occurs at approximately equal incidence rates among males and females.

Because the three leading causes of endemic bacterial meningitis are all spread via respiratory droplets, poverty and the resulting crowding increase the risk for this disease. Failure to breastfeed has been shown to be a risk factor in high-income countries and probably increases the risk in LMICs as well. Host factors also play an important role in determining the risk of endemic bacterial meningitis—most notably sickle cell disease and HIV infection. Sickle cell disease is associated with a markedly increased risk of infection with *S. pneumoniae*. Meningitis due to *S. pneumoniae*, like that caused by *Salmonella*, also occurs at a substantially increased rate among HIV-infected children and adults. In addition, malnutrition and anemia have been suspected to be risk factors for bacterial meningitis.

Current Approaches to Prevention and Control

Until recently, little could be done to prevent endemic bacterial meningitis in LMICs. Today, however, bacterial meningitis due to *H. influenzae* b and many serotypes of *S. pneumoniae* is preventable with vaccines produced by conjugating the bacterial polysaccharide to one of several protein molecules. The widespread use of these vaccines has led to the virtual elimination of meningitis due to *H. influenzae* b and to sharp reductions in the incidence of meningitis due to *S. pneumoniae* in high-income countries. Funding from GAVI is now making these vaccines more widely available in poor countries (see Exhibit 6-2). Although a portion of the cases of endemic meningitis attributable to *N. meningitidis* could be prevented with existing purified polysaccharide and conjugate vaccines against serogroups C, Y, and W-135 and more recently developed vaccines against serogroup B, these vaccines are in use in only a few LMICs (e.g., Brazil and Cuba are routinely using a vaccine against serogroup B *N. meningitidis*).

Chemoprophylaxis—that is, giving a short course of an antimicrobial agent to individuals in close contact with someone with bacterial meningitis due to *N. meningitidis* or *H. influenzae* type b—has been used with some success in high-income countries, but has not been widely advocated in LMICs because of the small percentage (less than 5%) of endemic cases that occur in close contacts of a known case, the cost and availability of the antimicrobial agents, the limited duration of the protection achieved, concerns about promoting the development of antimicrobial resistance, and logistical problems related to implementation.

Obstacles to Prevention and Control and Directions for Future Research

The primary obstacle to preventing endemic meningitis due to *H. influenzae* b and *S. pneumoniae* in LMICs has been the cost of the highly effective and safe conjugate vaccines now used in high-income countries. As economic barriers to the introduction of these vaccines in LMICs are removed, cases of endemic meningitis due to these organisms may be markedly reduced or nearly eliminated. Prevention of endemic meningococcal meningitis in most LMICs must await the widespread availability of serogroup B and C conjugate vaccines that are effective in infants, provide protection against multiple serotypes of serogroup B, and are affordable.

Epidemic Meningitis

Etiologic Agents

Epidemic bacterial meningitis is almost always caused by *N. meningitidis*. More specifically, serogroup A *N. meningitidis* causes most such epidemics, although

epidemics due to serogroup C and serogroup W-135 have also been well documented (Sidikou et al., 2016; WHO, 2013). Serogroup B *N. meningitidis*—one of the most important causes of endemic bacterial meningitis—has led to "epidemics" in a variety of high-income and LMICs, but these outbreaks are never of the scope and intensity of those caused by serogroup A, differing in overall attack rates by as much as two orders of magnitude.

Descriptive Epidemiologic Features and Risk Factors

Epidemic bacterial meningitis, generally caused by serogroup A *N. meningitidis,* is one of the most interesting but least understood infectious diseases. Epidemics involving hundreds of thousands of cases and cumulative incidence rates of almost 1,000 to 2,000 cases per 100,000 total population have been observed in the "meningitis belt" of Africa (**FIGURE 6-4**) for more than 100 years (Moore, 1992). An epidemic there in the 1990s was one of the largest ever recorded and extended into parts of Africa not previously considered to be in the meningitis belt. In this region of Africa, epidemics in a given area typically last for 2 or 3 years and recur every 5 to 15 years. They occur only during the hot, dry season (January to April) and dissipate when the rains and cooler weather arrive, only to return the next dry season. The interepidemic period can vary from a few years to a decade or more, likely reflecting a combination of the time it takes to reaccumulate enough susceptible individuals to sustain an epidemic, the introduction of a new virulent strain of *N. meningitidis*, prior use of meningococcal vaccine in the population, and other poorly defined factors (Moore, 1992).

Epidemics of meningococcal meningitis have also occurred in Asia and Latin America. In western China and Nepal, the epidemics have followed a pattern similar to that seen in Africa, except that they occur during the cold, dry season rather than the hot, dry season. In Latin America, Brazil has borne the brunt of such epidemics due to both serogroups A and C. Epidemics have also occurred in countries in the Middle East (e.g., Saudi Arabia) and the Pacific (e.g., New Zealand).

Current Approaches to Prevention and Control

Historically, efforts to reduce morbidity and mortality from epidemic meningococcal meningitis were reactive in nature. That is, once an epidemic was detected, a vaccination campaign was implemented as rapidly as possible, and the antimicrobial agents and other materials needed to treat cases appropriately were made available in the affected area. This approach was rightly criticized on the grounds that numerous delays arose when mounting such campaigns, during which time many cases (and resulting morbidity and mortality) occurred. Such reactive vaccination campaigns were almost inevitably disruptive of other health programs, and the extent to which they reduced the size of the epidemic was often

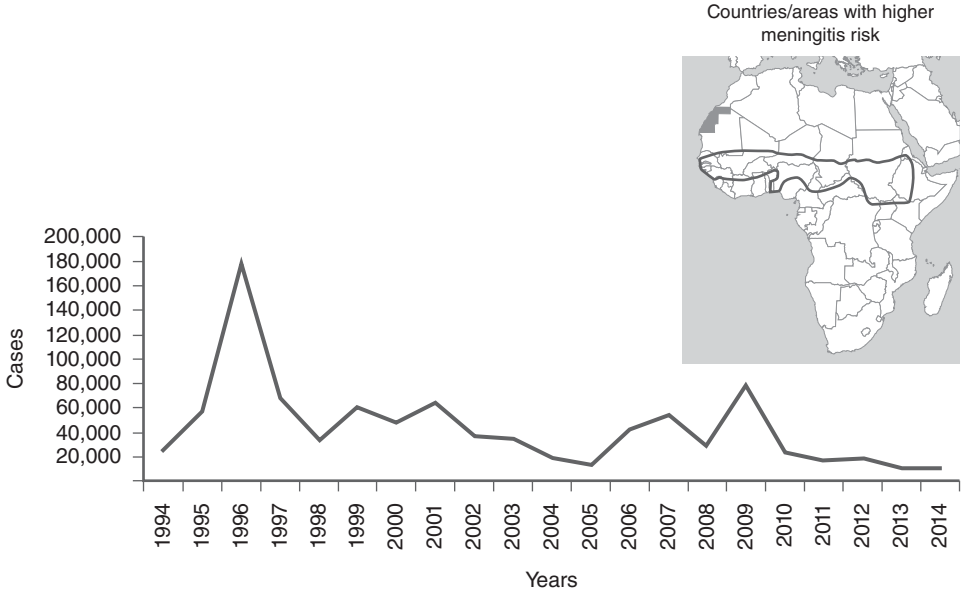


FIGURE 6-4 Trends of epidemic meningitis cases in the African Belt countries, 1994–2014.

Reproduced from World Health Organization. (2014). Global health observatory (GHO) data. Retrieved from http://www.who.int/gho/epidemic_diseases/meningitis/en/

debated. Furthermore, shortages of the purified polysaccharide vaccine used in these campaigns often occurred (Wakabi, 2009).

One alternative approach to improving the control of such epidemics that was suggested in the 1990s was to enhance surveillance for bacterial meningitis in areas susceptible to epidemics (e.g., the meningitis belt of Africa), use a predetermined threshold rate of cases to declare an epidemic and mount a vaccination campaign, and stockpile vaccines and other supplies and equipment in the immediate area (WHO, 1995). Unfortunately, projections suggest that even these steps resulted in the prevention of no more than 40% to 50% of the cases that would otherwise have occurred, and initial attempts to establish such early detection and response capabilities demonstrated the limitations of this approach.

Recognizing the limitations of an approach to controlling or preventing epidemic meningococcal meningitis epidemics using the purified polysaccharide vaccines against serogroup A, a group of vaccine manufacturers and philanthropic organizations developed a monovalent (i.e., serogroup A only) conjugate vaccine (MenAfriVac) using an approach similar to that employed in making the *H. influenzae* b and *S. pneumoniae* vaccines used to successfully immunize infants. MenAfriVac was introduced into countries in the "meningitis belt" beginning in 2010, targeting individuals 1 to 29 years of age. By the end of 2015, it had been introduced into 16 countries in sub-Saharan Africa. Since then, no epidemics of meningitis caused by serogroup A *N. meningitidis* have occurred in the region (WHO, 2013). Continued use of the vaccine in infants when they reach 1 year of age may help sustain this remarkable achievement.

Obstacles to Prevention and Control and Directions for Future Research

While use of MenAfriVac has, thus far, prevented the recurrence of epidemics of meningitis caused by serogroup A *N. meningitidis* in the countries in Africa's "meningitis belt," epidemics caused by both serogroup W-135 and serogroup C have occurred in Niger, Burkina Faso, and other countries in recent years (MacNeil et al., 2014; Sidikou et al., 2016). Thus, even if widespread use of MenAfriVac in the region continues to prevent outbreaks caused by serogroup A *N. meningitidis*, it may be necessary to add conjugate vaccines (or vaccine components) that provide protection against other serogroups of *N. meningitidis* to adequately control or prevent large-scale epidemics of meningitis in this region in the future.

Mycobacterial Infections

Overview

Although many species of mycobacteria can infect people, only two cause sufficient human illness in LMICs to warrant discussion here—*Mycobacterium tuberculosis*, the cause of tuberculosis (TB), and *Mycobacterium leprae*, the cause of leprosy. Although *Mycobacterium bovis* (which is closely related to *M. tuberculosis*) can cause TB in humans, it accounts for only a small percentage of cases, and these cases are generally not distinguishable from or in need of different treatment than cases caused by *M. tuberculosis.* Other than *M. leprae*, the various nontuberculous mycobacteria cause opportunistic infections that occur almost exclusively in immunocompromised individuals. Whereas these nontuberculous mycobacteria—particularly *Mycobacterium avium* complex—have caused substantial morbidity and mortality among AIDS patients in high-income countries, they appear to be uncommon in AIDS patients in LMICs.

Tuberculosis and infection with *M. tuberculosis* are, by every indicator available, among the most important public health problems in the world. Approximately one-third of the world's population (2 billion persons) is believed to be infected with *M. tuberculosis*, and in 2015 there were 10.4 million new cases of TB and 1.8 million deaths from TB, including 400,000 deaths of people with HIV infection (WHO, 2016c). Because suppression of the body's immune system is the most important determinant of which individuals infected with *M. tuberculosis* will subsequently develop TB, the AIDS epidemic has had disastrous consequences for the control of TB, which was underfunded and inadequate in most LMICs even before the arrival of HIV. In recognition of the gravity of the problems posed by this infection, WHO declared TB to be a global emergency in 1993.

A disease that many believe to have been leprosy was described in the Old Testament of the Bible. Leprosy occupies a unique position among human diseases, in part because of the disfigurement that it can produce and in part because of the belief in many cultures that it represents some form of divine punishment. Although leprosy was endemic to Europe during the eleventh through thirteenth centuries, it had virtually disappeared from there by the eighteenth century, long before modern medicine arrived. There is speculation that the rise of TB and infection with *M. tuberculosis* produced cross-immunity to *M. leprae* and led to the disappearance of leprosy from Europe. Some support for this theory comes from

the observation that BCG—the vaccine intended to prevent tuberculosis—appears to be at least partially effective in preventing leprosy. Whatever the cause of its near-total (but not complete) absence from high-income countries, leprosy today is largely confined to a shrinking number of LMICs.

Tuberculosis

Etiologic Agent, Clinical Features, and Vaccine Characteristics

Compared with other bacteria, mycobacteria are slow growing and have special nutritional needs. *M. tuberculosis* can be recovered from clinical specimens, particularly those from the respiratory tract, when appropriate artificial media and techniques are used, but the process takes many weeks and requires a laboratory with a modest level of sophistication and resources. Furthermore, it can be difficult to obtain a specimen of respiratory tract secretions, particularly from a child. Although monitoring of treatment progress for TB still relies on smear microscopy and culture, the availability of new technologies—including molecular tests such as Xpert MTB/RIF, whose use was endorsed by WHO in 2010—is dramatically changing the diagnostic landscape for TB. However, the relatively high cost of the latest technologies compared to smear microscopy has limited adoption of these technologies in some LMICs (Pai & Schito, 2015).

Descriptive Epidemiologic Features and Risk Factors

M. tuberculosis is spread via respiratory droplets produced when an individual with active pulmonary TB, particularly smear-positive TB (in which more organisms are present in the sputum), coughs or sneezes. Individuals in close contact with a person with untreated TB—particularly household contacts—are at highest risk of becoming infected. In LMICs, the highest incidence rates of pulmonary TB occur in men and women of reproductive age, meaning that there are often infants and young children in the household in close contact with individuals with active pulmonary TB. As a result, a high proportion of individuals will be exposed to and infected with *M. tuberculosis* in childhood.

Although a small proportion of infected infants and children will develop pulmonary or extrapulmonary TB soon after becoming infected, in most instances the immune system successfully walls off, but does not kill, all of the *M. tuberculosis* organisms. As these infected children grow up, various factors—particularly HIV infection and malnutrition—can reduce the ability of the immune system to keep the organisms in check, and reactivation TB can occur. Before the AIDS epidemic, it was estimated that 5% of persons infected as children developed TB soon after infection and another 5% developed TB at some point later in life. However, untreated HIV infection is such a potent inhibitor of cell-mediated immunity (the part of the immune system that holds *M. tuberculosis* infection in check) that a high proportion of untreated, dually infected individuals (i.e., infected with both HIV and *M. tuberculosis*) can be expected to develop TB unless they die of something else first or receive preventive therapy.

Almost 2 million people die from TB and 8 to 10 million people experience the onset of the disease each year, with 95% of these deaths and illnesses occurring in LMICs (WHO, 2016c). Although TB is a major public health problem in almost every LMIC, the burden of disease attributable to it varies by region and country, in part because of longstanding historical differences in the incidence of TB and the prevalence of infection with *M. tuberculosis*, in part because of local differences in the adequacy of TB control programs, and in part because of differences in the extent of the HIV/AIDS epidemic. Based on admittedly incomplete passive surveillance for reported cases of TB and on tuberculin skin test surveys, it appears that sub-Saharan Africa has the highest average annual risk of infection with *M. tuberculosis* and the highest crude incidence rate of cases of TB (**FIGURE 6-5**). The largest numbers of cases of TB are seen in India, Indonesia, China, Nigeria, Pakistan, and South Africa, which account for 60% of TB cases globally (WHO, 2016c). Because of the HIV/AIDS epidemic and unrelated demographic changes (e.g., population growth and increases in the numbers of individuals surviving to their thirties and forties), it has been projected that the number of new TB cases will increase in coming years in LMICs worldwide.

Although TB occurs in individuals in all socioeconomic strata, it is quintessentially a disease of poverty. Through its effect on crowding, poverty increases the risk of airborne transmission of *M. tuberculosis*. At the same time, through its negative effect on nutritional status, poverty increases the risk that someone infected with *M. tuberculosis* will develop TB. HIV infection is also more prevalent among poor populations, further increasing the likelihood that an infected individual will develop TB. Finally, through its effect on access to curative medical care, poverty increases the likelihood that a patient with symptomatic pulmonary TB will remain untreated and hence infectious for a longer period of time, which in turn increases the incidence of TB in LMICs.

Current Approaches to Prevention and Control

TB control programs exist in almost all LMICs, although enormous variability exists in terms of the resources at those countries' disposal and their effectiveness. The current approach to controlling TB in LMICs is based on a strategy of rapid detection of and provision of effective multidrug therapy to all infectious persons (i.e., patients with pulmonary TB, particularly smear-positive patients).

Since 1995, a key component of WHO recommendations has been directly observed treatment, short-course (DOTS), which involves the use of a standardized multidrug short-course regimen, with direct observation of drug ingestion for at least the first two months of treatment. In 2014, WHO adopted the End TB Strategy (**TABLE 6-4**), which includes three key pillars:

- Integrated patient-centered care and prevention
- Bold policies and supportive systems
- Intensified research and innovation (WHO, 2014b)

In the face of the limited resources available to ensure prompt diagnosis and treatment of infectious cases of TB among those who present spontaneously to health facilities, there has been an understandable reluctance in LMICs to devote scarce resources to actively searching for other cases in the households of affected individuals or in the community. Similarly, prophylactic treatment of latent TB infection in the general population has not been considered a high priority or an effective use of limited resources. In addition, the use of a single drug, such as isoniazid, to reduce the likelihood of TB in individuals infected with *M. tuberculosis* (as practiced in a number of high-income countries) has raised the specter of inadvertent single-drug therapy of patients with unrecognized TB and resultant promotion of the development of drug-resistant strains of *M. tuberculosis.* However, preventive therapy is effective in preventing TB among HIV-infected individuals in LMICs (Whalen et al., 1997). Thus, WHO recommends isoniazid preventive treatment for HIV-positive persons living in areas with a high prevalence of *M. tuberculosis* infection (i.e., where more than 30% of the population has latent *M. tuberculosis* infection) and for those with latent *M. tuberculosis* infection or exposure to an infectious case of pulmonary TB.

FIGURE 6-5A Estimated tuberculosis incidence rates, 2015.

World Health Organization. (2016). Global tuberculosis report 2016. Geneva, Switzerland. Reprinted with permission. http://www.who.int/tb/publications/global_report/en/

FIGURE 6-5B Estimated HIV prevalence in new and relapse tuberculosis cases, 2015.

World Health Organization. (2016). Global tuberculosis report 2016. Geneva, Switzerland. Reprinted with permission. http://www.who.int/tb/publications/global_report/en/

TABLE 6-4 Post-2015 Global Tuberculosis Strategy Framework

Vision	■ A world free of TB ■ Zero deaths, disease, and suffering due to TB
Goal	End the global TB epidemic
Milestones for 2025	■ 75% reduction in TB deaths (compared to 2015) ■ 50% reduction in TB incidence rate (fewer than 55 TB cases per 100,000 population) ■ No affected families facing catastrophic costs due to TB
Targets for 2035	■ 95% reduction in TB deaths (compared with 2015) ■ 90% reduction in TB incidence rate (approximately 10 TB cases per 100,000 population) ■ No affected families facing catastrophic costs due to TB

Principles

1. Government stewardship and accountability, with monitoring and evaluation
2. Strong coalition with civil society organizations and communities
3. Protection and promotion of human rights, ethics, and equity
4. Adaptation of the strategy and targets at the country level, with global collaboration

Pillars and Components

1. Integrated, patient-centered care and prevention
   A. Early diagnosis of TB, including universal drug-susceptibility testing and systematic screening of contacts and high-risk groups
   B. Treatment of all people with TB, including those with drug-resistant TB, and patient support
   C. Collaborative TB/HIV activities and management of comorbidities
   D. Preventive treatment of persons at high risk and vaccination against TB
2. Bold policies and supportive systems
   A. Political commitment, with adequate resources for TB care and prevention
   B. Engagement of communities, civil society organizations, and public and private care providers
   C. Universal health coverage policy and regulatory frameworks for case notification, vital registration, quality and rational use of medicines, and infection control
   D. Social protection, poverty alleviation, and actions on other determinants of TB
3. Intensified research and innovation
   A. Discovery, development, and rapid uptake of new tools, interventions, and strategies
   B. Research to optimize implementation and impact and promote innovations

Immunization at birth with BCG is a standard part of the EPI in every LMIC, and BCG coverage is high in almost every such country. Infant immunization with BCG appears to be quite effective at reducing the risk of disseminated TB (e.g., tuberculous meningitis) in infants and children, but the vaccine's efficacy against pulmonary TB in this age group is probably no better than 50% to 60%. Infant immunization with BCG probably has, at best, only a modest effect on the risk of developing pulmonary TB as an adult; consequently, it has little or no impact on the spread of *M. tuberculosis* in the community.

Obstacles to Prevention and Control and Directions for Future Research

The rapid growth of multidrug-resistant TB (MDR-TB; i.e., TB resistant to the first-line anti-TB drugs isoniazid and rifampicin) and the emergence of extremely drug-resistant TB (XDR-TB; i.e., TB resistant to isoniazid, rifampicin, fluoroquinolones, and kanamycin, capreomycin, or amikacin) have complicated TB control efforts. Both MDR-TB and XDR-TB require lengthy treatment with second-line drugs, which are both more expensive and more toxic than the first-line drugs. Although WHO supports universal access to treatment of drug-resistant TB, scaling up of treatment availability is not always possible, due to lack of laboratory capacity to diagnose the drug-resistant form of the disease and lack of funds for both capacity building and treatment. Capacity to detect MDR-TB, however, has been greatly enhanced with the rollout of Xpert MTB/RIF.

Both the numbers of deaths and the incidence rate of TB have declined since 2000. At present, the principal obstacles to reducing TB-related morbidity and mortality in LMICs are economic and operational in nature. Increased financial and technical assistance will be needed in many countries to ensure that currently available strategies for controlling TB are fully implemented.

At the same time, further research is needed in multiple areas if control of TB is to be achieved and sustained. Development of new drugs that are effective against *M. tuberculosis*, affordable and safe, and able to clear the infection with a shorter duration of treatment is a high priority, and many candidates are currently in the development pipeline. Similarly, development of a vaccine against *M. tuberculosis* that can prevent pulmonary TB is a high priority, with more than a dozen candidate vaccines currently in clinical trials. To support the development of an effective vaccine, more research is needed to establish immunologic correlate(s) of protection, and there is a great need for a human challenge model to help with the evaluation of vaccine candidates.

Leprosy

Etiologic Agent, Clinical Features, and Vaccine Characteristics

M. leprae cannot be grown on artificial media. In research laboratories, it can be isolated and propagated in the foot pad of a mouse or in armadillos, but these techniques have no relevance to diagnosing

leprosy. Thus, the diagnosis of leprosy is made on clinical grounds, together with histopathologic examination of tissue biopsy material. *M. leprae* grows even more slowly than *M. tuberculosis*. Its slow growth is highly relevant to the control and prevention of leprosy because of the consequent need to treat those who are infected for prolonged periods of time (months to years) and the resulting difficulty of ensuring compliance with antimicrobial therapy long after the individual feels well.

Descriptive Epidemiologic Features and Risk Factors

Although it is clear that prolonged, close contact (e.g., living in the same household) with someone with untreated leprosy—and particularly with someone having a high burden of organisms (i.e., multibacillary leprosy)—is associated with a substantially increased risk of acquiring the disease, the routes of transmission are ill defined. It is assumed that transmission occurs primarily through skin-to-skin contact or exposure to respiratory tract (e.g., nasal) secretions. Environmental or animal reservoirs of *M. leprae* are thought to have little or no role in human infections. The exceedingly long incubation period for leprosy, which is believed to be years to decades, makes any study of transmission very difficult.

Leprosy is primarily a disease of poverty. Even within a single relatively homogeneous community, it is disproportionately seen among the lowest-income members. In the past, leprosy has often been described as being more common in rural than in urban populations and as having an association with proximity to water (e.g., lakes) or humidity. Even so, clear differences in the prevalence of leprosy between neighboring communities are not well explained.

In the mid-1980s, the number of cases of leprosy worldwide was estimated to be between 10 and 12 million. The incidence of leprosy was 4 to 6 cases per 1,000 population, and the prevalence in affected countries often exceeded 10 individuals per 1,000 population. By 1998, control efforts had reduced the number of prevalent cases by more than 90%, to an estimated 829,000 cases worldwide (WHO, 1998b). This decline has continued, with WHO (2016d) reporting that there were 211,973 newly detected cases in 2015.

Although leprosy previously existed throughout the world, and a handful of individuals living in high-income countries such as the United States still develop leprosy each year (a few of whom have never traveled outside the United States), leprosy is now largely confined to a shrinking number of countries. In 2015, 14 countries accounted for 94% of new cases (WHO, 2016d).

Current Approaches to Prevention and Control

Early attempts to control leprosy were based on case finding and prolonged (i.e., multiyear) or lifetime treatment with dapsone, which had the advantages of being inexpensive and having few side effects. However, this approach failed to control leprosy, at least in part because *M. leprae* developed resistance to dapsone and because it was difficult to ensure ongoing patient compliance with treatment over many years.

In the early 1980s, multidrug treatment with two or three effective antimicrobial agents (depending on the stage of leprosy) was introduced and has produced a greater than 90% reduction in leprosy cases. WHO had set a goal of eliminating leprosy as a public health problem (i.e., achieving a prevalence of 1 case or fewer per 10,000) by the year 2000; that goal was met at the global level in 2000, but has not yet been met at the country level. Further progress in reducing leprosy burden has been hampered by reduced funding for treatment and control efforts. In addition, use of disease prevalence as an indicator of program success was recognized as problematic due to the lengthy incubation period of leprosy. Currently, progress is monitored using the annual number of new cases diagnosed.

In a renewed effort to combat this disease in 2015, WHO adopted the Global Leprosy Strategy, 2016–2020. Its goal is to further reduce the global and local leprosy burden through early detection of cases and prompt treatment to reduce disability and transmission (WHO, 2016d, 2016e).

Obstacles to Prevention and Control and Directions for Future Research

Because of its extremely long incubation period and the fact that infected individuals may transmit *M. leprae* for substantial periods of time prior to becoming symptomatic and receiving treatment, it is likely that incident cases of leprosy will continue to occur in substantial numbers for years to come, even if the current leprosy control strategies substantially reduce transmission (Meima, Smith, van Oortmarassen, Richardus, & Habbema, 2004).

The role, if any, for a vaccine in the immunotherapy or the prevention of leprosy remains uncertain. Numerous studies—both experimental and observational in nature—have suggested that BCG vaccine, which is given primarily to prevent various

forms of disseminated TB in children, offers some protection against leprosy. A meta-analysis performed in 2006 showed an overall protective effect of 61% in observational studies and an overall protective effect of 26% in experimental studies, providing further evidence that BCG is modestly effective against leprosy, and that two doses of BCG are more effective against leprosy than a single dose (Setia, Steinmaus, Ho, & Rutherford, 2006). Whether other candidate mycobacterial vaccines will offer even greater protection than BCG remains uncertain, as does their role, if any, in leprosy control.

Sexually Transmitted Infections and AIDS

Overview

Sexually transmitted infections have historically been one of the most neglected areas of medicine and public health in LMICs. Until the advent of the HIV/AIDS epidemic in the 1980s, remarkably little attention was paid to STIs in such countries, despite the fact that they collectively cause enormous morbidity, loss of productivity, and infertility, and result in substantial healthcare expenditures. The fact that human immunodeficiency virus—the cause of AIDS and all its attendant morbidity and mortality—is transmitted sexually, together with the fact that the sexual transmission of HIV is facilitated by the presence of other STIs, has focused attention on and brought an infusion of resources (albeit still inadequate) into this long-neglected area.

The myriad challenges confronting the treatment and control of STIs in LMICs are multifaceted and complex. Among the most daunting are the difficulty of changing human sexual behavior; the asymptomatic nature of many STIs, particularly among women; the lack of simple, inexpensive diagnostic tests for such infections; and the lack of readily accessible, inexpensive, easy-to-administer, single-dose treatment regimens for most STIs.

Etiologic Agents

A variety of viruses and bacteria can be transmitted sexually (**TABLE 6-5**), including some (e.g., HBV) that are not traditionally grouped with other STIs. In this chapter, HBV is discussed with the other viruses that cause hepatitis. Many infectious agents that are transmitted sexually can also be transmitted via contaminated blood or injection equipment, and vertically from a mother to her newborn infant.

Of the infectious agents transmitted sexually, some initially cause ulcerative lesions, primarily on or near the genitalia (e.g., *Herpes simplex, Treponema pallidum*, and *Haemophilus ducreyi*), others initially cause urethral or vaginal discharge (e.g., *Neisseria gonorrhoeae*, *Chlamydia trachomatis*, and *Trichomonas vaginalis*), and others cause only systemic manifestations (e.g., HIV). HPV, selected subtypes of which cause genital warts, plays an important role in the pathogenesis of cervical dysplasia and carcinoma, as well as in carcinoma of the anus, penis, and the head and neck, but HPV infections are initially silent. Asymptomatic or minimally symptomatic infection with many of the sexually transmitted agents is common, greatly exacerbating the problem of interrupting transmission and reducing the prevalence and incidence of infection.

Many of the agents transmitted sexually produce not only acute symptoms referable to the lower genital tract, but also other, often more serious manifestations. For example, untreated infections with *N. gonorrhoeae* and *C. trachomatis* can ascend through the genital tract and produce pelvic inflammatory disease, tubal infertility, and ectopic pregnancy. Untreated primary syphilis in young adults can lead to life-threatening cardiac and neurologic complications due to tertiary syphilis years later. In addition, congenital syphilis—the result of infection of a baby at the time of birth—produces profound systemic manifestations. As noted earlier, infection with HPV of selected subtypes is strongly associated with subsequent cervical dysplasia and cervical cancer, as well as other cancers. Finally, HIV causes profound damage to the host immune system and a virtually 100% case fatality ratio in the absence of treatment with antiretroviral drug regimens.

Descriptive Epidemiologic Features and Risk Factors

Data concerning the incidence of STIs in LMICs are considered highly unreliable due to substantial underdiagnosis and underreporting. According to one estimate, 357 million new adult cases of curable STIs (e.g., gonorrhea, chlamydia, syphilis, and trichomoniasis) occurred worldwide in 2012, the vast majority of which were in LMICs (**FIGURE 6-6**) (Newman et al., 2015).

Although difficult to estimate, the prevalence of viral STIs, including infections involving Herpes simplex type 2 (HSV2), HPV, and HIV, is high. In 2012, there were an estimated 417 million people infected with HSV2 and 291 million women with HPV infections (WHO, 2016f). The passive surveillance for AIDS in place in most LMICs is also not

TABLE 6-5 Sexually Transmitted Infections of Importance in Low- and Middle-Income Countries

Agent	Disease	Vertical Transmission to Newborn Babies	Transmission by Blood Products
Bacteria			
Treponema pallidum	Syphilis	Yes	Yes
Neisseria gonorrhoeae	Gonorrhea, pelvic inflammatory disease	Yes	Yes
Chlamydia trachomatis	Cervicitis, urethritis, lymphogranuloma venereum, pelvic inflammatory disease	Yes	No
Haemophilus ducreyi	Chancroid	No	No
Viruses			
Human immunodeficiency virus (HIV)	AIDS	Yes	Yes
Herpes simplex	Genital herpes	Yes	No
Human papillomavirus (HPV)	Genital warts, cervical dysplasia and cervical carcinoma, penile and anal carcinomas, and head and neck cancers	Yes	No
Hepatitis B virus[a]	Acute and chronic hepatitis, cirrhosis, and hepatocellular carcinoma	Yes	Yes
Other			
Trichomonas vaginalis	Vaginitis	No	No

[a] Discussed in the section on viral hepatitis.

very sensitive or specific, making it necessary to estimate the prevalence of HIV infection and the incidence of AIDS cases. The most recent estimate is that there were 36.7 million persons in the world living with HIV/AIDS in 2016, including 1.8 million people who became infected in that year (UNAIDS, 2017a). Sub-Saharan Africa continues to have the highest burden, with 70% of all people living with HIV residing there and nearly two-thirds of new infections occurring there (**FIGURE 6-7**) (UNAIDS, 2016a; WHO, 2016g). An estimated 35 million people had died of AIDS since the beginning of the epidemic and 14 million children had been orphaned (UNAIDS, 2016a, 2017a). The vast majority of AIDS-related deaths have occurred in LMICs, particularly sub-Saharan Africa.

The prevalence of various STIs, particularly HIV, has been investigated in many LMICs. The groups typically examined are those who can be studied easily and inexpensively, such as commercial sex workers, patients being treated for STIs or TB, injection-drug users, pregnant women or women giving birth, and blood donors. The prevalence of HIV infection among commercial sex workers is up to 10 times greater than the prevalence among the general population (UNAIDS, 2017b). Encouragingly, the prevalence of HIV infection among pregnant women in sub-Saharan Africa decreased

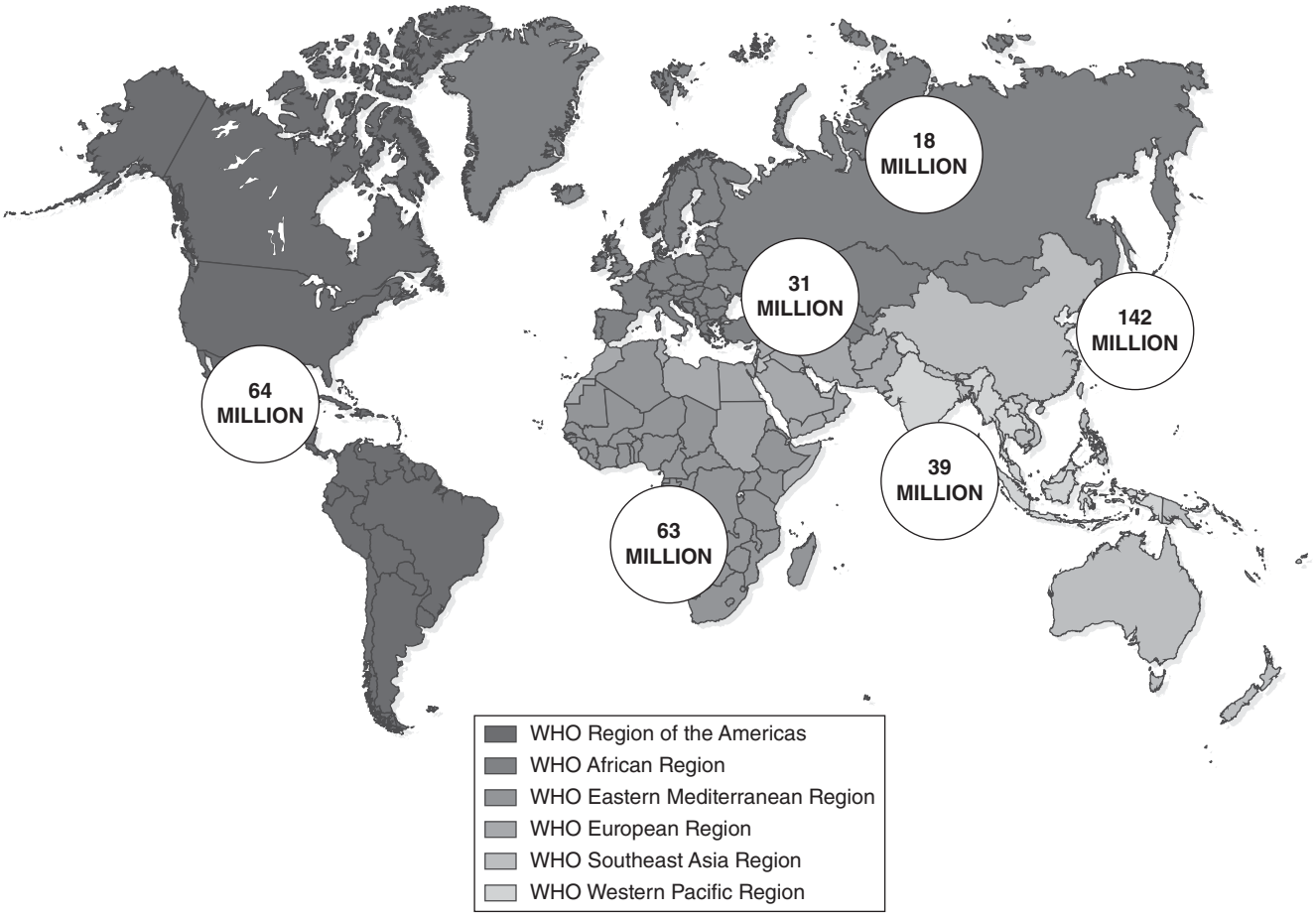


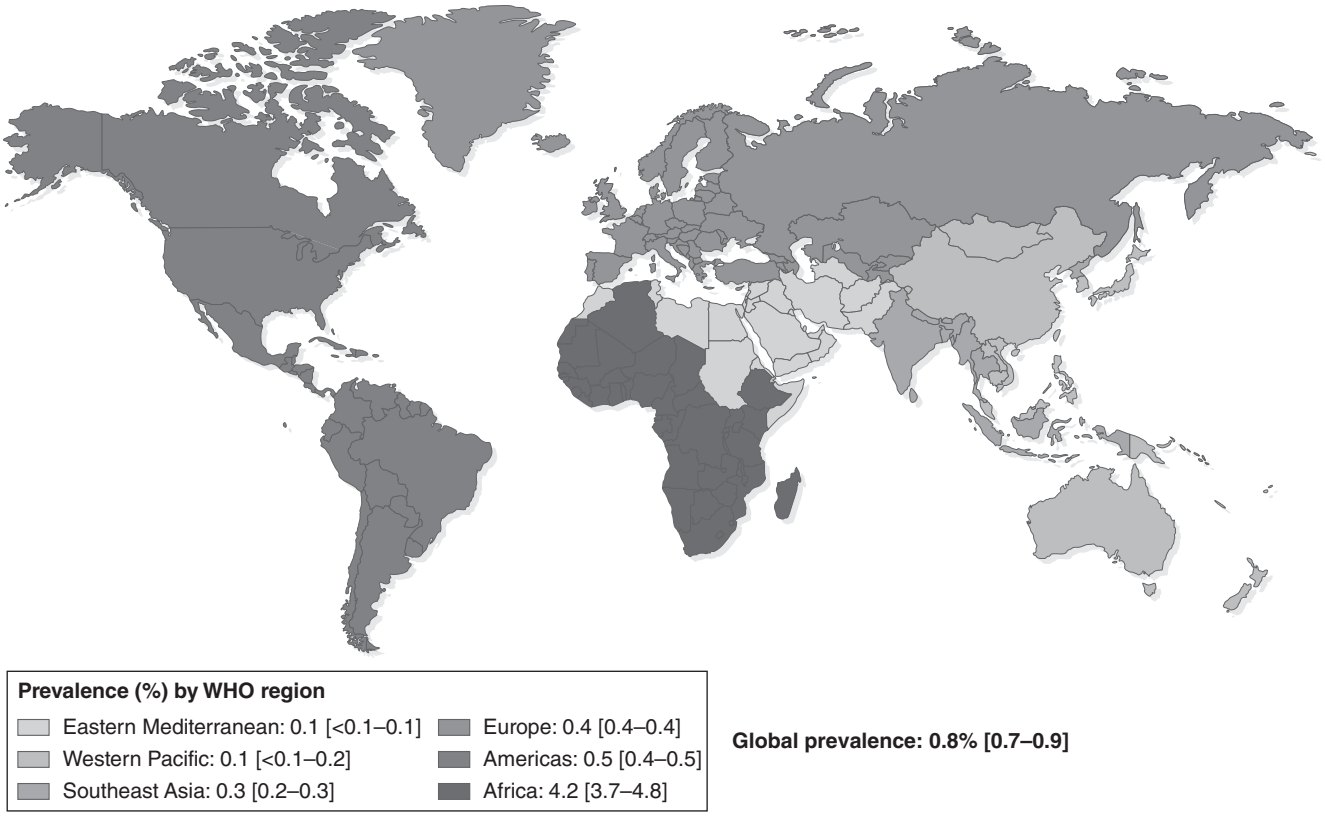
FIGURE 6-6 Estimated new cases of curable STIs, 2012.

Reprinted from World Health Organization. (2016). Global health sector strategy on sexually transmitted infections 2016–2021. http://apps.who.int/iris/bitstream/10665/246296/1/WHO-RHR-16.09-eng.pdf?ua=1

by 18% between 2003 and 2012 (Eaton et al., 2014); however, high prevalence of HIV infection is still seen in pregnant women in the worst-affected areas of sub-Saharan Africa. Because an estimated 15% to 45% of untreated HIV-infected pregnant women will transmit the virus to their newborn babies at or soon after delivery, large numbers of HIV-infected infants are also found wherever many women of reproductive age are infected. The result of widespread HIV infection among men and women of reproductive age and among infants in the most severely affected countries, in the absence of antiretroviral treatment (ART), was a reversal of prior gains in life expectancy, increases in infant and child mortality, an enormous increase in the number of orphaned children, and projected future declines in population size. However, better access to ART dramatically reduced the number of HIV-infected infants and led a 48% reduction in AIDS-related deaths globally between 2005 and 2016 (UNAIDS, 2016b, 2017a).

Given that the etiologic agents under discussion are transmitted through sexual contact, it is not surprising that the highest incidence and prevalence of these infections are seen among those men and women who are most sexually active, typically those 15 to 49 years of age. In addition, the most important risk factors for STIs—number of sexual partners, concurrency of sexual partnerships, type of sexual partners, and whether barrier protection (e.g., a male condom) is used—are directly or indirectly related to the likelihood of exposure to one of these etiologic agents. Thus, individuals with large numbers of sexual partners (e.g., commercial sex workers) who do not use barrier protection and those who have unprotected sex with such individuals, as well as those with multiple concurrent sexual partners, are at highest risk. In societies in which it is considered acceptable for men to frequent commercial sex workers or have multiple concurrent sexual partners while women are expected to have a single partner, many married women acquire STIs from their husbands. Other risk factors for STIs, including HIV, have also been demonstrated, such as lack of male circumcision.

It is important to note the complex interplay between HIV infection and other STIs. There is strong evidence that infection with both ulcerative



The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the world health organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

FIGURE 6-7 HIV prevalence in adults (15–49 years of age) by WHO region, 2016.

Reprinted from World Health Organization. (2017). Prevalence of HIV among adults ages 15 to 49, 2016 by WHO region. http://www.who.int/gho/hiv/hiv_013.jpg?ua=1

and non-ulcerative STIs increases the likelihood of HIV being transmitted sexually between partners, either through the presence of disrupted mucosa or due to the presence in the genital tract of increased numbers of inflammatory cells and lymphocytes that can bind HIV.

Current Approaches to Prevention and Control

The current approach to the prevention and control of STIs, including HIV, focuses on improving the availability of and access to high-quality diagnostic and treatment services (especially in the case of HIV/AIDS), changing sexual practices through education and health promotion, increasing the availability and use of barrier methods that reduce transmission (e.g., male and female condoms), encouraging male circumcision, shortening the time between the onset of symptoms and the seeking of appropriate care, reducing the stigma attached to STIs, and improving surveillance. Not surprisingly, given the enormous cultural diversity that exists around the world, the approaches to delivery of risk-reduction messages (e.g., school-based programs, billboards, and radio), the groups that are targeted (e.g., the entire population; individuals of reproductive age; school-age children; or high-risk groups, such as commercial sex workers, migrant workers, truck drivers, and adolescent girls), and the messages that are delivered (e.g., abstinence before marriage, monogamy, and condom use) have varied greatly. Additionally, while voluntary medical male circumcision has proved to be effective in reducing the acquisition of a number of STIs, including HIV, and could substantially reduce HIV transmission, scaling up the delivery of male circumcision remains a challenge for a number of reasons (Tobian, Kacker, & Quinn, 2014; WHO, 2016h).

Because the laboratory facilities and trained staff needed to identify a specific etiologic agent in a given patient with a suspected STI are often lacking in many LMICs, WHO (1991) has developed and promoted a

syndromic approach to the management of STIs. This approach relies on classifying patients with a suspected STI into various groups depending on their symptoms and findings on physical examination (e.g., women with a vaginal discharge, men with urethral discharge and dysuria, and patients with genital ulcer) and then treating them with a regimen designed to cover the treatable etiologic agents that are likely to be responsible (e.g., *N. gonorrhoeae* and *C. trachomatis* when a cervical or urethral discharge is present). **EXHIBIT 6-4** provides an example of this approach.

The finding that an intensive (and relatively expensive) regimen of AZT (zidovudine) given to HIV-infected pregnant women in the United States substantially reduced vertical transmission of HIV to their newborn babies led to trials of simpler and less expensive AZT regimens in Thailand and the Ivory Coast. These trials showed that even these simpler regimens, which are given entirely by mouth and for a shorter period of time, were partially effective in reducing vertical transmission of HIV. As a consequence, this regimen was made available in many LMICs (Shaffer et al., 1999; Wiktor et al., 1999). Since that time, recommendations for the prevention of mother-to-child transmission of HIV have rapidly evolved. In 2015, WHO released recommendations that all HIV-infected pregnant and breastfeeding women receive ART.

For many years, the prohibitive cost of antiretroviral drugs meant that the implementation of wide-scale AIDS treatment programs in LMICs such as those of sub-Saharan Africa was unthinkable. As a result, fewer than 10% of people in LMICs who needed ART had access to it in 2003. However, the introduction of generic competition in the global antiretroviral market precipitated spectacular decreases in the costs of these drugs, making it conceivable to include ART in the effort to reduce HIV-associated morbidity and avert early mortality for all affected groups. Efforts to expand access to ART, including the WHO's 3 by 5 initiative (a plan to treat 3 million people with ART by 2005); the establishment of the Global Fund to Fight HIV/AIDS, Tuberculosis, and Malaria; and the U.S. President's Emergency Plan for AIDS Relief (PEPFAR), which initially sought to treat 2 million people in selected LMICs, are now bearing fruit. In 2016, an estimated 20.9 million HIV-positive people were receiving ART (UNAIDS, 2017a). In hard-hit areas, such as eastern and southern Africa, the number of people on ART has nearly doubled since 2010 (UNAIDS, 2016b). Studies have demonstrated substantial declines in mortality among HIV-infected adults in LMICs resulting from such efforts (Jahn et al., 2008; Mermin et al., 2008), but substantial expansion of these programs is still needed to meet the needs of all HIV-infected persons who could benefit from such treatment.

Obstacles to Prevention and Control and Directions for Future Research

Enormous obstacles to changing sexual behaviors exist, although progress was clearly made in reducing the frequency of high-risk sexual behaviors and the incidence of HIV infection in Thailand (**EXHIBIT 6-5**) (Celentano et al., 1996; Nelson et al., 1996; Rojanapithayakorn & Hanenberg, 1996) and Uganda (Kilian et al., 1999) in response to the AIDS epidemic. In many societies, it remains socially

EXHIBIT 6-4 Syndromic Management of Genital Ulcer

A study in Lesotho illustrates the promise and drawbacks of the syndromic approach (Htun et al., 1998). In an attempt to validate STI flowcharts for the management of genital ulcer, researchers found that syndromic protocols would have provided adequate treatment for at least 90% of their patient population, while the traditional, clinically directed protocol would have provided adequate treatment for only 62% of those same patients. At the same time, syndromic protocols would have led to the overtreatment of primary syphilis in approximately 60% of patients, while the clinically directed protocols would not have resulted in any such overtreatment.

Several similar studies have shown that syndromic case management of STIs using flowcharts often leads to both improved treatment in many patients and overtreatment in some patients. In general, to determine the appropriateness of implementing the syndromic approach in a given region, the costs of overtreatment (including the cost of the drugs themselves), the risk of promoting drug resistance, and the stigma of an STI diagnosis (which can lead to domestic violence against women) must be weighed against the benefits of improved treatment, including reductions in STI and HIV transmission, decreases in sequelae from untreated infections, and increased patient satisfaction. In view of the high prevalence of STIs in many LMICs, this tradeoff is often considered acceptable.

EXHIBIT 6-5 A Successful Public Health Program: The Declining Spread of HIV Among Thai Military Conscripts

In the late 1980s, heterosexual commercial sex was found to contribute significantly to the rapid spread of HIV in Thailand. Thai authorities responded swiftly to this observation and implemented public health programs that substantially increased the number of commercial sex acts protected by condoms, which in turn led to significant reductions in the rate of HIV infection among young men in Thailand. This case study tells the story of this public health success.

The first national serosurvey for HIV conducted in Thailand found that HIV was exceedingly common among commercial sex workers. In June 1989, 44% of sex workers in the northern province of Chiang Mai were positive for HIV, a figure that would climb to 67% by June 1993 (Celentano et al., 1996). Commercial sex in Thailand was relatively common; for example, in the period from 1991 to 1993, more than 70% of Thai military conscripts were found to have engaged in at least one commercial sex act during their period of service. Because military conscripts are selected by a national lottery, these findings were applicable to the general population of young men in Thailand; thus, commercial sex was believed to be a common source of HIV infection.

To address this situation, Thai authorities implemented an HIV/AIDS prevention and control program that included the 100% Condom Campaign to promote condom use in commercial sex establishments. Under this campaign, free condoms were distributed to all sex establishments and the use of condoms was actively enforced by Thai authorities. The campaign was accompanied by mass advertising to promote condom use during commercial sex.

In the years following the initiation of the condom campaign, the use of condoms in sex establishments increased dramatically. National behavioral surveillance data revealed that the percentage of commercial sex acts in which condoms were used rose from approximately 14% in the years prior to 1989 to more than 90% by 1993 (Rojanapithayakorn & Hanenberg, 1996). Although the prevalence of HIV among sex workers remained high, the prevalence among newly inducted military conscripts declined from more than 10% in 1991 and 1993 to approximately 7% in 1995 (Nelson et al., 1996) and to 0.5% in 2009 (Rangsin et al., 2015). Further, this decline occurred in the absence of a visible AIDS epidemic (Rojanapithayakorn & Hanenberg, 1996). Experts attribute these heartening findings at least in part to the swift implementation of the Thai HIV/AIDS prevention and control program, and especially to the success of the 100% Condom Campaign.

acceptable—even desirable—for men to visit commercial sex workers and have multiple concurrent sex partners. Many men do not want to use condoms, and a woman may risk physical abuse, rejection, or loss of financial support if she tries to insist that a condom be used by her husband, boyfriend, or customer. Women who learn they are infected with HIV also risk abandonment or abuse if they share this information. Thus, changing sexual behaviors requires the education of men as well as women, and reducing the incidence of STIs, including HIV, requires raising the status and improving the power of women in society. Similarly, economic and other practices that contribute to or promote risky sexual behaviors (e.g., forcing men to live apart from their wives so that they can earn a living wage) need to be addressed.

Prevention of HIV infection in adolescent girls and women in sub-Saharan Africa is particularly challenging. Sub-Saharan African girls and young women (age 14–25) are being infected at two to four times the rate of boys and young men of the same age group (Shisana et al., 2014; UNAIDS, 2014), likely as the result of complex social, behavioral, and biological factors. Sexual relationships with older men have been shown to be a key risk factor for acquisition of HIV infection in women of this age group (Schaefer et al., 2017). Older men not only have a higher prevalence of HIV infection than younger men, but relationships between older men and young women also often cross socioeconomic strata. Such a power differential leaves these women in a poor position to negotiate condom use and may result in forced sex.

The global rise in antibiotic-resistant gonorrhea is complicating treatment of this once relatively easily treatable STI. Indeed, there is now widespread resistance to ciprofloxacin, once the first-line drug of choice, and increasing resistance to azithromycin and ceftriaxone (Wi et al., 2017). In 2016, WHO began recommending that azithromycin and ceftriaxone be used together as a first-line treatment of gonorrhea in an attempt to slow or prevent the emergence of strains resistant to all available antibiotics. More research is urgently needed to develop new antibiotics to treat gonorrhea (and other bacterial infections) and to make available rapid, point-of-care diagnostic tests, ideally ones that also detect antimicrobial resistance.

Vaccines against STIs, and particularly against HIV, are needed but have proved difficult to develop. Vaccines against HPV have been shown to be highly effective in preventing infection and cervical dysplasia caused by HPV types 16 and 18 and five other HPV genotypes that together cause approximately 90% of cervical cancer cases worldwide; it is expected that cervical cancer caused by these types will also be

prevented by use of the same vaccines (Castle & Maza, 2016; Chatterjee, 2014). However, it will be decades before declines in cervical cancer attributable to these vaccines will become apparent.

The development of vaccines against HIV has been beset by many problems, and there is continuing disagreement about which type of vaccine (i.e., inactivated whole virus, subunit, genetically engineered, live, attenuated, and so on) is likely to be safe and effective. Trials of a subunit HIV vaccine begun in the United States and Thailand in 1998–1999 showed that this vaccine was not effective in preventing acquisition of HIV infection (Cohen, 2003). To date, the only trial of an HIV vaccine to show any efficacy was a study in Thailand that showed at best a modest level of protection (Rerks-Ngarm et al., 2009). Although the vaccine had only modest efficacy, the trial did prove useful in increasing knowledge about the correlates of protection, and this knowledge has been used in designing subsequent candidate HIV vaccine evaluations. In 2016, a large efficacy trial of an HIV vaccine regimen began in South Africa, with results expected in 2020.

In the absence of vaccines against HIV and other STIs (other than HPV), other approaches to reducing morbidity and mortality beyond primary prevention of infection through behavior change are needed. For example, while women in high-income countries generally have access to regular screening for cervical dysplasia, which has been shown to be a highly effective tool for the secondary prevention of invasive cervical cancer, such programs are not available to most women in LMICs because of the cost and the need for moderately sophisticated laboratory support. In LMICs, the "see and treat" approach, which involves visual inspection of the cervix using acetic acid, is typically used, as it is less costly and invasive than a PAP smear and can be done in low-resource settings. However, scaling up this approach has proved difficult, leading to the exploration of other methods, such as primary screening for the presence of high-risk HPV types, as well as other methods, such as self-collection of samples, to make screening more culturally acceptable. Cervical cancer screening will continue to play an important role in cervical cancer prevention for years to come, despite the introduction of HPV vaccines.

As noted earlier, universal access to antiretroviral therapy to treat HIV infection in LMICs has become a top priority for international public health policy makers, not only to reduce morbidity and mortality among those already infected, but also because evidence shows that wide-scale treatment can reduce transmission of HIV at the population level (i.e., "treatment as prevention"). Although significant advances have been made in expanding access to ART, numerous challenges remain on this front. The current programs are not yet reaching many of those already infected with HIV, and large disparities in access to ART remain to be addressed. Thus, in the absence of an effective HIV vaccine, new approaches to identifying those at risk, reducing stigma, and assuring treatment are needed and will still play an important role once an HIV vaccine does become available.

Viral Hepatitis

Overview

Hepatitis, which entails inflammation of the liver, can be caused by many different viruses (as well as bacteria, protozoa, chemical agents, and some noninfectious diseases). However, at least five viruses specifically infect the liver—hepatitis viruses A, B, C, D, and E. Each belongs to a different family and has unique epidemiologic features, necessitating diverse approaches to control and prevention of hepatitis.

Viral hepatitis is a major global public health problem. All five primary hepatitis viruses can cause acute disease. HBV, HCV, and hepatitis D virus (HDV) can also produce chronic infection. In many LMICs, such persistent infections are the primary cause of serious liver disease, including chronic hepatitis, cirrhosis, and hepatocellular carcinoma; the last is a common cancer that is almost always fatal. Globally, an estimated 1.4 million deaths are attributable to acute and chronic infection with HBV and HCV (WHO, 2016i). In total, an estimated 240 million people are chronically infected with HBV and 130 to150 million people are chronically infected with HCV (WHO, 2016i).

Hepatitis A virus (HAV) and hepatitis E virus (HEV) cause acute self-limited disease; they do not cause chronic infection. Transmitted by the fecal–oral route, these viruses are endemic in many LMICs with suboptimal environmental sanitation. HAV infections are typically asymptomatic when they occur in infants. In older children and adults, they generally produce a self-limited illness that leads to few deaths. HEV infection results in more substantial morbidity and mortality, particularly when it occurs in pregnant women.

Etiologic Agents

The five primary hepatitis viruses (A, B, C, D, and E) account for almost all cases of viral hepatitis. Although other viruses (e.g., Epstein-Barr virus and cytomegalovirus) occasionally cause inflammation

of the liver, infections with these viruses do not principally involve the liver. GB virus C (formerly called hepatitis G virus) does not appear to cause liver disease (Bowden, 2001). The following discussion is limited to the five primary hepatitis viruses.

Descriptive Epidemiologic Features and Risk Factors

The prevalence of chronic HBV infection is moderate to high throughout LMICs. The highest prevalence of HBV infection is seen in East Asia and sub-Saharan Africa, with 6% of the adult population being chronically infected with this virus (WHO, 2017m). High prevalence and incidence of HBV infection are also seen in the Amazon Basin and the southern regions of central and eastern Europe. In these areas, most people who acquire HBV do so at the time of birth (through vertical transmission from an infected mother). Most such infections produce no acute clinical manifestations, but the likelihood of chronic infection and the risk of progression to chronic liver disease increase significantly as the age of acquisition of HBV decreases.

The most common routes of HBV transmission in LMICs are perinatally (from mother to child) and horizontally (from one child to another). Transmission of HBV through the use of nonsterile needles for medical injections is also thought to be significant. Less frequent modes of transmission of HBV in LMICs include sexual intercourse and needle sharing among injection drug users, the two principal modes of transmission in high-income and some middle-income countries. Additionally, practices such as tattooing, scarification, circumcision, body piercing, and acupuncture with nonsterile instruments can spread HBV.

HDV is a replication-defective virus, so that HDV infection can be acquired only in the presence of HBV infection. The distribution of HDV infection varies markedly by region, but it generally corresponds to the distribution of HBV infection, although some interesting exceptions to this pattern exist. For example, HDV infections are relatively rare in East and Southeast Asia, even though the prevalence of HBV infection in these regions is high (Margolis, Alter, & Hadler, 1997). The primary mode of transmission of HDV is through percutaneous exposure to blood, as may occur through the use of nonsterile needles and syringes for medical injections and the transfusion of unscreened blood. Perinatal and sexual transmission of HDV can occur, albeit infrequently.

HCV infection is widespread throughout the world. Although HCV is not as infectious as HBV, HCV infection is much more likely to become chronic; in fact, 55% to 85% of HCV-infected persons become chronically infected (WHO, 2017d). The prevalence of HCV infection varies by WHO region, with a majority of regions having a prevalence among adults ranging from 0.5% to 1.0%; a higher prevalence is seen in the WHO Eastern Mediterranean and European regions—2.3% and 1.5%, respectively. The principal modes of transmission of HCV in LMICs are believed to be the reuse of needles and syringes for medical injections, needle sharing among injection-drug users, and the transfusion of unscreened blood. HCV also can be spread perinatally and sexually, although transmission by these means is less common.

HAV infection is endemic in most LMICs. The primary mode of transmission of this virus is via the fecal–oral route, through either person-to-person contact or the ingestion of contaminated food or water. In areas with poor sanitation, HAV infection is essentially universal during the first few years of life, producing lifelong immunity. As sanitation improves, individuals are more likely to escape exposure to HAV in childhood. If they are infected as teenagers or adults, the clinical manifestations can to be more severe, albeit self-limited. Fatalities from HAV hepatitis are uncommon.

HEV is also endemic in many LMICs, where it causes substantial morbidity and mortality. HEV infection can be particularly deadly among pregnant women, with case fatality ratios approaching 30% among women infected during their last trimester (Jaiswal, Jain, Naik, Soni, & Chitnis, 2001; Khuroo, Teli, Skidmore, Sofi, & Khuroo, 1981). In endemic areas, HEV can produce cyclic, sometimes large outbreaks, as well as sporadic cases of hepatitis. HEV, like HAV, is acquired through the fecal–oral route, principally through ingestion of contaminated water. Person-to-person transmission of HEV can occur but is rare.

Current Approaches to Prevention and Control

A safe and effective vaccine for HBV has been available since 1982. In the early 1990s, WHO recommended integrating the HBV vaccine into the national immunization programs of all countries by 1997. In 2004, WHO further recommended that all infants receive a first dose of HBV vaccine as soon as possible following birth, with the goal being that all children would receive the vaccine within 24 hours of birth. This first dose is to be followed by two or three additional doses to complete the HBV vaccine series. As of 2015, an

estimated 84% of infants globally received three doses of HBV vaccine; however, coverage for receipt of the first dose of vaccine within 24 was lower, at only 39% (WHO, 2017m). Vaccination of older children and adults is given lower priority, with recommendations based on the epidemiologic features of HBV in the country, as well as vaccination levels in children.

Chronic HBV infection can be treated with interferon or antiviral drugs such as entecavir or tenofovir. With the advent of more effective drugs with limited side effects, as well as evidence that treatment can prevent liver failure and hepatocellular carcinoma, there have been calls for treatment to be provided to HBV-infected individuals in LMICs (Thursz, Cooke, & Hall, 2010).

Vaccination against HBV infection is also effective in preventing HDV coinfection. (As noted earlier, HDV infection requires the presence of HBV.) However, additional strategies are needed to prevent HDV superinfection among chronic carriers of HBV. Because there is no vaccine against HDV per se, the prevention of such superinfections depends on reducing percutaneous exposures to blood in medical and nonmedical settings among those persons who are chronically infected with HBV.

There is no vaccine against HCV at this time, and combating viral hepatitis has not historically been a high priority. However, with the adoption of the Sustainable Development Goals, viral hepatitis has been given high priority. WHO released its first global strategy for viral hepatitis in 2016 (WHO, 2016i). Treatment of chronic HCV infection has been rapidly evolving, with the current preferred treatment regimens including sofosbuvir, daclatasvir, and sofosbuvir/ledipasvir in combination. These regimens, which have remarkably few side effects and an excellent safety profile, can provide cure rates as high as 95%, a dramatic increase over previous regimens, which cleared the virus in only approximately 40% of patients. However, the prices of these drugs must be reduced substantially if large numbers of individuals in LMICs are to gain access to them.

Interrupting the transmission of HCV is also an important component of the strategy to decrease the prevalence of HCV infection globally. WHO (2017d) recommends the following measures to prevent HCV (and HBV) infection:

- Hand hygiene, including surgical hand preparation, hand washing, and use of gloves
- Safe and appropriate use of healthcare injections
- Safe handling and disposal of sharps and waste
- Provision of comprehensive harm-reduction services to people who inject drugs, including sterile injecting equipment
- Testing of donated blood for hepatitis B and C (as well as HIV and syphilis)
- Training of health personnel
- Promotion of correct and consistent use of condoms (WHO, 2017d)

The epidemiologic features of HAV infection vary by region and degree of sanitation, necessitating diverse approaches and control measures aimed at this virus. An effective vaccine against HAV is available and can be useful in controlling the periodic outbreaks among older children and adults that occur in areas of low endemicity. Although traditionally given as a two-dose series, the vaccine has been shown to be highly effective when only one dose is given to healthy individuals. However, this vaccine is licensed for use only in children age 1 year or older. In most LMICs, HAV infection is highly endemic and typically acquired during infancy and early childhood, when it produces little morbidity or mortality. Consequently, the prevention of HAV infection in LMICs through the use of HAV vaccine has not been a priority.

A vaccine for HEV has been developed and licensed in China, but is not currently available in other countries. Improved environmental sanitation—especially the provision of clean drinking water—remains the best strategy for preventing HEV infections.

Obstacles to Prevention and Control and Directions for Future Research

Preventing HBV and HCV infections remains a major global public health goal. Universal coverage with the HBV vaccine can dramatically reduce the incidence of HBV infection and serious liver disease and represents the best hope for decreasing the acute and delayed morbidity and mortality caused by this virus. Unfortunately, the cost of supplying and distributing this relatively inexpensive vaccine in settings that have many competing needs remains a significant barrier to achieving universal immunization. Depending on their cost, combined vaccines may help alleviate this difficulty in the future. However, giving the first dose of the vaccine within 24 hours of birth is difficult to achieve in many LMICs, especially where many births occur at home.

Many obstacles to preventing HBV and HCV infection through means other than vaccination exist. In spite of the existence of an efficacious HBV vaccine and an effective drug regimen for curing HCV infection, without a major expansion of current public health interventions, the number of people living with chronic HBV infection is expected to remain at high levels for the next four to five decades, while the

number of people living with chronic HCV infection is expected to increase. Nevertheless, given both tools that are currently available and those in the development pipeline, control of the viral hepatitis epidemic is possible. Control will require an investment in and expansion of interventions to provide vaccines for viral hepatitis, prevent mother-to-child transmission, increase the use of universal precautions, provide comprehensive care packages to injection drug users, and provide treatment to chronically infected individuals.

Malaria and Other Arthropod-Borne Diseases

Overview

Blood-sucking arthropods—including mosquitoes, flies, bugs, fleas, mites, and ticks—are efficient vectors for a host of pathogenic protozoa, bacteria, viruses, and worms that cause tremendous suffering and death around the world. Transmitted by mosquitoes, malaria is undoubtedly the most important parasitic disease in tropical regions of LMICs. Malaria caused an estimated 429,000 deaths in 2015, primarily among African children (WHO, 2016j). Mosquitoes also transmit dengue fever (an estimated 390 million infections per year; WHO, 2017l), Zika, yellow fever, filariasis, chikungunya, and Japanese encephalitis, while various other arthropods spread trypanosomiasis, leishmaniasis, onchocerciasis (river blindness), and plague, to name but a few arthropod-borne diseases. The public health importance of arthropod-borne diseases can scarcely be overstated: They contribute substantially to morbidity and mortality in affected countries and significantly retard social, economic, and developmental progress in those regions. The 2015–2016 Zika virus and 2013–2015 chikungunya pandemics demonstrate that emerging arthropod-borne diseases will continue to be a threat worldwide.

Etiologic Agents and Clinical Features

As noted earlier, arthropod-borne diseases are caused by a wide variety of pathogens. This section limits its discussion to malaria, dengue fever, yellow fever, Zika, American trypanosomiasis, and African trypanosomiasis. Note, however, that many of the prevention and control issues discussed in the context of these diseases are applicable to other arthropod-borne diseases. Onchocerciasis is discussed in the section on infectious causes of blindness.

Malaria is a febrile disease caused by five species of the parasitic *Plasmodium* protozoa: *P. malariae, P. falciparum, P. vivax, P. ovale,* and *P. knowlesi*. Dengue is caused by the four dengue viruses, which produce a spectrum of disease ranging from undifferentiated fever; to classic dengue fever, which is self-limiting and rarely fatal; to dengue hemorrhagic fever, which is characterized by plasma leakage and hemorrhage that can progress to shock and death. Yellow fever is caused by a virus of the same name; it is characterized by fever and, in severe cases, hemorrhage, jaundice, and liver and kidney involvement. Zika is caused by a cousin of the dengue and yellow fever viruses, Zika virus. It generally produces a mild, self-limited illness; however, in pregnant women, infection can lead to severe birth defects or loss of the pregnancy. In addition, Zika virus infection can lead to neurologic problems in adults, including Guillain-Barré syndrome and encephalitis.

American trypanosomiasis (Chagas disease) is caused by the protozoan parasite *Trypanosoma cruzi*. Acute infections are usually mild, but approximately one-third of infected individuals develop more severe chronic manifestations after several years of asymptomatic infection, including gastrointestinal tract damage, neurologic involvement, and cardiac damage leading to heart failure.

African trypanosomiasis (also known as sleeping sickness) is caused by two subspecies of the protozoan parasite *Trypanosoma brucei*—namely, *T.b.* rhodesiense and *T.b.* gambiense. Individuals infected with *T.b.* rhodesiense develop symptoms within weeks to months, whereas those infected with *T.b.* gambiense develop symptoms over a period of months to years. In both cases, the disease follows a course of central nervous system derangement, coma, and certain death if left untreated.

Descriptive Epidemiologic Features and Risk Factors

Although malaria is found in 91 countries around the world, more than 90% of cases occur in Africa (WHO, 2016j). The parasite is transmitted between humans (who serve as the reservoir) by the female *Anopheles* mosquito. In endemic areas, where transmission is constant, individuals who survive gradually develop immunity to severe disease. As a consequence, young children (who have not yet developed immunity) and pregnant women (who have depressed immune function) experience the highest rates of malaria morbidity and mortality. Susceptible individuals who enter endemic areas (e.g., migrant laborers, displaced persons, and travelers) are also at risk. Of the five species of malaria parasite, *P. falciparum* causes the most severe disease, with *P. vivax, P. malariae,* and *P. ovale*

causing milder forms of illness. Evidence suggests that infection with *P. knowlesi* may be more widely distributed than initially thought, as well as more severe, with approximately 1 out of 10 patients developing life-threatening disease (Cox-Singh et al., 2008; Daneshvar et al., 2009).

As many as 3.9 billion people live in areas where they are at risk for dengue virus infection, resulting in an estimated 390 million infections each year, of which 96 million present with clinically apparent dengue infection (Bhatt et al., 2013; Brady et al., 2012). Major epidemics of severe dengue have largely been limited to South and Southeast Asia and Latin America, whereas epidemics of dengue fever are more widespread.

Four dengue virus serotypes exist, all of which produce the same spectrum of clinical disease. Infection with one serotype produces lifelong immunity to that serotype, but confers only transient cross-protection against the others. Preexisting antibodies, resulting from either a prior infection with another dengue serotype or transfer to infants of maternal dengue antibodies, are a risk factor for developing severe disease. In addition, viral virulence factors are thought to influence the clinical outcome of infection. Dengue viruses have an urban or peri-urban transmission cycle (human–mosquito–human) in which the mosquitoes *Aedes aegypti* and *Aedes albopictus* are the principal vectors. Frequently, multiple dengue virus serotypes cocirculate in endemic urban cycles that periodically erupt in widespread epidemics (Gubler, 1998). A rural transmission cycle is also possible, but is less important from the public health perspective.

Yellow fever virus is endemic in 34 countries in sub-Saharan Africa and 13 countries in South America (WHO, 2018b). A recent modeling study estimated that the burden of yellow fever in 2013 in Africa, where a majority of the cases are seen, was 130,000 severe cases with 78,000 deaths (Garske et al., 2014). Yellow fever has an urban transmission cycle in which mosquitoes (typically *A. aegypti*) transmit the yellow fever virus between humans. Major epidemics of yellow fever occur when the virus is acquired by the urban vector and introduced to susceptible populations. Yellow fever also follows sylvatic and rural transmission cycles.

American trypanosomiasis (Chagas disease) occurs mainly in the Americas and is an important public health problem in Latin America, where it is estimated that 6 to 77 million persons are infected with *T. cruzi* (WHO, 2017e). Historically, *T. cruzi* has been (and continues to be) endemic in rural areas, where a variety of wild and domestic animals serve as reservoirs for the parasite. This parasite is transmitted to humans by the bite of triatomine bugs, which infest houses built with materials that shelter the bugs (e.g., thatched roofs or cracked mud walls). In recent decades, rural migration has introduced the disease to urban areas, where *T. cruzi* then spreads via transfusion of infected blood. Congenital infection also occurs. Despite these alternative transmission routes, the triatomine vector remains the most important source of infection.

African trypanosomiasis is found exclusively in sub-Saharan Africa, where the disease occurs in rural, endemic pockets in 36 countries. In recent years, control efforts have led to a massive reduction in the number of new cases, with an estimated 50,000 to 70,000 cases in 2006 declining to 2,804 reported cases in 2015 (WHO, 2017f). Although African trypanosomiasis was nearly eliminated by the early 1960s, its incidence soared in the following decades, and devastating epidemics occurred in the wake of civil disturbance, war, and population movements. Recently, a dramatic reduction in incidence and associated mortality has been achieved through improved case finding and treatment. The parasite that causes African trypanosomiasis, *T. brucei*, is transmitted to humans by the bite of the tsetse fly. Cattle and wild animals are the major reservoirs for *T.b.* rhodesiense, which is found in eastern and southern Africa, whereas humans are the major reservoir for *T.b.* gambiense, which is found in western and central Africa.

Current Approaches to Prevention and Control

The historical context of the current malaria control strategy is helpful in understanding how approaches to prevention and control evolve over time (Trigg & Kondrachine, 1998). In the 1950s, malaria prevention programs focused on parasite eradication through the use of antimalarial drugs (e.g., chloroquine) to eliminate human infection and reduce the parasite reservoir, and on residual insecticides (e.g., dichlorodiphenyltrichloroethane [DDT]) to interrupt malarial transmission via its mosquito vector. By the late 1960s, the incidence of malaria had been greatly reduced in areas that implemented these programs, including Latin America and tropical Asia. These improvements proved unsustainable, however—in part because of the rigidity of the eradication program, which was vertically structured and failed to consider regional differences in malaria epidemiology and public health infrastructure. Malaria resurged in the following decades (along with other mosquito-borne diseases, including dengue fever and yellow fever) as vector-control programs were not sustained.

Global malaria eradication was declared a failure and abandoned.

In the early 1990s, the international community identified malaria as a leading public health priority (again) and began to formulate new strategies for its control. In 1998, the Roll Back Malaria (RBM) Partnership was established to coordinate an international effort to reduce the global burden of malaria—the first such campaign in more than three decades. By 2009, thousands of organizations throughout the world were participating in RBM. The discovery of artemisinin and the development of artemisinin-based combination treatments (ACTs), which can be mass-produced relatively inexpensively and are highly effective against malaria parasites that have developed resistance to chloroquine and sulfadoxine–pyrimethamine, resulted in WHO recommending these agents as the first-line treatment for *P. falciparum* infection. A resultant increase in the deployment of ACTs has led to enhanced control of malaria.

In 2008, RBM released the Global Malaria Action Plan, which used a number of key tools, including long-lasting insecticide–treated bed nets, indoor residual spraying, intermittent preventive treatment during pregnancy, other vector-control strategies, improved diagnosis, and prompt provision of treatment. As a result, the incidence of malaria was reduced by 41% between 2000 and 2015. In 2015, the WHA adopted the Global Technical Strategy for Malaria 2016–2030 (GTS), with the goals of reducing the incidence of malaria by 90% by 2030 compared to the 2015 incidence, eliminating malaria from 35 countries where it was endemic in 2015, and preventing the reestablishment of malaria in malaria-free countries.

Currently, there is no licensed malaria vaccine, although several are currently under development. The most promising malaria vaccine to date is RTS,S, an adjuvanted recombinant vaccine that has been shown to reduce clinical malaria cases by 26% following a three-dose primary series and 39% when a fourth dose is administered 18 months after the primary series (RTS,S Clinical Trials Partnership, 2015). Vaccine efficacy waned following the fourth dose, however, so additional booster doses may be necessary to provide long-term protection. Pilot studies of such regimens are set to begin in a number of countries in Africa in 2018.

There is no specific treatment for dengue fever or dengue hemorrhagic fever beyond attentive clinical management. In 2015, Dengvaxia, a tetravalent vaccine for dengue virus, was licensed. In 2016, WHO recommended that countries highly endemic for dengue consider introducing this vaccine. By the end of that year, seven countries had approved use of Dengvaxia. Unfortunately, this vaccine is of limited use in the control of dengue, as it is currently licensed only for children aged 9 years and older and adults, due to its low efficacy in sero-naïve individuals. In addition, the efficacy of the vaccine is lower against dengue virus serotype 2. Five additional dengue vaccine candidates were under development in 2017, including two that are in Phase III clinical trials. Until an efficacious vaccine that protects against all four serotypes is widely used, the control and prevention of dengue fever will largely depend on mosquito control, disease surveillance, and epidemic preparedness. A key component of the control of *Aedes* mosquitos is source reduction through elimination of mosquito breeding sites.

Immunization is the most important prevention measure against yellow fever. The yellow fever vaccine, designated as 17D, is safe, effective, and affordable; it requires only a single dose to provide multiple-year, likely lifelong protection. To prevent and control yellow fever in endemic areas, WHO (1998c) recommends routine infant immunization against yellow fever (within the EPI), mass immunization campaigns (e.g., preventive catch-up campaigns), vigilant case and vector surveillance, reactive immunization to contain outbreaks, and careful management of the vaccine supply. Since 2004, WHO has maintained a stockpile of approximately 6 million doses of yellow fever vaccine for emergency response, though this supply may not be enough if several concomitant large outbreaks occur. In response to the large-scale outbreaks in Angola and Democratic Republic of the Congo in 2016, the global stockpile was increased; however, the global supply has been stretched to keep up with increased demand.

For American trypanosomiasis (Chagas disease), it is currently recommended that all acute infections be treated, as well as chronic infections in children. A mounting body of evidence suggests that treatment of all chronic infections may reduce the severity of illness; however, the drugs used in the treatment of acute manifestations (benznidazole and nifurtimox) are toxic, so treatment of chronic infections in adults depends on each individual's health status. Prevention and control of American trypanosomiasis currently focus on interrupting vector-related transmission and preventing transmission via the transfusion of infected blood. Vector-control efforts center on eradicating strictly domestic triatomine bugs and controlling domestic infestations of other (sylvatic) triatomine bugs, predominantly through indoor residual insecticide spraying in areas at high risk for American trypanosomiasis. The prevention

of transfusion-related transmission is accomplished by screening all blood donors for *T. cruzi* antibodies and rejecting blood donations that test positive. The results of these efforts are encouraging, with the global burden of American trypanosomiasis having dropped substantially since the 1990s.

In 2007, WHO launched the WHO Global Network for Chagas Elimination, with the goal of eliminating American trypanosomiasis by 2010. That goal was not met, and, the past decade has seen spread of Chagas disease into areas previously considered non-endemic. In addition to the previously mentioned control measures, the WHO network is working to increase surveillance and access to diagnosis and treatment, as well as to develop new treatments and improved diagnostic tests.

The current strategy for the control and prevention of African trypanosomiasis caused by *T.b.* gambiense involves active case finding and early treatment, which reduces the human reservoir by curing the infection. Passive surveillance for *T.b.* gambiense infection is problematic because infected individuals are often infectious for months to years before they develop symptoms that lead to (passive) detection. In contrast, passive surveillance and epidemic preparedness are often adequate to control the transmission of *T.b.* rhodesiense, which quickly leads to symptoms of infection. The drugs used to treat African trypanosomiasis have significant drawbacks, including high levels of toxicity. The first-stage drugs, pentamidine and suramin, and second-stage drugs, melarsoprol and eflornithine, are provided by WHO to countries in which this disease is endemic.

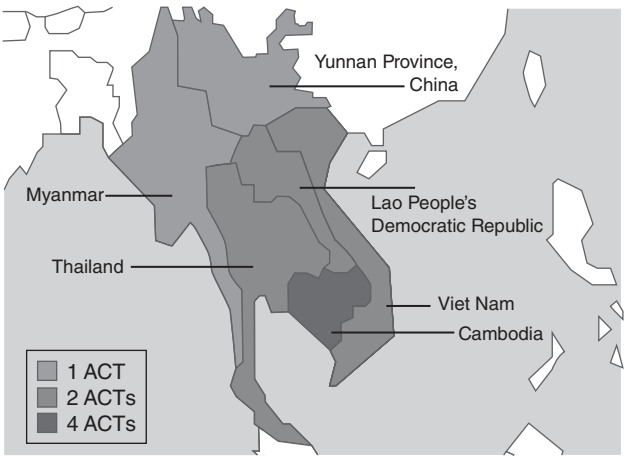
Obstacles to Prevention and Control and Directions for Future Research

Insecticide resistance presents a major challenge to prevention strategies based on vector control. Vectors for malaria, dengue, and yellow fever have all developed resistance to multiple insecticides, including both the classic insecticides (e.g., DDT) and their replacements (e.g., organophosphates). Alternative insecticides have been developed but are often more costly and toxic than the first-line agents. The dynamic ecology of vector-borne pathogens is an additional barrier to effective vector control. For example, changing land-use patterns, shifting weather patterns, urbanization, and population movements may result in new foci for disease transmission. Moreover, national and international travel and commerce present continuing opportunities for reinfestation, necessitating continuous surveillance, even in areas that have achieved good vector control.

Drug resistance poses a serious problem for control strategies that rely on chemotherapy to treat disease and subsequently decrease transmission through the reduction of parasite load in humans. Nowhere is this problem more evident than in the history of the management of malaria. Drugs that once provided inexpensive and effective means of antimalarial prophylaxis and therapy became severely compromised by resistance in many parts of the world, resulting in substantially increased morbidity and mortality. The discovery and use of ACTs, in combination with other interventions, has resulted in a renewed ability to control malaria—but resistance to ACTs was first detected in Cambodia in 2008 and has spread since that time (**FIGURE 6-8**). In addition, sporadic reports of ACT-resistant malaria continue to crop up, making careful monitoring to detect drug resistance essential for continued success.

Other barriers to the prevention and control of arthropod-borne diseases include the intermittent disruption of public health services as a consequence of civil disturbance or war and the ever-present problem of limited resources. Problems include the difficulty of sustaining support for successful control measures (e.g., indoor insecticide spraying to control the vector-borne spread of American trypanosomiasis) once the disease has been brought under control and, therefore, is no longer a public health priority.

Future prospects for the improved prevention of arthropod-borne disease rest primarily on vaccine development. Current research is focused on developing additional tetravalent dengue fever vaccines and vaccines against malaria and Zika virus.



ACT, artemisinin - based combination therapy

FIGURE 6-8 Distribution of malarial multidrug resistance, 2016.

Reprinted from World Health Organization. (2016). *World malaria report, 2016*. http://apps.who.int/iris/bitstream/10665/252038/1/9789241511711-eng.pdf?ua=1

Helminthiasis

Overview

Approximately 1.5 billion people are believed to be chronically infected with parasitic helminths (WHO, 2017g). The vast majority of these individuals live in LMICs, where environmental sanitation is poor. Each parasite produces different manifestations, depending on the site, intensity, and length of infection. The host response also influences the clinical course of infection. In general, children experience the heaviest worm burden, and persistent infection throughout childhood is common in LMICs. Heavy, prolonged infection adversely affects growth, development, and educational achievement, and significantly increases childhood morbidity. In adults, helminthiases can produce acute and chronic morbidity, leading to impaired productivity, chronic disability, and reduced quality of life.

In terms of global prevalence, morbidity, and mortality, the most important helminthic infections are schistosomiasis (207 million infections), ascariasis (807 million to 1.2 billion infections), trichuriasis (800 million infections), hookworm (576 to 740 million infections), onchocerciasis (26 million infections), and lymphatic filariasis (120 million infections) (CDC, 2013, 2017, 2018; WHO, 2017h). Although many of these infections are minimally symptomatic, the number of clinically significant cases is substantial. This section focuses on schistosomiasis and the most common intestinal helminthiases—that is, ascariasis, trichuriasis, and hookworm. Dracunculiasis, a disease of considerable historical importance, is also considered.

Etiologic Agents and Clinical Features

Human schistosomiasis (also known as bilharziasis) is caused by a group of blood trematodes (flukes) known as schistosomes. The three main species that infect humans are *Schistosoma mansoni, Schistosoma japonicum,* and *Schistosoma haematobium*. Two other species, *Schistosoma mekongi* and *Schistosoma intercalatum*, also parasitize humans, but such infections are uncommon and will not be discussed further. Acute schistosomiasis (Katayama fever) is characterized by fever and chills, abdominal pain, diarrhea, and enlargement of the spleen or other organs. Infection with *S. japonicum* is most commonly associated with acute disease, although any type of schistosomiasis can cause these findings. Chronic disease is initiated by the deposition of schistosome eggs in the body, which induces inflammation and scarring. Chronic infection with *S. mansoni* or *S. japonicum* can lead to hepatosplenic and intestinal involvement, although most infected individuals remain asymptomatic. Chronic infection with *S. haematobium* is associated with urinary tract disease, and a much higher proportion (50% to 70%) of infected individuals are symptomatic (WHO, 1993). *S. haematobium* infection is also associated with an increased risk of bladder cancer.

The most common intestinal helminthiases of humans are caused by *Ascaris lumbricoides*, *Trichuris trichiura* (whipworm), and two species of hookworm, *Ancylostoma duodenale* and *Necator americanus*. These parasites produce a diverse range of clinical manifestations, although persons with light infections generally remain asymptomatic. When worm burdens are high, *A. lumbricoides* infection is associated with malnutrition and stunted growth. *A. lumbricoides* is also the largest of the intestinal helminths, measuring 15 to 35 centimeters in length; thus, single worms can cause obstruction and inflammation of the appendix, bile duct, or pancreatic duct, and a bolus of worms can cause intestinal obstruction leading to death. Trichuriasis is associated with abdominal pain, diarrhea, and general malaise when moderate worm loads are present, whereas nutritional deficiencies, anemia, and stunted growth can result from heavier worm loads. Hookworms attach to the intestinal mucosa leading to chronic blood loss (Hotez et al., 2006). The major clinical manifestations of hookworm infection are iron-deficiency anemia and hypoalbuminemia.

Dracunculiasis is caused by *Dracunculus medinensis* (Guinea worm), a long (60 to 90 centimeters), thin nematode. Symptoms of the disease are caused by the migration of the worm to the subcutaneous tissues and its eruption through the skin. Clinical manifestations can include the formation of a blister or bleb at the site where the worm will emerge (usually the ankle or foot), hives, nausea, vomiting, diarrhea, and asthma. The blister ulcerates as the worm emerges, and local pain typically persists until the worm is expelled or extracted, which can take several weeks. Secondary infection of the worm track is common. Severe disability generally lasting one to three months occurs in most persons who become infected, with 0.5% of infected individuals experiencing permanent disability (Hunter, 1996). Even short-term disability can have a substantial economic impact if the worm emerges during the months when important agricultural work is performed (Hopkins, 1984).

Descriptive Epidemiologic Features and Risk Factors

Schistosomiasis is the single most important helminthic disease and the second most important parasitic disease (after malaria) worldwide. An estimated 218 million persons in 78 countries and territories harbor this infection, resulting in an estimated 3.3 million disability-adjusted life-years (DALYs) lost each year (Hotez et al., 2014; WHO, 2017h). *S. mansoni*, the most widespread of the schistosomes, is found in parts of Africa, the Middle East, the Caribbean, and South America; *S. haematobium* is restricted to Africa and the Middle East; and *S. japonicum* is currently found only in the Western Pacific countries. Overall, the vast majority of people requiring treatment for schistosome infections (90%) reside in sub-Saharan Africa (WHO, 2017h).

The life cycle of the schistosome is quite complex, and a complete description is beyond the scope of this chapter. Nevertheless, a brief discussion is necessary to understand how infection plays out in human hosts. Adult worms live, mate, and deposit their eggs in the blood vessels lining the human intestines (*S. mansoni* and *S. japonicum*) or bladder (*S. haematobium*). These eggs migrate into the intestine or bladder and are passed out in the excreta. When deposited in fresh water, the eggs develop into immature parasites (miracidia) that infect fresh-water snails. The immature parasites then develop into infective larvae (cercariae). When released into the surrounding water, the larvae swim about. Those coming in contact with humans during this time penetrate through the skin and eventually reach the vascular system, completing the cycle. Thus, the maintenance of schistosomes in a given area requires the presence of the intermediate host, the contamination of water with egg-laden excreta, and the exposure of humans to contaminated water.

Because schistosomes do not multiply in the human host, the intensity of infection is largely determined by the rate at which new worms are acquired. Studies have shown that the infection rate for schistosomes typically peaks during childhood and then declines with advancing age (Butterworth, Dunne, Fulford, Ouma, & Sturrock, 1996; Hagan, 1996). This finding is thought to reflect age-dependent changes in both water exposure and immunity to infection.

The public health impact of intestinal helminthiases on LMICs is substantial. As noted earlier, *A. lumbricoides*, *T. trichiura*, and hookworm each infect 740 million to 1.2 billion individuals globally (CDC, 2013, 2017, 2018). Although the morbidity and mortality associated with intestinal helminthiases are relatively low, the burden of illness is considerable due to the extremely high prevalence of these infections (WHO, 2017g).

A. lumbricoides, T. trichiura, and hookworm share a relatively simple life cycle. The parasites mature and mate in the human intestine and produce eggs that are passed out in the feces. When deposited on moist soil, these eggs develop into the infective form of the parasite. (Because this part of the life cycle is spent in soil, these worms are collectively referred to as geohelminths.) Humans acquire the parasite when contaminated soil is ingested or, in the case of hookworm, when the infective larvae in the soil penetrate exposed skin. Thus, the maintenance of these helminths requires the deposition of feces in soil and the exposure of humans to that soil. For *A. lumbricoides* and *T. trichiura*, the peak intensity of infection typically occurs among young children; in contrast, for hookworm, the peak occurs among adults (Bundy, Hall, Medley, & Savioli, 1992).

Dracunculiasis is contracted by drinking contaminated water. Parasitic larvae then migrate to the abdominal cavity, where they mature into adult worms over the course of a year. After copulation, the gravid female worm migrates through the body to the subcutaneous tissue and then secretes irritants that ulcerate the skin and expose the worm. When the affected body site is immersed in water, the worm expels clouds of larvae. The larvae are consumed by copepods (minute fresh-water crustaceans, also known as water fleas), which are in turn ingested by humans who drink contaminated water. Dracunculiasis occurs in individuals of all ages, but is less common in very young children, possibly because breastfeeding reduces their level of water consumption (Hunter, 1996).

A global campaign to eradicate dracunculiasis was launched in the 1980s as a part of the United Nations International Drinking Water Supply and Sanitation Decade (1981–1990). Since then, the incidence of this disease has declined from 4 million reported cases in 1981 (Hopkins, Ruiz-Tiben, & Ruebush, 1997) to 25 confirmed cases in 2016 (WHO, 2017i). Endemic disease is now confined to just three African countries, and dracunculiasis is likely to be the next disease that is eradicated.

Current Approaches to Prevention and Control

The major strategies to control helminthic infections focus on preventive chemotherapy, environmental sanitation and health education, and control or eradication of the intermediate host. No antihelminthic vaccines are currently available.

For the control of schistosomiasis and infection with geohelminths, WHO emphasizes controlling morbidity by reducing the intensity of infection (the number of worms per person) rather than reducing the prevalence of infection (the number of persons with worms). The former is achieved in the short term through regular deworming, despite reinfection, whereas the latter requires improved environmental sanitation, intermediate-host control (in the case of schistosomiasis, for example), or both, and is viewed as a long-term goal. WHO recommends preventive treatment for schistosomiasis, soil-transmitted helminthiases, lymphatic filariasis, and onchocerciasis. This approach is possible because safe, effective, and inexpensive oral treatments are available for these common helminthiases. Five neglected tropical diseases—schistosomiasis, lymphatic filariasis, onchocerciasis, soil-transmitted helminthiases, and trachoma—are currently controlled through an integrated approach to tailored preventive chemotherapy (**FIGURE 6-9**) (WHO, 2017j). In this approach, six drugs are used in differing combinations, depending on which diseases are present, to both treat infected individuals and reduce transmission in the communities. This approach has not only reduced morbidity, but also led to increased control and is contributing to elimination of these diseases. More than 1.5 billion people were in need of preventive chemotherapy for these diseases in 2015.

Ultimately, most helminthiases are diseases of poor sanitation. Providing for the sanitary disposal of human excreta prevents the contamination of soil and water, thereby breaking the chain of transmission. However, environmental improvements must be made at the community level to be effective. Improvements in environmental sanitation are complemented by health education to promote healthy behaviors. For example, WHO has determined that schistosomiasis could be largely prevented by eliminating indiscriminate urination and defecation, and increasing compliance with medical interventions (WHO, 1993).

Control of the intermediate host species for certain helminths, such as snail control for schistosomiasis, can also play an important role in preventing transmission. The primary strategies for snail control include application of molluscicidal agents and environmental management. Recent reviews have indicated that regular applications of molluscicides can significantly contribute to elimination efforts (King, Sutherland, & Bertsch, 2015; Sokolow et al., 2016). Control of the intermediate host plays a crucial role in preventing the transmission of dracunculiasis. Much of the success of the dracunculiasis eradication campaign has been achieved by teaching people to pour drinking water through a finely woven cloth or to use a small straw-like device to filter out the tiny fleas that carry the *Dracunculus* parasite. Chemical eradicants have also played a role.

Obstacles to Prevention and Control and Directions for Future Research

Major obstacles to the effective implementation of chemotherapy-based interventions are related to the logistical and financial difficulties of treating the large number of individuals in need. In addition, chemotherapy does not address the conditions that led to the

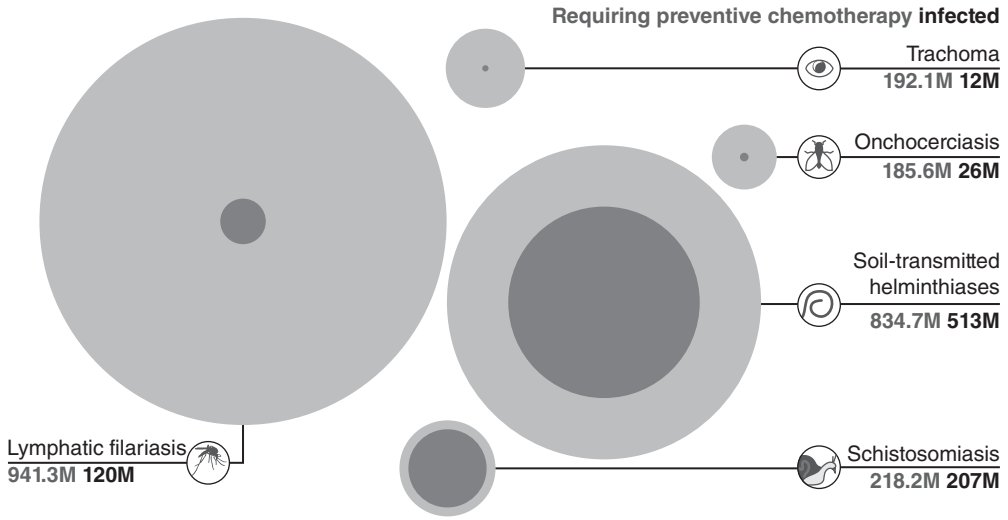


FIGURE 6-9 Five neglected tropical diseases amenable to preventive chemotherapy.

Reprinted from World Health Organization. (2017). *Crossing the Billion.* Geneva, Switzerland: Author.

primary infection; thus, reinfection is likely to occur, necessitating repeated treatments and further costs.

Adequate environmental sanitation is generally viewed as the only long-term solution to the control of intestinal helminthiases and schistosomiasis. However, providing safe drinking water and adequate disposal facilities in LMICs has proved to be an extremely difficult, costly, and lengthy process. Current control efforts emphasize targeted chemotherapeutic interventions to relieve the immediate burden of disease while maintaining universal sanitation as a long-term goal.

Although a massive reduction in the burden of illness caused by the *Dracunculus* parasite has been achieved, as with any disease, the last cases are the most difficult to find, particularly in areas with political instability or conflict. Furthermore, the recent finding of *D. medinensis* infection in dogs in endemic areas poses a major obstacle to achieving eradication. Further research is needed to understand risk factors for infection in dogs and the role of dogs in the transmission cycle.

Zoonoses

Overview

Zoonoses are defined as diseases and infections that are naturally transmitted between vertebrate animals and humans. Zoonotic agents include a wide variety of bacteria, viruses, protozoa, and helminths, while their nonhuman hosts include wild animals and birds, food and draft animals, and household pets.

More than 200 different zoonoses are currently recognized (Hart, Trees, & Duerden, 1997). Zoonotic infections have a worldwide distribution and collectively produce significant morbidity and mortality. For example, indigenous rabies, which has a global distribution, causes more than 59,000 deaths annually (WHO, 2017k). *Yersinia pestis*, the bacterium that causes plague, remains enzoonotic in many parts of Africa, Asia, and South America, and in the southwestern United States. Although only 3,248 cases of human plague were reported to WHO between 2010 and 2015, the control of human plague remains a public health priority because of its epidemic potential and relatively high case fatality ratio. In LMICs, important zoonoses include rabies, plague, anthrax, leptospirosis, leishmaniasis, African trypanosomiasis, and a number of hemorrhagic fever viruses. Foodborne zoonoses, such as salmonellosis, campylobacteriosis, and certain *E. coli* infections, also contribute to the global burden of disease, while occupational zoonoses such as brucellosis, echinococcosis, and Q fever have regional significance. Finally, zoonotic infections are a rich source of emerging diseases, such as severe acute respiratory syndrome (SARS), avian influenza in humans, Nipah virus encephalitis, variant Creutzfeldt-Jakob disease, and Ebola hemorrhagic fever.

Although zoonoses have a worldwide distribution, LMICs bear a greater burden of these diseases than high-income countries. This divergence reflects, in part, differences in opportunities for exposure to zoonotic pathogens. Residents of LMICs typically experience more frequent and intimate contact with animals, and they often live in situations of suboptimal environmental sanitation (which promotes exposure to infective material, such as contaminated animal excreta). In addition, many control measures that are readily available in high-income countries (e.g., mass veterinary vaccination of domesticated animals) may not be available or affordable in LMICs.

Zoonotic pathogens are transmitted to humans by five major routes:

1. Inhalation: Transmission occurs when infective materials are aerosolized and inhaled.
2. Ingestion: Transmission occurs when humans consume contaminated meat, milk, or blood from infected animals or when food-stuffs (e.g., fruits and vegetables), drinking water, or hands are contaminated with infective materials, which are then ingested.
3. Nontraumatic contact: Transmission typically occurs when pathogens enter through the skin (or mucosal surfaces or conjunctivae) as a result of direct or indirect contact with animal hides, hair, excreta, blood, or carcasses.
4. Traumatic contact: Transmission occurs via animal bites or scratches.
5. Arthropod: Transmission occurs by biting arthropods that feed on animals and humans.

For some zoonoses, humans are incidental, dead-end hosts. Human-to-human transmission of rabies or anthrax, for example, is extremely rare. For other zoonoses, humans may serve as a reservoir for infection, transmitting pathogens to other humans or even back to animals.

Zoonoses produce a wide variety of diseases with distinct clinical and epidemiologic characteristics. In addition, they are associated with many different animals occupying a number of ecologic niches, both urban and rural. Approaches to the control and prevention of zoonoses necessarily reflect these differences. The general principles underlying these strategies

often include good animal husbandry (including vaccination when appropriate), environmental sanitation, vector control, and the control or elimination of animal reservoirs (e.g., rats or other wild animals). The One Health approach recognizes the interplay of human health, animal health, and the environment, and encourages collaborative efforts across multiple disciplines to improve the health of all (see the *Public Health Infrastructure* chapter for more information on this approach).

Rather than attempting to survey the entire range of zoonoses, this section is limited to a detailed discussion of rabies and leptospirosis, which are important zoonoses in many LMICs.

Etiologic Agents and Clinical Features

Rabies is caused by the rabies virus, which is a Lyssa virus belonging to the Rhabdoviridae family. Typically transmitted by the bite of a rabid animal, human rabies is an acute, encephalitic disease that is almost always fatal. Following inoculation, there is an asymptomatic incubation period typically lasting one to three months, although periods as short as a few days and as long as several years have been reported (Fishbein, 1991). Post-exposure prophylaxis must be initiated during this period to be effective.

The prodromal period for rabies begins with the early signs of disease, which include nonspecific symptoms (e.g., fever and malaise) and abnormal sensations near the site of inoculation. This stage is followed by a 2- to 10-day period of acute neurologic dysfunction that manifests as furious rabies in approximately 80% of cases and paralytic (dumb) rabies in the remainder (Fishbein, 1991). Furious rabies is characterized by periods of extreme agitation and hyperactivity interspersed with periods of normalcy. Hydrophobia, aerophobia, combativeness, and hallucination may also occur. Features of paralytic rabies include paresthesias, weakness, and paralysis. The almost inevitable final stage of rabies (both furious and paralytic) is coma followed by death.

Leptospirosis is an acute febrile disease caused by pathogenic bacteria of the genus *Leptospira*. The clinical course of the disease is variable: The majority of infections are asymptomatic or result in mild disease, but 5% to 10% of persons who become infected will develop more severe manifestations, including kidney failure and pulmonary hemorrhage (Levett, 2001). Ocular involvement may occur as well.

Leptospira bacteria are classified serologically into 200 serovars arranged in 24 serogroups; alternatively, the organism may be classified genotypically into genomospecies (Vinetz, 2001). The serologic classification scheme is more frequently used in epidemiologic studies. Certain serovars were once thought to be associated with more severe disease, but evidence supporting this hypothesis is lacking (Levett, 2001; Vinetz, 2001).

Descriptive Epidemiologic Features and Risk Factors

Rabies is a disease of animals; humans are incidentally infected and only rarely transmit the virus. Animal reservoirs include dogs, cats, and wild animals (notably foxes, skunks, wolves, coyotes, raccoons, mongooses, and bats). Most human rabies cases are acquired from dogs, which are the source of more than 95% of fatal rabies infections in humans in LMICs. In countries where immunization of domestic animals is common, wild animals constitute the principal reservoirs of infection and human rabies is extremely rare.

Rabies virus is present in the saliva of infected animals and is typically transmitted to humans by the bite of a rabid animal. It may also be transmitted when intact mucous membranes are exposed to infective saliva. Not every exposure results in infection. Nevertheless, if infection occurs, it is almost always fatal in the absence of post-exposure prophylaxis.

Leptospirosis has a worldwide distribution, but its incidence is difficult to assess. In general, the incidence and prevalence of infection are low in high-income countries and higher in LMICs. For example, in some LMICs, prevalence of antibodies to leptospirosis in the range of 18% to 48% has been reported, suggesting that exposure to the causative organism is common in such settings (Ellis, 1998). As with rabies, human infection with *Leptospira* is incidental, and humans do not contribute to the transmission of the bacteria. The main animal reservoirs for *Leptospira* serovars that cause human disease are rats, dogs, pigs, and cattle.

Transmission of leptospires to humans typically occurs in one of two ways: (1) through contact with water that has been contaminated with the urine of infected animals or (2) through direct contact with infective animal urine. The bacteria infect humans by entering through broken skin, water-softened intact skin, mucosal surfaces, or conjunctivae. Human leptospirosis is an occupational disease among those whose professions involve contact with host animals (e.g., dairy farmers) or water (e.g., rice farmers). Home and recreational exposures are also becoming increasingly

important, particularly in LMICs. In addition, periodic flooding due to heavy rains can produce large epidemics of leptospirosis.

Current Approaches to Prevention and Control

The control of human rabies is achieved by the prevention of human exposure to the virus and the prevention of disease through post-exposure prophylaxis when exposure does occur. Historically, rabies vaccines were not deemed suitable (i.e., sufficiently inexpensive, safe, and effective) for mass pre-exposure immunization of humans. However, research is currently under way to gather evidence to support investment in rabies vaccination programs in Africa and Asia (WHO, 2017k).

The prevention of human exposure to rabies in LMICs primarily depends on the control of dog rabies, which can be achieved through the widespread use of the canine rabies vaccine. The mass vaccination of pet dogs and the elimination of stray or feral dogs has been successful in reducing human and dog rabies in many countries. For example, in the Americas, the incidence of human rabies was reduced by more than 95% and dog rabies was reduced by more than 98% between 1983 and 2016 (WHO, 2017k).

The prevention of human disease once exposure to rabies virus has occurred depends on good local wound care (e.g., flushing with soap and water) and post-exposure prophylaxis, which entails passive immunization with immunoglobulin and active immunization with the rabies vaccine. Although the complete post-exposure regimen almost always prevents disease, this regimen must be delivered during the incubation period: Neither immunization nor other treatments alter the course of disease once symptoms develop. There are no diagnostic tests that can detect rabies infection in humans prior to the onset of symptoms; thus, all individuals with suspected or possible infections must receive prophylactic treatment.

The prevention of human leptospirosis focuses on interrupting the transmission of leptospires to humans. Preventive measures include the vaccination of certain host animals (e.g., cattle) or the elimination of others (e.g., rats). Human vaccination has met with limited success and is not widely applied. Environmental control strategies aimed at reducing hazards such as stagnant bodies of water or the periodic flooding of residential areas, as well as educational campaigns aimed at decreasing unnecessary water exposures, may enhance prevention efforts. Occupational improvements that curtail contact with host animals or contaminated water are also desirable. Antibiotic therapy and supportive care remain the treatment of choice, although evidence supporting the effectiveness of antibiotic therapy is limited.

Obstacles to Prevention and Control and Directions for Future Research

The major obstacle to the improved control of rabies is the difficulty of achieving adequate vaccine coverage among dog populations in LMICs. At least 70% of the canine population must be vaccinated in endemic areas to control rabies. Ensuring the delivery of vaccine to stray and feral dogs (or eliminating these dogs altogether) is especially problematic. With sustained effort, WHO considers the global elimination of urban (canine) rabies to be an attainable goal.

The future control of leptospirosis is less promising. Animal vaccination may not be economically feasible for many farmers, and vaccines are not available or appropriate for every host species. The elimination of wildlife hosts (e.g., rats) is usually not feasible. Improvements in working and living conditions to reduce contact with animals and their excreta are desirable (for many reasons), yet difficult to achieve in LMICs. Until these difficulties are overcome, the control of leptospirosis must rely on health education to reduce risky behavior, veterinary education to promote good animal husbandry, and medical education to ensure the prompt diagnosis and treatment of leptospirosis.

Viral Hemorrhagic Fevers

Overview

The viruses that cause hemorrhagic fever (HF) belong to five different families (**TABLE 6-6**), and the illnesses they produce have distinct epidemiologic features. Despite their differences, these viruses produce a common clinical picture that is characterized by fever and hemorrhage, as their name suggests. Hemorrhagic manifestations can include petechiae, ecchymoses, bleeding gums, nosebleeds, vaginal bleeding, and bleeding from other mucosal surfaces, producing bloody urine, stool, and vomit. Complications can include cardiovascular and neurologic disturbances, shock, and death.

The spectrum of disease typically associated with each type of virus varies substantially. For example, Lassa virus infections are inapparent or result

TABLE 6-6 Distribution and Modes of Transmission of Viral Hemorrhagic Fevers

Family, Genus, and Virus	Disease	Principal Means of Transmission	Principal Locations
Flaviviridae			
Flavivirus			
Dengue	Dengue HF	Mosquito	Asia, Latin America
Yellow fever	Yellow fever	Mosquito	sub-Saharan Africa, South America
Kyasanur Forest disease	Kyasanur Forest	Tick	India
Omsk HF	Omsk HF	Tick, muskrat	Russia
Alkhurma HF	Alkhurma HF	Tick	Egypt, Saudi Arabia
Arenaviridae			
Arenavirus			
Lassa	Lassa fever	Rodent, person-to-person	West Africa
Junin	Argentine HF	Rodent	Argentina
Machupo	Bolivian HF	Rodent	Bolivia
Guanarito	Venezuelan HF	Rodent	Venezuela
Sabio	Brazilian HF	Unknown	Brazil
Lujo	Lujo HF	Rodent, person-to-person	South Africa
Chapare	Chapare HF	Rodent	Bolivia
Bunyaviridae			
Phlebovirus			
Rift Valley fever	Rift Valley fever	Mosquito	Africa
Nairovirus			
Crimean–Congo HF	Crimean–Congo HF	Tick, person-to-person	Africa, Asia, Eastern Europe, Middle East
Hantavirus			
Hantaan, Seoul, and others	HF with renal syndrome	Rodent	Asia, Europe

Filoviridae			
Filovirus			
Marburg	Marburg HF	Person-to-person	Africa
Ebola	Ebola HF	Person-to-person	Africa
Paramyxoviridae			
Henipavirus			
Nipah		Pig, bat, person-to-person	Malaysia, Singapore
Hendra		Horse	Australia

Abbreviation: HF, hemorrhagic fever.

in mild symptoms in many people, yet cause severe life-threatening disease in others. Ebola virus and Marburg virus, by comparison, appear to cause severe disease in virtually all those infected. The severity of the clinical illness that results from infection also varies according to differences in host response, viral virulence factors, and dose.

Infection with the HF viruses is, in general, relatively rare (Lassa virus is a notable exception). Although outbreaks are dramatic and lead to major responses on the part of public health and other officials, such infections do not typically have the same public health impact in terms of morbidity and mortality as the infectious diseases that more commonly afflict inhabitants of most LMICs (WHO, 1985). However, because most HF viruses are extremely virulent and capable of causing epidemics, developing strategies to control these viruses is a public health priority. In addition, some of the viruses, such as the South American HF viruses, pose regional public health hazards within their areas of endemicity.

All HF viruses are thought to be zoonotic. These viruses have been proven or are suspected to be transmitted from animals to humans by an arthropod vector (e.g., ticks or mosquitoes) or by direct or indirect contact with the animal reservoir (e.g., rodents or bats) of the virus (see Table 6-7). In endemic areas, the temporal distribution of many viral HFs follows seasonal changes in the activity and density of the vector or animal reservoir or seasonal changes in human activity. The age and sex distributions of infection and disease caused by some HF viruses reflect differences in exposure to the vector or reservoir, whereas other viruses affect persons of all ages and both sexes.

Several of the HFs have only recently been recognized and had their agents characterized. For example, the South American HFs, Ebola HF, and Marburg HF appear to have emerged in just the past 50 years, presumably as a result of increased human activity or settlement in areas where these viruses were already circulating among their zoonotic hosts. In some instances, such as with Argentine HF, the geographic range of the virus has expanded beyond its initial focus (Vainrub & Salas, 1994). The development of control and prevention strategies for newly emerging infections can pose special difficulties because investigators have not yet conducted the research needed to formulate treatment modalities, vaccines, vector-control strategies, and other preventive measures.

Etiologic Agents and Clinical Features

Although at least 19 different HF viruses exist (see Table 6-6), the public health significance of each virus differs substantially. The most prominent (in terms of incidence of disease) are the Lassa, yellow fever, and dengue viruses. The latter two are discussed in the section on arthropod-borne diseases. Ebola virus and Marburg virus are notable for their epidemic potential and extraordinarily high case fatality ratios, whereas the South American HFs caused by the Junin, Machupo, and Guanarito viruses have regional public health importance. Detailed discussion of the remaining HF viruses is beyond the scope of this chapter.

Descriptive Epidemiologic Features and Risk Factors

Ebola HF emerged in 1976 when concurrent epidemics occurred in Democratic Republic of the Congo (formerly Zaire) and Sudan. These outbreaks were caused by distinct subtypes of the virus and were characterized by high case fatality ratios (88% in Democratic Republic of the Congo and 53% in Sudan). In total, 602 persons were infected and 431 died in the two outbreaks. A number of subsequent outbreaks occurred in Africa between 1979 and 2012. These outbreaks were limited in size and scope, resulting in fewer than 500 cases, and many of them were much smaller, with variable case fatality ratios (25% to 82%). Then, in December 2013, an unprecedented Ebola outbreak began in West Africa, resulting in a total of 28,616 suspected, probable, or confirmed cases and leading to 11,310 deaths (**EXHIBIT 6-6**).

The natural reservoir for Ebola has yet to be identified, but it is generally believed that the first person in an outbreak becomes infected through a spillover event—for example, due to contact with an infected primate or fruit bat. Epidemics of Ebola HF have been sustained by person-to-person transmission through direct physical contact with infected patients or corpses (or with their bodily fluids or tissues) and the use of unsterile needles for medical injections. Sexual transmission of Ebola virus through contact with semen from a man who had recovered from Ebola was suspected in the recent West African epidemic.

The first recognized outbreak of Marburg HF occurred in Marburg, Germany, in 1967 among laboratory workers (and their contacts), who acquired the infection from monkeys imported from Uganda. Since then, small numbers of cases have been reported in South Africa (1975), Kenya (1980 and 1987), and Uganda (2007 and 2008); in addition, more than 200 cases were identified in Democratic Republic of the Congo from 1998 to 2000 (Bausch et al., 2003). In 2004–2005, the largest outbreak to date occurred in Angola, involving 252 cases and 227 deaths (case fatality ratio of 90%) (Towner et al., 2006). The case fatality ratio of Marburg HF ranges from 21% to 90%.

Like Ebola, Marburg virus is also transmitted by person-to-person contact. Fortunately, airborne transmission does not appear to be common for either virus. Bats and nonhuman primates also play a role in the transmission of Marburg virus (Towner et al., 2007).

Lassa virus is widely distributed across West Africa, where it causes substantial morbidity and mortality. Although its incidence is poorly defined, it has been estimated that 100,000 to 300,000 new infections occur each year, with 5,000 deaths (Ogbu, Ajuluchukwu, & Uneke, 2007). Maternal death, fetal death, and permanent deafness are common complications of this disease (Cummins et al., 1990; Monson et al., 1987; Richmond & Baglole, 2003).

Lassa virus is maintained in a rodent reservoir that is commonly found in the home. Virus is shed in rodent urine and droppings, and rodent-to-human transmission is believed to occur by aerosolization or ingestion of rodent excreta or by inoculation through broken skin. Rodent-to-human transmission may also occur when infected rodents are consumed as food (Ter Meulen et al., 1996). Person-to-person transmission occurs in community and hospital settings, and makes a substantial contribution to epidemics of Lassa fever. This mode of transmission requires direct contact with infected persons. Person-to-person airborne

EXHIBIT 6-6 Ebola Outbreak, 2013–2016

The Ebola outbreak that occurred between December 2013 and April 2016 is by far the largest Ebola outbreak to date. The outbreak consisted of more than 28,000 reported confirmed, suspected, or probable cases, including 11,310 deaths. However, given the difficulty in collecting data on this outbreak, these numbers likely significantly underestimate the toll.

The outbreak has been traced back to an initial case that occurred in late December 2013 in a 2-year-old boy in Guinea, who likely acquired the virus through zoonotic transmission. All subsequent transmissions most likely occurred through the person-to-person route. Unlike previous outbreaks, which had occurred in rural and isolated areas, this outbreak spread into a large city, Conakry, and from there onward. Cases appeared in Liberia in late March 2014 and in Sierra Leone in May 2014. In August 2014, WHO declared the Ebola outbreak to be an international public health emergency. By that point, due to the spread of the virus into urban areas with mobile populations, outbreaks in all three affected countries were undergoing exponential growth. A massive international response, which was criticized by many as being too little, too late, was needed to end the outbreak.

Ultimately, the outbreak was controlled using traditional methods: finding symptomatic cases; isolating them; when possible, admitting patients to Ebola treatment centers and providing supportive care; tracing contacts of cases; and providing safe burials for the dead. Additionally, vaccination with a vaccine developed and deployed on an emergency basis, which began in Guinea in April 2015, helped to reduce transmission and prevent flare-ups.

transmission occurs rarely, if ever. Sexual transmission can occur, but the importance of this mode of transmission is unknown.

Argentine HF, which is caused by Junin virus, was first described in 1955 in agricultural workers in the Argentine pampas. Several hundred cases of Argentine HF occur each year in large, primarily agricultural regions of the pampas. The region of endemicity is expanding, however, and is now nearly 10 times larger than the initial compass (Vainrub & Salas, 1994).

Bolivian HF, which is caused by Machupo virus, was described in northeastern Bolivia, which is the only known endemic area. Outbreaks of Bolivian HF occurred in the 1960s and early 1970s, including large epidemics that affected hundreds of individuals. Although no cases were reported from 1976 to 1992, possibly due to effective host control, a small outbreak in 1994 and subsequent outbreaks in 2007, 2008, and 2011 have marked its reemergence.

Venezuelan HF, which is caused by Guanarito virus, was first recognized in 1989. During the period from September 1989 to January 1997, 165 cases were reported within the small region of central Venezuela where Guanarito virus is endemic (De Manzione et al., 1998).

For infections caused by the Argentine, Bolivian, and Venezuelan HF viruses, the case fatality ratio is in the range of 15% to 33% (De Manzione et al., 1998; Doyle, Bryan, & Peters, 1998). Each of these viruses is associated with a rodent reservoir that maintains the virus in the wild. As with Lassa fever, rodent-to-human transmission occurs by the aerosolization or ingestion of virus-laden rodent excreta or by inoculation through broken skin. The rodent that carries Junin virus typically dwells in agricultural fields, whereas the rodent that carries Machupo virus readily enters the home. Rodent-control strategies must take such differences into account. Person-to-person transmission of Junin, Machupo, and Guanarito viruses is considered rare, and nosocomial outbreaks are uncommon.

Current Approaches to Prevention and Control

Field trials have demonstrated that a live, attenuated Junin virus vaccine is safe and provides effective protection against Argentine HF (Maiztegui et al., 1998); it may also provide cross-protection against Bolivian HF. Vaccines are not available for Lassa fever or Marburg HF, although several candidate Lassa fever vaccines are under development. A vaccine for Ebola, called VSV-EBOV, has been shown to be very efficacious and is currently available for use in flare-ups in Africa (Henao-Restrepo et al., 2017). Several other candidate Ebola vaccines are currently in development or undergoing clinical trials.

In the absence of vaccines, reducing the morbidity and mortality caused by HF viruses depends on preventing primary transmission by limiting exposures to virus reservoirs and vectors and controlling secondary transmission (e.g., person-to-person transmission in the hospital, household, or community setting) through patient isolation and barrier nursing. In addition, the use of antiviral drugs or convalescent serum is effective in some instances.

Strategies to limit exposure to the virus are determined by the unique characteristics of the associated animal reservoir or vector and the distinct ways in which the various viruses are transmitted. For example, the rodent that carries Machupo virus is frequently found in and around the home, and aggressive rodent eradication measures through trapping and poisoning appear to have been quite successful in controlling Bolivian HF (Kilgore et al., 1995). Conversely, the rodent reservoir of Junin virus lives in crop fields, where trapping and poisoning are difficult, necessitating the development of alternative rodent abatement strategies. Eradication (or even control) of the rodent reservoir of Lassa virus in West Africa is not considered feasible due to the density and wide distribution of the rodent that carries the virus. Thus, preventing the primary transmission of Lassa virus has relied on educating at-risk communities about ways to reduce opportunities for exposure, such as never leaving food items uncovered and never consuming rodents as food.

Historically, nosocomial and person-to-person transmission of Lassa fever, Ebola HF, and Marburg HF have contributed significantly to devastating outbreaks of these diseases. (The South American HF viruses are rarely transmitted by these routes.) Field experience indicates that epidemic control of limited outbreaks can be achieved through simple barrier nursing techniques (e.g., wearing gloves, gowns, and masks; sterilizing equipment; and isolating patients), and epidemiologic studies support this conclusion. For example, serologic studies in Sierra Leone found that hospital personnel who used barrier techniques when caring for patients with Lassa patients had no greater risk of infection than the local population (Helmick, Webb, Scribner, Krebs, & McCormick, 1986). Traditional Ebola control measures include rapid diagnosis, isolation of cases, contact tracing, provision of supportive treatment, and safe burial practices.

At present, few specific treatments are available for the viral HFs. During the 2013–2016 Ebola epidemic, two experimental drugs, ZMapp and TKM-Ebola,

were used on a limited emergency basis. A subsequent small trial of ZMapp failed to meet the preset superiority cut-point, but the therapy did appear to be beneficial (PREVAIL II Writing Group, 2016). Ribavirin (an antiviral drug) is effective in the treatment of Lassa fever (McCormick et al., 1986). Laboratory data suggest that ribavirin may also be effective in treating South American HFs, although clinical data supporting this usage are incomplete (Doyle et al., 1998). Convalescent serum is useful in the treatment of Argentine HF (WHO, 1985), but donors are not plentiful. Most people in LMICs are not able to afford these therapies.

Obstacles to Prevention and Control and Directions for Future Research

The major obstacle to containing outbreaks of viral HFs (especially Ebola HF, Marburg HF, and Lassa fever) is inadequate disease surveillance, which results in delayed response and increased opportunities for epidemic spread. In many LMICs, disease surveillance is impeded by the difficulty of making an early diagnosis in areas where illnesses with similar initial manifestations (e.g., malaria, influenza, typhoid fever, leptospirosis, meningococcemia, and hepatitis) are prevalent. The lack of ready access to diagnostic laboratories exacerbates this difficulty. In addition, because epidemics are unpredictable in time and place, surveillance efforts are difficult to maintain. Other obstacles to the control and prevention of viral HFs include the costliness of sustaining readiness for infection control measures (e.g., maintaining supplies for barrier nursing), lack of information about the vectors and reservoirs of the Ebola and Marburg viruses, the difficulty of developing and maintaining vector- and rodent-control programs, and the limited availability of ribavirin (especially for the treatment of Lassa fever).

Infectious Causes of Blindness

Overview

Severely decreased visual acuity or complete blindness is profoundly disabling in any setting, but perhaps even more so in LMICs. Among the known causes of blindness, two infectious agents play important etiologic roles in selected regions of the world: *C. trachomatis*, the cause of trachoma, and *Onchocerca volvulus*, the cause of onchocerciasis, also known as river blindness.

Trachoma

Etiologic Agents

C. trachomatis is a small bacterium that lives within selected types of human cells and is difficult to grow in the laboratory. Although *C. trachomatis* is also the cause of STIs, as discussed previously, different immunotypes of the bacterium cause trachoma and genital tract infections. Those that cause trachoma are spread via a person-to-person route, most probably through eye and possibly nasal secretions on the hands. *C. trachomatis* is also spread mechanically by flies and probably by fomites such as washrags and handkerchiefs. Repeated episodes of infection in young (preschool) children lead to scarring of the eyelids, which then causes in-turned eyelashes that abrade the corneal surface, leading to subsequent corneal opacification and reduced visual acuity or blindness in adults.

Descriptive Epidemiologic Features and Risk Factors

Trachoma is an important cause of preventable blindness across the world. It is estimated that 229 million people live in areas endemic for trachoma, spread across 53 countries (Taylor, Burton, Haddad, West, & Wright, 2014). Approximately 21 million persons have active trachoma, of whom 2.2 million are blind or severely visually impaired. Trachoma is a disease of poverty that was described by the ancient Egyptians and previously was found throughout the world. It disappeared from Europe and almost all of the United States long before antimicrobial agents became available in the 1930s and 1940s; improved standards of living and personal hygiene are credited with its disappearance.

Trachoma is not a reportable condition, and what is known of its descriptive epidemiologic features comes from numerous surveys. This disease persists in hot LMICs, particularly in North Africa, the Middle East, sub-Saharan Africa, and drier regions of India and Southeast Asia. In hyperendemic areas, infection of the eye is generally universal in children by their fifth birthday, but active disease is seen largely in older children. Repeated reinfection in children leads to the permanent damage to the eyes that results in subsequent blindness or visual impairment in adulthood. Although infection in childhood appears to be equally common in boys and girls, the blinding complications appear to be more common in women, perhaps because of repeated exposure to infected children.

Risk factors for trachoma in children largely relate to facial cleanliness, the presence of flies, and cultural practices that lead to an increased likelihood of

person-to-person transmission of the etiologic agent, such as sharing washcloths and ways in which eye makeup is applied.

Current Approaches to Prevention and Control

Intervention studies have demonstrated that mass treatment with a variety of topical or oral antimicrobial agents and health educational programs that lead to improved facial cleanliness can substantially reduce the prevalence of trachoma in a community, as can fly control (Bhosai, Bailey, Gaynor, & Lietman, 2012; Emerson et al., 2004; House et al., 2009). Reductions in disease incidence in LMICs in the absence of a specific control program have also been documented as access to water, access to health care, and hygiene have improved.

The current approach to reducing trachoma-associated blindness in endemic areas is summarized by the acronym SAFE:

- **S**urgery to correct eyelid deformity
- **A**ntibiotics to treat acute eye infection and reduce sources of infection in the community
- **F**acial cleanliness
- **E**nvironmental change that enhances availability of water and reduces the prevalence of flies

In 1997, WHO launched a new trachoma control program—called Global Elimination of Trachoma by 2020 (GET 2020)—based on this approach. In 2016, more than 260,000 people with trachomatous trichiasis were provided with corrective surgery and more than 86 million people received antibiotics to eliminate trachoma through the integrated approach to preventive chemotherapy.

Obstacles to Prevention and Directions for Future Research

Trachoma is likely to remain a persistent problem in endemic areas until rising socioeconomic conditions result in better access to water, improved personal hygiene and sanitation, reductions in the numbers of flies, and improved access to healthcare services. Community-wide treatment with antimicrobial agents can lead to marked reductions in trachoma in such areas, but these reductions have proved difficult to sustain unless such treatment is made a routine part of regularly available health services and is accompanied by improvements in hygiene. Although a vaccine against trachoma has been discussed for many years, it remains unclear whether an effective vaccine can or will ever be developed.

Onchocerciasis

Etiologic Agents

O. volvulus is a filarial parasite that is spread through the bite of one of several species of *Simulium* black flies. During the bite of an infected female fly, larvae enter the body. There, they ultimately develop into adult worms that form nodules, usually over bony prominences. Adult worms can survive inside these nodules for as long as 15 years. The female adult worm produces microfilariae that migrate to the skin and the eye and are ingested by female flies when they bite an infected person, thereby completing the cycle.

In the skin, an inflammatory response to dead and dying microfilariae can lead to incapacitating itching and various types of degenerative, often unsightly, skin changes. In the eye, heavy and prolonged infection of the cornea with the microfilariae leads to opacification and reduced visual acuity or total blindness. The microfilariae can be detected by taking small snips of skin, immersing them in saline, and examining the saline microscopically.

Descriptive Epidemiologic Features and Risk Factors

Onchocerciasis is found only in a band of sub-Saharan African countries, parts of South America, and the Eastern Mediterranean. Approximately 26 million people worldwide are believed to be currently infected with the parasite, and 185.6 million people require preventive chemotherapy (WHO, 2017j). Onchocerciasis is the world's second leading cause of preventable blindness due to infection. Within affected regions, onchocerciasis occurs in clusters, largely determined by distance from the black fly breeding sites. The intensity of infection (and hence the risk of visual impairment) increases with age, as the burden of adult female worms producing microfilariae increases. This burden tends to be greater in men than in women, perhaps reflecting work-related exposures to the flies.

Current Approaches to Prevention and Control

Approaches to the control of onchocerciasis and prevention of the blindness it causes have included vector control, mass treatment of infected individuals, and nodulectomy (removal of nodules to reduce the source and number of microfilariae that migrate to the eyes). Early attempts to control onchocerciasis targeted the *Simulium* flies that serve as vectors, the immature stages of which require running water (e.g., rivers and streams) for their development. Initially, DDT was the pesticide added to rivers that served as

the breeding grounds for the vector. Beginning in the 1970s, other agents that target the larval stages of the fly (e.g., temefos) were used with great success, particularly in West Africa. These programs permitted resettlement of fertile areas that had been abandoned because of onchocerciasis, but the flies' development of resistance to temefos required switching to other larvicidal agents in some areas.

The control of onchocerciasis was revolutionized in the late 1980s with the introduction of ivermectin, a single dose of which eliminates microfilariae for a number of months. Because ivermectin does not kill the adult worms, repeated treatment (e.g., every 6 to 12 months) of infected individuals over many years is needed to provide continued suppression of the number of microfilariae and to prevent visual damage. Treatment every 3 months may be even more effective at reducing both the number of female worms in nodules and the severity of symptoms (Gardon et al., 2002). In a noteworthy humanitarian gesture, the manufacturer of ivermectin has made a commitment to provide the drug for free "for as long as necessary to as many as necessary." Globally, control is managed through an integrated approach of preventive chemotherapy. Currently, nearly 400 million doses of ivermectin are distributed each year for preventive treatment of onchocerciasis or the combination of onchocerciasis and lymphatic filariasis (WHO, 2017j).

Obstacles to Prevention and Directions for Future Research

Currently, the goal is to eliminate onchocerciasis worldwide. Since the implementation of an integrated approach in 2008, major gains have been made toward achieving this goal, including the elimination of the disease in four countries in the Americas.

Antimicrobial Resistance

Antimicrobial resistance (AMR) is one of the greatest public health challenges of the twenty-first century. Antimicrobial agents that kill or inhibit the growth of microorganisms consist of three main classes: disinfectants, antiseptics, and antibiotics. The term "antibiotics," while initially used to refer only to antibacterial agents, now also includes antiviral, antiparasitic, and antifungal agents. Resistance to antimicrobial agents used to treat specific infections, such as gonorrhea, malaria, and tuberculosis, has been addressed in the relevant chapter sections and thus will not be addressed here.

Antibiotics are used to treat a wide variety of infections. Global consumption of antibiotics has been estimated at more than 70 billion doses per year (Van Boeckel et al., 2014), with use increasing due to economic growth and the resultant increased access to and use of antibiotics in LMICs.

In spite of the massive increase in antibiotic consumption globally, insufficient access to effective and affordable antibiotics remains a major public health problem in low-resource settings, where lack of or delayed access to antibiotics is responsible for many more deaths than antibiotic resistance (Laxminarayan et al., 2016). Nevertheless, AMR is clearly a growing problem that contributes to the burden of infectious disease morbidity and mortality, although the magnitude of the problem is difficult to quantify.

Public health measures—particularly vaccination and sanitation—play an important role in helping reduce antibiotic use. For instance, introduction of pneumococcal conjugate vaccine in the United States resulted in reduced antibiotic consumption and fewer infections caused by antibiotic-resistant strains of *S. pneumoniae* (Kyaw et al., 2006). In countries with a high burden of diarrheal illness, access to clean water and adequate sanitation could dramatically reduce the use of antibiotics for diarrheal illnesses. Improving laboratory capacity in low-resource settings, as well as providing access to point-of-care diagnostic tests, would both reduce antibiotic consumption and improve access to appropriate treatment (Mendelson et al., 2016).

Nowhere is a One Health approach more needed than in combating AMR. As is true in humans, antibiotics are a necessary part of maintaining animal health and need to be available for appropriate veterinary use. However, the use of low-dose antibiotics for growth promotion in food animals and as prophylaxis in fish farming are important contributors to the rise of AMR. Indeed, the use of antimicrobial agents in animals to increase food production greatly exceeds the use of these drugs in humans (WHO, 2012b). In 2006, the European Union banned the use of antimicrobial agents for growth promotion, although they are still widely used for that purpose globally.

Addressing the rise of AMR and ensuring continued access to treatment with effective antimicrobial agents will require global action. In 2015, WHO endorsed an action plan to combat AMR that includes five strategic objectives:

- Improve awareness and understanding of AMR
- Strengthen knowledge through surveillance and research
- Reduce the incidence of infections
- Optimize the use of antimicrobial agents

- Develop the case for sustainable investment and increase investment in the development of new antimicrobial agents (WHO, 2015b)

The way forward will require that we improve surveillance for AMR, decrease human consumption of antibiotics while improving access to appropriate treatment, adopt responsible animal husbandry practices, expand measures to prevent and control infections in humans, and foster innovations such as new medicines, vaccines, and diagnostic tests, as well as sustained political commitment on a global scale. The chapters on *Public Health Infrastructure* and *Pharmaceuticals* provide additional discussion of AMR.

Emergence of New Infectious Disease Threats

The availability of a growing number of antimicrobial agents and vaccines beginning in the 1950s and continuing through the 1970s simultaneously led to the burgeoning infectious disease prevention and control activities described earlier in this chapter and to a widely shared assumption that infectious disease threats to human (and animal) health would diminish over time. A sobering note was injected into this otherwise optimistic conversation with the unexpected emergence of new infectious diseases and the reemergence of previously controlled diseases that became apparent in the 1980s and 1990s. Their appearance made it clear that efforts to control existing infectious diseases needed to be accompanied by an improved global capacity for early detection and rapid response to newly emergent or reemergent infectious diseases. Concern about the possible intentional release of infectious agents in a deliberate effort to frighten, harm, or kill gave added impetus to efforts in this area.

The concept of emerging and reemerging infections, which began receiving prominent attention after the U.S. Institute of Medicine released a report entitled *Emerging Infections: Microbial Threats to Health in the United States* in 1992, encompasses several distinct phenomena that can produce unexpected and sometimes urgent infectious disease threats. Included under the rubric of emerging and reemerging infectious disease threats are diseases caused by microbial agents not previously known to cause illness in humans (e.g., SARS, Nipah virus encephalitis, and H5N1 [avian] influenza in humans), the appearance in a new location of an infectious agent (e.g., the spread of West Nile virus to the United States or Zika virus to South America), the appearance of a new epidemiologic pattern of disease caused by an infectious agent (e.g., epidemic meningococcal meningitis caused by serogroup W-135 *N. meningitidis*), the appearance or spread of new variants of an infectious agent (e.g., multidrug- and extensively drug-resistant *M. tuberculosis*), and the resurgence of an infectious disease previously under good control (e.g., diphtheria in parts of the former Soviet Union).

The 1992 report by the Institute of Medicine, together with a follow-up report in 2003 (Smolinski, Hamburg, & Lederberg, 2003), outlined a number of factors that could promote the emergence or reemergence of infectious disease threats. In addition to microbial adaptation and change in response to selection pressures, numerous human and environmental factors may contribute to this trend—for example, increased human susceptibility to infection; human demographic conditions and behaviors; international travel and commerce; economic development and changes in land use; technological change; climate and ecologic factors; breakdown of public health measures; war and famine; poverty and social inequality; lack of political will; and intent to harm. Individually and collectively, these factors may increase the likelihood that humans will come in contact with various infectious agents, produce a greater susceptibility to infection, enhance and accelerate the spread of infectious agents, and reduce the capacity of communities to detect and respond effectively to infectious disease threats. Although the result can be dramatic outbreaks that command widespread public attention and political response (e.g., SARS, Ebola, and Zika), other equally serious infectious disease threats (e.g., multidrug- and extensively drug-resistant tuberculosis) often go largely unnoticed by the general public and decision makers because they do not produce explosive epidemics.

A dramatic example of an emerging infectious disease threat was the appearance of SARS. This life-threatening form of pneumonia appears to have originated in southern China in late 2002 (CDC, 2003). Before the significance of these cases of pneumonia became apparent, the illness had spread to multiple other parts of Asia and eventually to Europe, North America, and elsewhere. By the middle of 2003, more than 8,400 probable cases of SARS and more than 800 SARS-related deaths had been reported from 30 countries, with China, Hong Kong, and Taiwan bearing the brunt of the epidemic. SARS is caused by a previously unknown member of the Coronaviridiae family that almost certainly originated in one or more animal

reservoirs and initially spread to humans through direct exposure to infected animals in southern China. The virus then proved to be transmissible from person to person, particularly in healthcare settings. Spread of the SARS-associated coronavirus was also documented on commercial airplanes, and the ease of modern travel from one part of the world to another clearly facilitated the rapid dissemination of the virus across the globe.

In addition to the morbidity and mortality it caused, the SARS epidemic had an enormous economic impact. The total cost of this epidemic to Asian economies alone was estimated at approximately $60 billion. In the absence of a vaccine, preventing further spread of the SARS epidemic required stringent isolation of patients, quarantine of exposed individuals, restrictions on travel, and various other measures that were highly disruptive to commerce, travel, and other aspects of life.

Experience with influenza demonstrates many of the challenges related to detecting and responding rapidly to an emergent infectious disease threat. The appearance in Hong Kong in 1997 of human cases of influenza caused by the highly pathogenic avian influenza virus A (H5N1) heralded widespread epizootics (i.e., epidemics in animals) in domestic poultry and widespread infection (often asymptomatic) in many species of wild birds that ultimately reached much of Asia and parts of Africa, Europe, and the Middle East. By 2008, several hundred human infections with H5N1 had been documented in Asia (Vietnam, Thailand, Cambodia, Indonesia, and China), as had a smaller number of infections in Egypt and Turkey (Writing Committee of the Second World Health Organization Consultation on Clinical Aspects of Human Infection with Avian Influenza A [H5N1] Virus, 2008).

The H5N1 influenza virus is highly pathogenic in people. More than 50% of the people with documented illnesses died from their disease, prompting well-justified concerns about a human epidemic or pandemic caused by this virus and resulting in efforts to make and stockpile a safe and effective H5N1 vaccine. To date, however, there have been very few instances of person-to-person transmission of H5N1 virus (almost all human cases are the result of animal-to-human transmission), and it remains uncertain whether this virus is capable of both retaining its high pathogenicity and becoming readily transmissible from person to person.

While public health officials and vaccine manufacturers were preparing for a possible H5N1 influenza pandemic, a novel H1N1 influenza A virus of swine origin arose. Although this H1N1 virus was first detected in ill individuals in southern California and in Mexico in early 2009 (Novel Swine-Origin Influenza A [H1N1] Virus Investigation Team, 2009), its origins remain unknown. By the time the virus was first detected, it undoubtedly had undergone many generations of transmission among humans and had already become widely disseminated, making attempts to contain it futile. Within a short period of time after its discovery, the novel H1N1 influenza A virus was circulating and causing disease throughout North America, Europe, and Asia in the northern hemisphere and causing large epidemics in numerous countries in South America, southern Africa, and Australia and New Zealand in the southern hemisphere. Although a vaccine containing this virus was developed and produced with

EXHIBIT 6-7 Global Outbreak Alert and Response Network

Throughout the 1970s, 1980s, and 1990s, outbreaks of novel or high-impact infectious diseases (e.g., Ebola, cholera, and meningococcal meningitis) in LMICs often led to ad hoc responses by multiple national, international, academic, research, and private-sector organizations. Unfortunately, many of these efforts were uncoordinated and sometimes duplicative or even competing in nature. Although countries often looked to WHO for leadership and technical assistance in such instances, historically this organization has had very limited trained personnel and resources to devote to such efforts, as well as political and legal constraints that could limit its access.

In recognition of these problems, in 2000 WHO created the Global Outbreak Alert and Response Network (GOARN), which has its headquarters in Geneva, Switzerland. GOARN is a network of more than 200 institutions and networks that provides rapid assistance to identify and control infectious disease outbreaks and public health emergencies worldwide. This network has proved useful as SARS, influenza, yellow fever, and other outbreaks have appeared since 2000. Revisions to the International Health Regulations that give WHO improved access when important outbreaks occur should improve the functioning of this system.

exceptional speed, neither the vaccine nor other approaches (e.g., antiviral drugs, isolation and quarantine, and social distancing measures) prevented a pandemic from unfolding.

Fortunately, this influenza A virus proved to be no more virulent than the influenza A and B viruses that cause annual epidemics, although it still produced widespread morbidity and mortality. A more virulent influenza A virus would have exacted a substantially higher human toll, and would have been equally difficult to contain.

As these examples demonstrate, there is an ongoing threat of new infectious agents entering and causing disease in human populations, particularly from animal sources. Even as progress is made in eradicating or controlling historically important infectious diseases, the need persists for vigilance, preparedness, and a high level of international cooperation to detect and respond to new infectious disease threats that might emerge. WHO and others are working to improve global capacity in this area (**EXHIBIT 6-7**).

Global Health Security Agenda

In 2014, the Global Health Security Agenda (GHSA) was launched by multiple partners, including almost 50 countries, diverse international organizations, and various nongovernmental stakeholders, in response to the recognition that substantial parts of the world were not adequately prepared and did not have adequate resources to prevent, detect, and respond to infectious disease threats (Sands, Mundaca-Shah, & Dzau, 2016). GHSA is committed to supporting collaborative capacity-building efforts by helping countries develop and maintain the core animal and human disease surveillance, pandemic preparedness, and outbreak response capacities mandated by WHO's International Health Regulations, the World Organization of Animal Health's Performance of Veterinary Services Pathway, and other global health security frameworks. GHSA, which has received substantial initial financial commitments, is working to improve epidemiologic and laboratory capabilities in dozens of LMICs around the world, through workforce development and training; laboratory improvement and expansion; and enhancement of communications and computing capabilities. Whether it can have a sustained impact will likely depend on whether funding agencies continue to make such efforts a priority, once the initial funds committed to GHSA have been expended (see the *Public Health Infrastructure* chapter for more on GHSA).

Conclusion

The current status of infectious diseases in LMICs reflects both the dramatic progress that has been made in controlling some diseases and the disappointing results to date in controlling others. The eradication of smallpox, the expected imminent eradication of polio, and impressive gains made against measles and neonatal tetanus all demonstrate what can be accomplished, even in the lowest-income countries, with an effective vaccine when concerted efforts are made to ensure that the vaccine reaches those in need. Similar progress in reducing morbidity from dracunculiasis and onchocerciasis demonstrates that, under the right conditions and with available resources, infectious diseases can be controlled through a combination of vector control and avoidance and treatment. At the same time, the reductions in the morbidity and mortality from diarrheal diseases and acute respiratory infections that have been achieved provide clear evidence that a combination of improved knowledge and access to reasonably inexpensive treatment modalities can also be highly effective.

Remarkable progress has been made in substantially reducing the geographic extent of and mortality from malaria, despite the absence of an effective vaccine. Similarly, by making antiretroviral treatment with inexpensive, generic medications widely available, the morbidity and mortality from HIV/AIDS have been reduced in many LMICs. Progress against TB and dengue, however, has proved harder to achieve. Moreover, recent explosive outbreaks of Ebola virus and Zika virus infections have demonstrated how quickly an infectious agent can spread, reinforcing the need for ongoing vigilance in the form of disease surveillance and adequate global preparedness to respond quickly to regional and global infectious disease threats.

The progress made to date in controlling the morbidity and mortality from infectious diseases in LMICs demonstrates that much can be accomplished even in the absence of marked improvements in socioeconomic conditions. Ultimately, though, widespread improvements in education and socioeconomic conditions will be needed if such progress is to be maintained.

Acknowledgments

The authors acknowledge Christina Phares and Kelly Walblay for their assistance in the preparation of this chapter.

Discussion Questions

1. What are the major types of approaches that have been used to prevent morbidity and mortality from infectious diseases in LMICs?
2. What are the major obstacles that have had, and remain, to be overcome in implementing various approaches to preventing morbidity and mortality from infectious diseases in LMICs?
3. In the current year, which infectious diseases account for the most mortality in LMICs? The most morbidity?
4. If you were working in the Ministry of Health of a low-income country and needed to set priorities concerning resource allocation, how could you go about determining the relative importance of various infectious diseases as causes of mortality in your country? The causes of morbidity/disability?

References

Aaby, P., Samb, B., Simondon, F., Knudsen, K., Seck, A. M., Bennett, J., . . . Whittle, H. (1994). Sex-specific differences in mortality after high-titre measles immunization in rural Senegal. *Bulletin of the World Health Organization*, 72, 761–770.

Bach, J. F., Chalons, S., Forier, E., Elana, G., Jouanelle, J., Kayemba, S., . . . Berchel, C.(1996). 10-year educational programme aimed at rheumatic fever in two French Caribbean islands. *Lancet*, 347, 644–648.

Bausch, D. G., Borchert, M., Grein, T., Roth, C., Swanepoel, R., Libande, M. L., . . . Rollin, P. E. (2003). Risk factors for Marburg hemorrhagic fever, Democratic Republic of the Congo. *Emerging Infectious Diseases*, 9(12), 1531–1537.

Berman, S. (1991). Epidemiology of acute respiratory infections in children of developing countries. *Reviews of Infectious Diseases, 13,* S454–S462.

Bhatt, S., Getting, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., . . . Hay, S. I. (2013). The global distribution and burden of dengue. *Nature, 496,* 504–507.

Bhosai, S. J., Bailey, R. L., Gaynor, B. D., & Lietman, T. M. (2012). Trachoma: An update on prevention, diagnosis, and treatment. *Current Opinion in Ophthalmology, 23*(4), 288–295.

Bhutta, Z. A., Das, J. K., Walker, N., Rizvi, A., Campbell, H., Rudan, I., Lancet Diarrhoea and Pneumonia Interventions Study Group. (2013). Interventions to address deaths from childhood pneumonia and diarrhea equitably: What works and at what cost? *Lancet, 381*, 1417–1429.

Black, R. E. (1993). Persistent diarrhea in children of developing countries. *Pediatric Infectious Disease Journal, 12,* 751–761.

Black, R. E., Morris, S. S., & Bryce, J. (2003). Where and why are 10 million children dying every year? *Lancet, 361,* 2226–2234.

Bowden, S. (2001). New hepatitis viruses: Contenders and pretenders. *Journal of Gastroenterology and Hepatology, 16*(2), 124–131.

Brady, O. J., Gething, P. W., Bhatt, S., Messina, J. P., Brownstein J. S., Hoen, A. G., . . . Hay, S. I. (2012). Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Neglected Tropical Diseases, 6*(8), e1790.

Bundy, D. A., Hall, A., Medley, G. F., & Savioli, L. (1992). Evaluating measures to control intestinal parasitic infections. *World Health Statistics Quarterly, 45,* 168–179.

Butterworth, A. E., Dunne, D. W., Fulford, A. J., Ouma, J. H., & Sturrock, R. F. (1996). Immunity and morbidity in *Schistosoma mansoni* infection: Quantitative aspects. American Journal of Tropical Medicine and Hygiene, 55, 109–115.

Castle, P. E., & Maza, M. (2016). Prophylactic HPV vaccination: Past, present, and future. *Epidemiology & Infection, 144,* 449–468.

Celentano, D. D., Nelson, K. E., Suprasert, S., Eiumtrakul, S., Tulvatana, S., Kuntolbutra, S., . . . Theetranont, C. (1996). Risk factors for HIV-1 seroconversion among young men in northern Thailand. *Journal of the American Medical Association*, 275, 122–127.

Centers for Disease Control and Prevention (CDC). (2002). Progress toward global eradication of poliomyelitis. *Morbidity and Mortality Weekly Report*, 52, 366–369.

Centers for Disease Control and Prevention (CDC). (1996). Progress toward elimination of neonatal tetanus: Egypt, 1988–1994. *Morbidity and Mortality Weekly Report, 45*, 89–92.

Centers for Disease Control and Prevention (CDC). (2003). Outbreak of severe acute respiratory syndrome—worldwide, 2003. *Morbidity and Mortality Weekly Report, 52,* 226–228.

Centers for Disease Control and Prevention (CDC). (2013). Parasites: Hookworm. Fact sheet. Retrieved from https://www.cdc.gov/parasites/hookworm/

Centers for Disease Control and Prevention (CDC). (2017). DPDx: Trichuriasis. Fact sheet. Retrieved from https://www.cdc.gov/dpdx/trichuriasis/index.html

Centers for Disease Control and Prevention (CDC). (2018). Parasites: Ascariasis. Fact sheet. Retrieved from https://www.cdc.gov/parasites/ascariasis/index.html

Chatterjee, A. (2014). The next generation of HPV vaccines: Nonavalent vaccine V503 on the horizon. *Expert Review of Vaccines, 13*(11), 1279–1290.

Cholera Working Group, International Centre for Diarrhoeal Diseases Research, Bangladesh. (1993). Large epidemic of cholera-like disease in Bangladesh caused by *Vibrio cholerae* O139 synonym Bengal. *Lancet*, 342, 387–390.

Clasen T. F., Alexander K. T., Sinclair D., Boisson S., Peletz R., Chang H. H., . . . Cairncross, S. (2015). Interventions to improve water quality for preventing diarrhoea. *Cochrane Database of Systematic Reviews, 10,* CD004794. doi:10.1002/14651858.CD004794.pub3

Cohen, J. (2003). AIDS vaccine still alive as booster after second failure in Thailand. *Science, 302,* 1309–1310.

Cox-Singh, J., Davis, T. M., Lee, K. S., Shamsul, S. S., Matusop, A., Ratnam, S., . . . Singh, B. (2008). *Plasmodium knowlesi* malaria in humans is widely distributed and potentially life threatening. Clinical Infectious Diseases, 46, 165–71.

Cummins, D., McCormick, J. B., Bennett, D., Samba, J. A., Farrar, B., Machin, S. J., & Fisher-Hoch, S. P. (1990). Acute sensorineural deafness in Lassa fever. *Journal of the American Medical Association, 264,* 2093–2096.

Dabbagh, A., Patel, M. K., Dumolard, L., Gacic-Dobo, M., Strebel, P. M., Mulders, M. N., . . . Goodson, J. L. (2017). Progress towards regional measles elimination—worldwide, 2000–2016. *Morbidity and Mortality Weekly Report, 43*, 649–660.

Daneshvar, C., Davis, T. M., Cox-Singh, J., Rafa'ee, M. Z., Zakaria, S. K., Divis, P. C., & Singh, B. (2009). Clinical and laboratory features of human *Plasmodium knowlesi infection. Clinical Infectious Diseases, 49*, 852–860.

De Manzione, N., Salas, R. A., Paredes, H., Godoy, O., Rojas, L., Araoz, F., . . . Tesh, R. B. (1998). Venezuelan hemorrhagic fever: Clinical and epidemiological studies of 165 cases. *Clinical Infectious Diseases, 26*, 308–313.

Dowdle, W. R., & Hopkins, D. R. (1998). *The eradication of infectious diseases: Report of the Dahlem Workshop on the Eradication of Infectious Diseases, Berlin, March 16–22, 1997.* Chichester, UK: John Wiley & Sons.

Doyle, T. J., Bryan, R. T., & Peters, C. J. (1998). Viral hemorrhagic fevers and hantavirus infections in the Americas. *Infectious Disease Clinics of North America, 12,* 95–110.

Eaton, J. W., Rehle, T. M., Jooste, S., Nkambule, R., Kim, A. A., Mahy, M., & Hallett, T. B. (2014). Recent HIV prevalence trends among pregnant women and all women in sub-Saharan Africa: Implications for HIV estimates. *AIDS, 28*(4), S507–S514.

Ellis, W. A. (1998). Leptospirosis. In S. R. Palmer, E. J. L. Soulsby, & D. I. H. Simpson (Eds.), *Zoonoses: Biology, clinical practice, and public health control* (pp. 115–126). New York, NY: Oxford University Press.

El-Rafie, M., Hassouna, W. A., Hirschhorn, N., Loza, S., Miller, P., Nagaty, A., . . . Riyad, S. (1990). Effect of diarrhoeal disease control on infant and childhood mortality in Egypt: Report from the National Control of Diarrheal Diseases Project. *Lancet, 335,* 334–338.

El Samani, E. F., Willett, W. C., & Ware, J. H. (1988). Association of malnutrition and diarrhea in children aged under five years: A prospective follow-up study in a rural Sudanese community. *American Journal of Epidemiology, 128,* 93–105.

Emerson, P. M., Lindsay, S. W., Alexander, N., Bah, M., Dibba, S.-M., Faal, H. B., . . . Bailey, R. L. (2004). Role of flies and provision of latrines in trachoma control: Cluster-randomised controlled trial. *Lancet, 363,* 1093–1098.

Fine, P. E. M., Oblapenko, G., & Sutter, R. W. (2004). Polio control after certification: Major issues outstanding. *Bulletin of the World Health Organization, 82,* 47–52.

Fishbein, D. B. (1991). Rabies in humans. In G. M. Baer (Ed.), *The natural history of rabies* (2nd ed., pp. 519–549). Boca Raton, FL: CRC Press.

Gardon, J., Boussinesq, M., Kamgno, J., Gardon-Wendel, N., Demanga-Ngangue, B., & Duke, O. L. (2002). Effects of standard and high doses of ivermectin on adult worms on *Onchocerca volvulus*: A randomized controlled trial. *Lancet, 360*, 203–210.

Garske, T., Van Kerkhove, M. D., Yactayo, S., Ronveaux, O., Lewis, R. F., Staples J. E., . . . Yellow Fever Expert Committee. (2014). Yellow fever in Africa: Estimating the burden of disease and impact of mass vaccination from outbreak and serological data. *PLoS Medicine, 11*(5), e1001638.

Gera, T., Shah, D., Garner, P., Richardson, M., & Sachdev, H. S. (2016). Integrated management of childhood illness (IMCI) strategy for children under five. *Cochrane Database of Systematic Reviews, 6*. doi:10.1002/14651858.CD010123.pub2

Global Burden of Disease (GBD) 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age–sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet, 385*, 117–171.

Global Burden of Disease Study (GBD) 2015 Child Mortality Collaborators. (2016). Global, regional, national and selected subnational levels of stillbirths, neonatal, infant, and under-5 mortality, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *Lancet, 388,* 1725–1774.

Global Burden of Disease Study (GBD) 2015 Disease and Injury Incidence and Prevalence Collaborators. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. *Lancet, 388*, 1545–1602.

Global Burden of Disease Study (GBD) 2015 Mortality and Causes of Death Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *Lancet, 388*, 1405–1416.

Gove, S. (1997). Integrated management of childhood illness by outpatient health workers: Technical basis and overview. The WHO Working Group on Guidelines for Integrated Management of the Sick Child. *Bulletin of the World Health Organization*, *75*, 7–24.

Greenwood, B. M. (1987). The epidemiology of acute bacterial meningitis in tropical Africa. In J. D. Williams & J. Burnie (Eds.), *Bacterial meningitis* (pp. 61–91). London, UK: Academic Press.

Gubler, D. J. (1998). Dengue and dengue hemorrhagic fever. *Clinical Microbiology Reviews, 11,* 480–496.

Hagan, P. (1996). Immunity and morbidity in infection due to *Schistosoma haematobium. American Journal of Tropical Medicine and Hygiene, 55,* 116–120.

Hammitt, L. L., Crane, R. J., Karani, A., Mutuku, A., Morpeth, S. C., Burbidge, P., . . . Scott, J. A. (2016). Effect of *Haemophilus influenzae* type b vaccination without a booster dose on invasive *H. influenzae* type b disease, nasopharyngeal carriage, and population immunity in Kilifi, Kenya: A 15-year regional surveillance study. *Lancet Glob Health, 4,* e185–e194.

Hampton, L. M., Farell, M., Ramirez-Gonzalez, A., Menning, L., Shendale, S., Lewis, I., . . . Immunization Systems Management Group of the Global Polio Eradication Initiative. (2016). Cessation of trivalent oral poliovirus vaccine and introduction of inactivated poliovirus vaccine—worldwide, 2016. *Morbidity and Mortality Weekly Report*, 65(35), 934-938.

Hart, C. A., Trees, A. J., & Duerden, B. I. (1997). Zoonoses. *Journal of Medical Microbiology, 46,* 4–6.

Helmick, C. G., Webb, P. A., Scribner, C. L., Krebs, J. W., & McCormick, J. B. (1986). No evidence for increased risk of Lassa fever infection in hospital staff. *Lancet, 2,* 1202–1205.

Henao-Restrepo, A. M., Camacho, A., Longini, I. M., Watson, C. H., Edmunds, W. J., Egger, M., . . . Kieny, M. P. (2017). Efficacy and effectiveness of an rVSV-vectored vaccine in

preventing Ebola virus disease: Final results from the Guinea ring vaccination, open-label, cluster-randomised trial. *Lancet, 389,* 505–518.

Heron, M. (2016). Deaths: Leading causes for 2014. *National Vital Statistics Reports, 65*(5).

Hopkins, D. R. (1984). Eradication of dracunculiasis. In P. G. Bourne (Ed.), *Water and sanitation: Economic and sociological perspectives* (pp. 93–114). Orlando, FL: Academic Press.

Hopkins, D. R., Ruiz-Tiben, E., & Ruebush, T. K. (1997). Dracunculiasis eradication: Almost a reality. *American Journal of Tropical Medicine and Hygiene, 57,* 252–259.

Hotez, P. J., Alvarado, M., Basáñez, M., Bolliger, I., Bourne, R., Boussinesq, M., . . . Naghavi, M. (2014). The Global Burden of Disease Study 2010: Interpretation and implications for the neglected tropical diseases. *PLoS Neglected Tropical Diseases, 8*(7), e2865.

Hotez, P. J., Bundy, D. A. P., Beegle, K., Brooker, S., Drake, L., de Silva, N., . . . Savioli, L. (2006). Helminth infections: Soil-transmitted helminth infections and schistosomiasis. In D. T. Jamison, W. H. Mosley, A. R. Measham, & J. L. Bobadilla (Eds.), *Disease control priorities in developing countries* (2nd ed.). New York, NY: Oxford University Press. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK11748/

House, J., Ayele, F., Porco, T. C., Zhou, Z., Hong, K. C., Gebre, T., . . . Lietman, T. M. (2009). Assessment of herd protection against trachoma due to repeated mass antibiotic distributions: A cluster-randomised trial. *Lancet, 373,* 1111–1118.

Htun, Y., Morse, S. A., Dangor, Y., Fehler, G., Radebe, F., Trees, D. L., . . . Ballard, R. C. (1998). Comparison of clinically directed, disease specific, and syndromic protocols for the management of genital ulcer disease in Lesotho. *Sexually Transmitted Infections, 74,* S23–S28.

Hull, H. F., Ward, N. A., Hull, B. P., Milstien, J. B., & de Quadros, C. (1994). Paralytic poliomyelitis: Seasoned strategies, disappearing disease. *Lancet, 343,* 1331–1337.

Hunter, J. M. (1996). An introduction to Guinea worm on the eve of its departure: Dracunculiasis transmission, health effects, ecology and control. *Social Science and Medicine, 43,* 1399–1425.

Hussey, G. D., & Klein, M. (1990). A randomized, controlled trial of vitamin A in children with severe measles. *New England Journal of Medicine, 323,* 160–164.

Institute of Medicine. (1992). *Emerging infections: Microbial threats to health in the United States.* Washington, DC: National Academy Press.

Jahn, A., Floyd, S., Crampin, A. C., Mwaungulu F., Mvula, H., Munthali, F., . . . Glynn, J. R. (2008). Population-level effect of HIV on adult mortality and early evidence of reversal after introduction of antiretroviral therapy in Malawi. *Lancet, 371,* 1603–1611.

Jaiswal, S. P., Jain, A. K., Naik, G., Soni, N., & Chitnis, D. S. (2001). Viral hepatitis during pregnancy. *International Journal of Gynaecology and Obstetrics, 72*(2), 103–108.

Keusch, G. T., & Cash, R. A. (1997). A vaccine against rotavirus: When is too much too much? [Editorial]. *New England Journal of Medicine, 337,* 1228–1229.

Khuroo, M. S., Teli, M. R., Skidmore, S., Sofi, M. A., & Khuroo, M. I. (1981). Incidence and severity of viral hepatitis in pregnancy. *American Journal of Medicine, 70,* 252–255.

Kilgore, P. E., Peters, C. J., Mills, J. N., Rollin, P. E., Armstrong, L., Khan, A. S., & Ksiazek, T. G. (1995). Prospects for the control of Bolivian hemorrhagic fever [Editorial]. *Emerging Infectious Diseases, 1,* 97–100.

Kilian, A. H., Gregson, S., Ndyanabangi, B., Walusaga, K., Kipp, W., Sahlmuller, G., . . .von Sonnenburg, F. (1999). Reductions in risk behaviour provide the most consistent explanation for declining HIV-1 prevalence in Uganda. *AIDS, 13,* 391–398.

King, C. H., Sutherland, L. J., & Bertsch, D. (2015). Systematic review and meta-analysis of the impact of chemical-based molluscicide for control of *Schistosoma mansoni* and *S. haematobium* transmission. *PLoS Neglected Tropical Diseases, 9*(12), e0004290.

Kosek, M., Bern, C., & Guerrant, R. L. (2003). The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bulletin of the World Health Organization, 81,* 197–204.

Kyaw, M. H., Lynfield, R., Schaffner, W., Hadler, J., Reingold, A., Thomas, A. R., . . . Active Bacterial Core Surveillance of the Emerging Infections Program Network. (2006). Effect of introduction of the pneumococcal vaccine on drug-resistant *Streptococcus* pneumonia. *New England Journal of Medicine, 354,* 1455–1463.

LaForce, F. M., Lichnevski, M. S., Keja, J., & Henderson, R. H. (1980). Clinical survey techniques to estimate prevalence and annual incidence of poliomyelitis in developing countries. *Bulletin of the World Health Organization, 58,* 609–620.

Lanata, C. F., Fischer-Walker, C. L., Olascoaga, A. C., Torres, C. X., Aryee, M. J. & Black, R. E. (2013). Global causes of diarrheal disease mortality in children <5 years of age: A systematic review. *PLoS One, 8*(9), e72788.

Laxminarayan, R., Matsoso, P., Pant, S., Brower, C., Rottingen, J., Klugman, K., & Davies, S. (2016). Access to effective antimicrobials: A worldwide challenge. *Lancet, 387*, 168–175.

Lennon, D., Anderson, P., Kerdemelidis, M., Farrell, E., Mahi, S. C., Percival, T., . . . Stewart, J. (2017). First presentation acute rheumatic fever is preventable in a community setting: A school based intervention. *Pediatric Infectious Disease Journal, 36*(12), 1113–1118.

Levett, P. N. (2001). Leptospirosis. *Clinical Microbiology Reviews, 14*(2), 296–326.

Levine, M. M. (2009). Typhoid vaccines ready for implementation. *New England Journal of Medicine, 361,* 403–404.

Liu, J., Platts-Mills, J. A., Juma, J., Kabir, F., Nkeze, J., Okoi, C., . . . Houpt, E. R. (2016). Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: A reanalysis of the GEMS case-control study. *Lancet, 388,* 1291–1301.

Loharikar, A., Dumolard, L., Chu, S., Hyde, T., Goodman, T., & Mantel, C. (2016). Status of new vaccine introduction—worldwide, September 2016. *Morbidity and Mortality Weekly Report, 65*(41), 1136–1140.

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., . . . Memish, Z. A. (2013). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. *Lancet, 380*(9859), 2095–2128.

Luquero, F. J., Rondy, M., Boncy, J., Munger, A., Mekaoui, H., Rymshaw, E., . . . Ciglenecki, I. (2016). Mortality rates during cholera epidemic, Haiti, 2010–2011. *Emerging Infectious Diseases*, *22*(3), 410–416.

MacNeil, J. R., Medah, I., Koussoubé, D., Novak, R. T., Cohn, A. C., Diomandé, F. V. K., . . . Messonnier, N. E. (2014). Neisseria meningitidis serogroup W, Burkina Faso, 2012. *Emerging Infectious Diseases, 20*(3), 394–399.

Madhi, S. A., Cunliffe, N. A., Steele, D., Witte, D., Kirsten, M., Louw, C., . . . Neuzil, K. M. (2010). Effect of human rotavirus

vaccine on severe diarrhea in African infants. *New England Journal of Medicine, 362*, 289–298.

Madhi, S. A., Peterson, K., Madhi, A., Khoosal, M., & Klugman, K. (2000). Increased disease burden and antibiotic resistance of bacteria causing severe community-acquired lower respiratory tract infections in human immunodeficiency virus type 1–infected children. *Clinical Infectious Diseases, 31*, 170–176.

Maiztegui, J. I., McKee, K. T., Jr., Barrera Oro, J. G., Harrison, L. H., Gibbs, P. H., Feuillade, M. R., . . . Peters, C. J. (1998). Protective efficacy of a live attenuated vaccine against Argentine hemorrhagic fever. AHF Study Group. *Journal of Infectious Diseases, 177,* 277–283.

Margolis, H. S., Alter, M. J., & Hadler, S. C. (1997). Viral hepatitis. In A. S. Evans & R. A. Kaslow (Eds.), *Viral infections of humans: Epidemiology and control* (4th ed., pp. 363–418). New York, NY: Plenum Medical Books.

Markowitz, L. E., Nzilambi, N., Driskell, W. J., Sension, M. G., Rovira, E. Z., Nieburg, P., & Ryder, R. W. (1989). Vitamin A levels and mortality among hospitalized measles patients, Kinshasa, Zaire. *Journal of Tropical Pediatrics, 35,* 109–112.

McCormick, J. B., King, I. J., Webb, P. A., Scribner, C. L., Craven, R. B., Johnson, K. M., . . . Belmont-Williams, R. (1986). Lassa fever: Effective therapy with ribavirin. *New England Journal of Medicine, 314,* 20–26.

Medina, E., & Yrarrazaval, M. (1983). Typhoid fever in Chile: Epidemiological considerations. *Revista Medica de Chile, 111,* 609–615.

Meima, A., Smith, W. C. S., van Oortmarassen, G. J., Richardus, J. H., & Habbema, J. D. F. (2004). The future incidence of leprosy: A scenario analysis. *Bulletin of the World Health Organization, 82,* 373–380.

Mendelson, M., Rottingen, J., Gopinatian, U., Hamer, D. H., Wertheim, H., Basnyat, B., . . . Balasegaram, M. (2016). Maximising access to achieve appropriate human antimicrobial use in low-income and middle-income countries. *Lancet, 387*, 188–198.

Mermin, J., Were, W., Ekwaru, J. P., Moore, D., Downing, R., Behumbiize, P., . . . Bunnell, R. (2008). Mortality in HIV-infected Ugandan adults receiving antiretroviral treatment and survival of their HIV-uninfected children: A prospective cohort study. *Lancet, 371*(9614), 752–759.

Mogasale, V., Maskery, B., Ochiai, R. L., Lee, J. S., Mogasale, V. V., Ramani, E., . . . Wierzba, T. F. (2014). Burden of typhoid fever in low-income and middle-income countries: A systematic, literature-based update with risk-factor adjustment. *Lancet Global Health, 2*, e570–e580.

Monson, M. H., Cole, A. K., Frame, J. D., Serwint, J. R., Alexander, S., & Jahrling, P. B. (1987). Pediatric Lassa fever: A review of 33 Liberian cases. *American Journal of Tropical Medicine and Hygiene, 36,* 408–415.

Moore, P. S. (1992). Meningococcal meningitis in sub-Saharan Africa: A model for the epidemic process. *Clinical Infectious Diseases, 14,* 515–525.

Mulholland, K., Hilton, S., Adegbola, R., Usen, S., Oparaugo, A., Omosigho, C., . . . Greenwood, B. (1997). Randomised trial of *Haemophilus influenzae* type b tetanus protein conjugate for prevention of pneumonia and meningitis in Gambian infants. *Lancet, 349*, 1191–1197.

Nair, G. B., Ramamurthy, T., Bhattacharya, S. K., Mukhopadhyay, A. K., Garg, S., Bhattacharya, M. K., . . . Deb, B. C. (1994). Spread of *Vibrio cholerae* O139 Bengal in India. *Journal of Infectious Diseases, 169*, 1029–1034.

Nelson, K. E., Celentano, D. D., Eiumtrakol, S., Hoover, D. R., Beyrer, C., Suprasert, S., . . . Khamboonruang, C. (1996). Changes in sexual behavior and a decline in HIV infection among young men in Thailand. *New England Journal of Medicine, 335,* 297–303.

Newman, L., Rowley, J., Vander Hoorn, S., Wijesooriya, N. S., Unemo, M., Low, N., . . . Temmerman, M. (2015). Global estimates of the prevalence and incidence of four curable sexually transmitted infections in 2012 based on systematic review and global reporting. *PLoS One*, *10*(12), e0143304.

Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team. (2009). Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *New England Journal of Medicine, 360,* 2605–2615.

O'Brien, K. L., Wolfson, L., Watt, J. P., Henkle, E., Deloria-Knoll, M., McCall, N., . . . Hib and Pneumococcal Global Burden of Disease Study Team. (2009). Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: Global estimates. *Lancet, 374*, 893–902.

Ogbu, O., Ajuluchukwu, E., & Uneke, C. J. (2007). Lassa fever in West African sub-region: An overview. *Journal of Vector Borne Diseases, 44,* 1–11.

Omer, S. B. (2017). Maternal immunization. *New England Journal of Medicine, 376,* 1256–1267.

Otten, M. W., Jr., Deming, M. S., Jaiteh, K. O., Flagg, E. W., Forgie, I., Sanyang, Y., . . . Gowers, P. (1992). Epidemic poliomyelitis in the Gambia following the control of poliomyelitis as an endemic disease. I. Descriptive findings. *American Journal of Epidemiology, 135*, 381–392.

Pai, M., & Schito, M. (2015). Tuberculosis diagnostics in 2015: Landscape, priorities, needs, and prospects. *Journal of Infectious Diseases, 211*(S2), S21–S28.

Patriarca, P. A., Wright, P. F., & John, T. J. (1991). Factors affecting the immunogenicity of oral poliovirus vaccine in developing countries: Review. *Reviews of Infectious Diseases, 13,* 926–939.

Pavia, A. T. (2011). Viral infections of the lower respiratory tract: Old viruses, new viruses, and the role of diagnosis. *Clinical Infectious Diseases, 52,* S284–S289.

Pilishvili, T., Chernyshova, L., Bondarenko, A., Lapiy, F., Sychova, I., Cohen, A., . . . Hajjeh, R. (2013). Evaluation of the effectiveness of *Haemophilus influenzae* type b conjugate vaccine introduction against radiologically-confirmed hospitalized pneumonia in young children in Ukraine. *Journal of Pediatrics, 163*, S12–S18.

Platts-Mills, J. A., Babji, S., Bodhidatta, L., Gratz, J., Haque, R., Havt, A., . . . MAL-ED Network Investigators. (2015). Pathogen-specific burdens of community diarrhoea in developing countries: A multisite birth cohort study (MAL-ED). *Lancet Global Health, 3*(9), e564–e575.

PREVAIL II Writing Group; Multi-National PREVAIL II Study Team. Davey, R. T., Jr., Dodd, L., Proschan, M. A., Neaton, J., Neuhaus Nordwall, J., Koopmeiners, J. S., . . . Malvy, D. (2016). A randomized, controlled trial of ZMapp for Ebola virus infection. *New England Journal of Medicine, 375*, 1448–1456.

Rangsin, R., Kana, K., Chuenchitra, T., Sunantarod, A., Mungthin, M., Meesiri, S., . . . Nelson, K. E. (2015). Risk factors for HIV infection among young Thai men during 2005–2009. *PLoS One, 10*(8), e0136555.

Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S., Kaewkungwal, J., Chiu, J., Paris, R., . . . MOPH-TAVEG Investigators. (2009). Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *New England Journal of Medicine, 361,* 2209–2220.

Richardson, V., Hernandez-Pichardo, J., Quintanar-Solares, M., Esparza-Aguilar, M., Johnson, B., Gomez-Altamirano, C. M., . . . Patel, M. (2010). Effect of rotavirus vaccination on death

from childhood diarrhea in Mexico. *New England Journal of Medicine, 362,* 299–305.

Richmond, J. K., & Baglole, D. J. (2003). Lassa fever: Epidemiology, clinical features, and social consequences. *British Medical Journal, 327,* 1271–1275.

Rojanapithayakorn, W., & Hanenberg, R. (1996). The 100% Condom program in Thailand. *AIDS, 10*, 1–7.

RTS,S Clinical Trials Partnership. (2015). Efficacy and safety of RTS,S/A01 malaria vaccine with or without a booster dose in infants and children in Africa: Final results of a phase 3, individually randomised, controlled trial. *Lancet, 386,* 31–45.

Sands, P., Mundaca-Shah, C., & Dzau, V. J. (2016). The neglected dimension of global security: A framework for countering infectious-disease crises. *New England Journal of Medicine, 374*, 1281–1287.

Sazawal, S., & Black, R. E. (1992). Meta-analysis of intervention trials on case-management of pneumonia in community settings. *Lancet, 340,* 528–533.

Schaefer, R., Gregson, S., Eaton, J. W., Mugurungi, O., Rhead, R., Takaruza, A., . . . Nyamukapa, C. (2017). Age-disparate relationships and HIV incidence in adolescent girls and young women: Evidence from Zimbabwe. *AIDS, 31*, 1461–1470.

Schofield, F. (1986). Selective primary health care: Strategies for control of disease in the developing world. XXII. Tetanus: A preventable problem. *Reviews of Infectious Diseases, 8,* 144–156.

Scott, S., Altanseseg, D., Sodbayer, D., Nymadawa, P., Bulgan, D., Mendsaikhan, J., . . . Edmond, K. M. (2013). Impact of *Haemophilus influenzae* type b conjugate vaccine in Mongolia: Prospective population-based surveillance, 2002–2010. *Journal of Pediatrics, 163,* S8–S11.

Selwyn, B. J. (1990). The epidemiology of acute respiratory tract infection in young children: Comparison of findings from several developing countries. Coordinated Data Group of BOSTID Researchers. *Reviews of Infectious Diseases, 12*, S870–S888.

Setia, M. S., Steinmaus, C., Ho, C. S., & Rutherford, G. W. (2006). The role of BCG in prevention of leprosy: A meta-analysis. *Lancet Infectious Diseases, 6,* 162–170.

Shaffer, N., Chuachoowong, R., Mock, P. A., Bhadrakom, C., Siriwasin, W., Young, N. L., . . . Simonds, R. J. (1999). Short-course zidovudine for perinatal HIV-1 transmission in Bangkok, Thailand: A randomised controlled trial. *Lancet, 353*, 773–780.

Shann, F. (1986). Etiology of severe pneumonia in children in developing countries. *Pediatric Infectious Disease, 5,* 247–252.

Shisana, O., Rehle, T., Simbayi, L. C., Zuma, K., Jooste, S., Zungu, N., . . . Onoya, D. (2014). *South African national HIV prevalence, incidence and behaviour survey, 2012.* Cape Town, South Africa: Human Sciences Research Council, South Africa.

Sidikou, F., Zaneidou, M., Alkassoum, I., Schwartz, S., Issaka, B., Obama, R., . . . MenAfriNet consortium. (2016). Emergence of epidemic *Neisseria meningitides* serogroup C in Niger, 2015: An analysis of national surveillance data. *Lancet, 16*(11), 1288–1294.

Sinha, A., Sazawal, S., Kumar, R., Sood, S., Reddaiah, V. P., Singh, B., . . . Bhan, M. K. (1999). Typhoid fever in children aged less than 5 years. *Lancet,* 354, 734–737.

Smolinski, M. S., Hamburg, M. A., & Lederberg, J. (Eds.). (2003). *Microbial threats to health: Emergence, detection, and response.* Washington, DC: National Academies Press.

Sokolow, S. H., Wood, C. L., Jones, I. J., Swartz, S. J., Lopez, M., Hsieh, M. H., . . . De Leo, G. A. (2016). Global assessment of schistosomiasis control over the past century shows targeting the snail intermediate host works best. *PLoS Neglected Tropical Diseases, 10*(7), e0004794.

Sridhar, S. (2009). An affordable cholera vaccine: An important step forward. *Lancet.* doi:10.1016/S0140-6736(09)61418-5

Stanfield, J. P., & Galazka, A. (1984). Neonatal tetanus in the world today. *Bulletin of the World Health Organization, 62,* 647–669.

Sur, D., Lopez, A. L., Kanungo, S., Paisley, A., Manna, B., Ali, M., . . . Clemens, J. D. (2009). Efficacy and safety of a modified killed whole cell oral cholera vaccine in India: An interim analysis of a cluster-randomised, double-blind, placebo-controlled trial. *Lancet, 374,* 1694–1702.

Sur, D., Ochial, R. L., Bhattacharya, S. K., Ganguly, N., Ali, M., Manna, B., . . . Clemens, J. D. (2009). A cluster-randomized effectiveness trial of Vi typhoid vaccine in India. *New England Journal of Medicine, 361,* 335–344.

Sutter, R. W., Cáceres, V. M., & Mas Lago, P. (2004). The role of routine polio immunization in the post-certification era. *Bulletin of the World Health Organization, 82,* 31–39.

Sutter, R. W., Patriarca, P. A., Brogan, S., Malankar, P. G., Pallansch, M. A., Kew, O. M., . . . Alexander, J. P. (1991). Outbreak of paralytic poliomyelitis in Oman: Evidence for widespread transmission among fully vaccinated children. *Lancet, 338,* 715–720.

Taylor, H. R., Burton, M. J., Haddad, D., West, S., & Wright, H., (2014). Trachoma. *Lancet, 384,* 2142–2152.

Ter Meulen, J., Lukashevich, I., Sidibe, K., Inapogui, A., Marx, M., Dorlemann, A., . . . Schmitz, H. (1996). Hunting of peridomestic rodents and consumption of their meat as possible risk factors for rodent-to-human transmission of Lassa virus in the Republic of Guinea. *American Journal of Tropical Medicine and Hygiene, 55*, 661–666.

Thursz, M., Cooke, G. S., & Hall, A. J. (2010). Hepatitis B treatment in resource poor settings: Time for action. *Tropical Medicine and International Health, 15,* 2–4.

Tobian, A. A. R., Kacker, S., & Quinn, T. C. (2014). Male circumcision: A globally relevant but under-utilized method for prevention of HIV and other sexually transmitted infections. *Annual Review of Medicine, 65,* 293–306.

Towner, J. S., Khristova, M. L., Sealy, T. K., Vincent, M. J., Erickson, B. R., Bawiec, D. A., . . . Nichol, S. T. (2006). Marburg virus genomics and association with a large hemorrhagic fever outbreak in Angola. *Journal of Virology, 80*, 6497–6516.

Towner, J. S., Pourrut, X., Albariño, C. G., Nkogue, C. N., Bird, B. H., Grard, G., . . . Leroy, E. M. (2007). Marburg virus infection detected in a common African bat. *PLoS One, 2,* e764.

Trigg, P. I., & Kondrachine, A. V. (1998). Commentary: Malaria control in the 1990s. *Bulletin of the World Health Organization, 76,* 11–16.

UNAIDS. (2013). *2013 global fact sheet.* Geneva, Switzerland: Author.

UNAIDS. (2014). *The gap report 2014.* Geneva, Switzerland: Author.

UNAIDS. (2016a). *AIDS by the numbers.* Geneva, Switzerland: Author.

UNAIDS. (2016b). *Global AIDS update 2016.* Geneva, Switzerland: Author.

UNAIDS. (2017a). *UNAIDS data 2017.* Geneva, Switzerland: Author.

UNAIDS. (2017b). *Protecting the rights of sex workers.* Geneva, Switzerland: Author.

Vainrub, B., & Salas, R. (1994). Latin American hemorrhagic fever. *Infectious Disease Clinics of North America, 8,* 47–59.

Van Boeckel, T. P., Gandra, S., Ashok, A., Caudron, Q., Grenfell, B. T., Levin, S. A., & Laxminarayan, R. (2014). Global antibiotic consumption 2000 to 2010: An analysis of national pharmaceutical sales data. *Lancet Infectious Diseases, 14*, 742–750.

Vinetz, J. M. (2001). Leptospirosis. *Current* Opinion in Infectious Diseases, 14(5), 527–538.

Wakabi, W. (2009). West African meningitis outbreak strains vaccine supplies. *Lancet, 373,* 1836.

Walsh, J. A. (1983). Selective primary health care: Strategies for control of disease in the Watt, J. P., Wolfson, L. J., O'Brien, K. L., Henkle, E., Deloria-Knoll, M., McCall, N., . . . Hib and Pneumococcal Global Burden of Disease Study Team. (2009). Burden of disease caused by *Haemophilus influenzae* type b in children younger than 5 years: Global estimates. *Lancet, 374*, 903–911.

Whalen, C. C., Johnson, J. L., Okwera, A., Hom, D. L., Huebner, R., Mugyenyi, P., . . . Ellner, J. J. (1997). A trial of three regimens to prevent tuberculosis in Ugandan adults infected with the human immunodeficiency virus. Uganda–Case Western Reserve University Research Collaboration. *New England Journal of Medicine, 337,* 801–808.

Wi, T., Lahra, M. M., Ndowa, F., Bala, M., Dillon, J. R., Ramon-Pardo, P., . . . Unemo, M. (2017). Antimicrobial resistance in *Neisseria gonorrhoeae*: Global surveillance and a call for international collaborative action. *PLoS Medicine*. https://doi.org/10.1371/journal.pmed.1002344

Wiktor, S. Z., Ekpini, E., Karon, J. M., Nkengasong, J., Maurice, C., Severin, S. T., . . . Greenberg, A. E. (1999). Short-course oral zidovudine for prevention of mother-to-child transmission of HIV-1 in Abidjan, Côte d'Ivoire: A randomised trial. *Lancet, 353,* 781–785.

World Health Organization (WHO). (1981). Clinical management of acute respiratory infections in children: A WHO memorandum. *Bulletin of the World Health Organization, 59,* 707–716.

World Health Organization (WHO). (1985). *Viral haemorrhagic fevers: Report of a WHO Expert Committee* (WHO Technical Report No. 721). Geneva, Switzerland: Author.

World Health Organization (WHO). (1991). *Management of patients with sexually transmitted diseases: Report of a WHO Study Group* (WHO Technical Report No. 810). Geneva, Switzerland: Author.

World Health Organization (WHO). (1993). *Control of schistosomiasis: The second report of the WHO Expert Committee* (WHO Technical Report No. 830). Geneva, Switzerland: Author.

World Health Organization (WHO). (1994). Expanded programme on immunization, Global Advisory Group. Part II. Achieving the major disease control goals. *Weekly Epidemiological Record,* 69, 29–31, 34–35.

World Health Organization (WHO). (1995). *Control of epidemic meningococcal disease: WHO practical guidelines*. Lyon, France: WHO & Foundation Marcel Mérieux.

World Health Organization (WHO). (1998a). *EPI information system: Global summary, September 1998* (Document WHO/EPI/GEN/98.10). Geneva, Switzerland: Author.

World Health Organization (WHO). (1998b). *Action programme for the elimination of leprosy: Status report 1998* (Document WHO/LEP/98.2). Geneva, Switzerland: Author.

World Health Organization (WHO). (1998c). *Yellow fever technical consensus meeting* (Document WHO/EPI/GEN/98.08). Geneva, Switzerland: Author.

World Health Organization (WHO). (2012a). *Recommendations for the management of common childhood conditions.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2012b). The evolving threat of antimicrobial resistance: Options for action. Retrieved from http://apps.who.int/iris/bitstream/10665/44812/1/9789241503181_eng.pdf

World Health Organization (WHO). (2013). Meningococcal disease in countries of the African meningitis belt, 2012: Emerging needs and future perspectives. *Weekly Epidemiological Record,* 12(88), 129–136.

World Health Organization (WHO). (2014a). *Antimicrobial resistance: Global report on surveillance.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2014b). The end TB strategy. Retrieved from http://www.who.int/tb/strategy/End_TB_Strategy.pdf?ua=1

World Health Organization (WHO). (2015a). Immunization, vaccines and biologicals: Typhoid. Retrieved from http://www.who.int/immunization/diseases/typhoid/en/

World Health Organization (WHO). (2015b). Global action plan on antimicrobial resistance. Retrieved from http://apps.who.int/iris/bitstream/10665/193736/1/9789241509763_eng.pdf

World Health Organization (WHO). (2016b). Integrated Management of Childhood Illness (IMCI). Retrieved from http://www.who.int/maternal_child_adolescent/topics/child/imci/en/

World Health Organization (WHO). (2016c). Global tuberculosis report 2016. Retrieved from http://apps.who.int/iris/bitstream/10665/250441/1/9789241565394-eng.pdf?ua=1

World Health Organization (WHO). (2016d). Leprosy: Fact sheet. Retrieved from http://www.who.int/mediacentre/factsheets/fs101/en/

World Health Organization (WHO). (2016e). *Global leprosy strategy 2016–2020.* New Delhi, India: WHO Regional Office for South-East Asia.

World Health Organization (WHO). (2016f). Overview of the guidelines for the prevention, treatment and management of STIs. In *WHO Guidelines for the Treatment of Treponema pallidum (Syphilis).* Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK384903/

World Health Organization (WHO). (2016g). HIV/AIDS. Retrieved from http://www.who.int/mediacentre/factsheets/fs360/en/

World Health Organization (WHO). (2016h). *Global health sector strategy on HIV 2016–2021.* Geneva, Switzerland: WHO, Department of HIV/AIDS.

World Health Organization (WHO). (2016i). *Global health sector strategy on viral hepatitis 2016–2021.* Geneva, Switzerland: Global Hepatitis Programme, Department of HIV/AIDS.

World Health Organization (WHO). (2016j). World malaria report 2016. Retrieved from http://apps.who.int/iris/bitstream/10665/252038/1/9789241511711-eng.pdf?ua=1

World Health Organization (WHO). (2017a). Immunization, vaccines, and biologicals: Diphtheria. Retrieved from http://www.who.int/immunization/monitoring_surveillance/burden/diphtheria/en/

World Health Organization (WHO). (2017b). Immunizations, vaccines, and biologicals: Pertussis. Retrieved from http://www.who.int/immunization/monitoring_surveillance/burden/vpd/surveillance_type/passive/pertussis/en/

World Health Organization (WHO). (2017c). Neonatal tetanus: Fact sheet. Retrieved from http://www.who.int/immunization

/monitoring_surveillance/burden/vpd/surveillance_type/active/neonatal_tetanus/en/

World Health Organization (WHO). (2017d). Hepatitis C. Retrieved from http://www.who.int/mediacentre/factsheets/fs164/en/

World Health Organization (WHO). (2017e). Chagas disease (American trypanosomiasis). Retrieved from http://www.who.int/mediacentre/factsheets/fs340/en/

World Health Organization (WHO). (2017f). Trypanosomiasis, human African (sleeping sickness). Retrieved from http://www.who.int/mediacentre/factsheets/fs259/en/

World Health Organization (WHO). (2017g). Soil-transmitted helminth infections. Retrieved from http://www.who.int/mediacentre/factsheets/fs366/en/

World Health Organization (WHO). (2017h). Schistosomiasis: Fact sheet. Retrieved from http://www.who.int/mediacentre/factsheets/fs115/en/

World Health Organization (WHO). (2017i). Dracunculiasis (Guinea-worm disease): Fact sheet. Retrieved from http://www.who.int/mediacentre/factsheets/fs359/en/

World Health Organization (WHO). (2017j). *Crossing the billion*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2017k). Human rabies: 2016 updates and call for data. *Weekly Epidemiological Record, 92,* 77–88.

World Health Organization (WHO). (2017l). Dengue and severe dengue: Fact sheet. Retrieved from http://www.who.int/mediacentre/factsheets/fs117/en/

World Health Organization (WHO). (2017m). Hepatitis B. Retrieved from http://www.who.int/mediacentre/factsheets/fs204/en/

World Health Organization (WHO). (2018a). Immunization coverage: Fact sheet. Retrieved from http://www.who.int/mediacentre/factsheets/fs378/en/

World Health Organization (WHO). (2018b). Yellow fever. Retrieved from http://www.who.int/mediacentre/factsheets/fs100/en/

Writing Committee of the Second World Health Organization Consultation on Clinical Aspects of Human Infection with Avian Influenza A (H5N1) Virus. (2008). Update on avian influenza A (H5N1) virus infection in humans. *New England Journal of Medicine, 358,* 261–272.

CHAPTER 7

Nutrition

Christine Stewart, Jessica Fanzo, Keith West, and Robert Black

Introduction

Nutritional concerns in low- and middle-income countries (LMICs) are diverse, ranging from deprivation, hunger, and micronutrient deficiencies that impair health, quality of life, and survival, to a rising tide of obesity and ensuing risks of noncommunicable diseases (NCDs) in every region of the world. Nutritional disorders in early life can have significant health consequences, including increased risk of infectious disease and mortality, poor child cognitive development and performance in school, and increased risk of NCDs in adulthood. Undernutrition, reflected by high prevalence of wasting, stunting, and micronutrient deficiencies, is the predominant form of malnutrition throughout southern Asia and most of sub-Saharan Africa, while overweight and obesity, often in the presence of stunting and micronutrient malnutrition, is now a growing concern globally. Undernutrition and obesity both are problems arising from a complex array of societal, community, and individual risk factors. Food quality and dietary intake are certainly immediate contributors, but so, too, are environmental risks from infectious disease and poor access to health care. From practices within the home to food systems and markets, an individual's nutritional status is influenced by many factors. This presents a challenge for programs to prevent and respond to nutritional problems, as one approach is unlikely to meet all needs in all contexts. Nevertheless, across a wide range of interventions and programs, there is strong evidence of success. Focused, nutrition-specific programs and integrated or multisectoral programs both offer opportunities for improving nutrition, and thereby improving health and well-being throughout life. This chapter covers the burden of malnutrition including who is affected and where, the causes and consequences of malnutrition, and the interventions and approaches to address the disease burden.

Spectrum of Malnutrition and Its Consequences Across the Life Course

Malnutrition may manifest in many forms throughout the life course. Undernutrition in childhood may lead to a pattern of slow, stunted growth and inadequate weight gain. During or prior to pregnancy, undernutrition or inadequate weight gain may have negative health consequences on the mother as well as the developing fetus. These overt patterns of undernutrition may be accompanied by either clinical or subclinical deficiencies in essential micronutrients, described as a form of "hidden hunger." Layered onto this situation in many countries is an increasing problem of excess weight gain leading to overweight and obesity. These problems may co-occur in populations and even within individuals, who may be overweight while

also suffering from a micronutrient deficiency, for example. This has come to be referred to as the "triple burden of malnutrition," reflecting the coexistence of three distinct, but related problems of undernutrition, micronutrient deficiencies, and obesity (Pinstrup-Andersen, 2007).

Undernutrition

Undernourishment comes in many forms. Individuals who are undernourished may have experienced a recent event, such as a food shortage or period of infection that led to an acute period of undernutrition. They may experience periods of wasting, characterized by the loss of muscle and adipose tissue. Sometimes the problem may be longer term in nature, resulting from chronic deprivation, malabsorption in the gut, or chronic illness. Among children, their growth rate may slow or micronutrient deficiencies may develop, as their stored nutrient reserves become depleted. Among pregnant women, undernutrition may lead to poor birth outcomes

Undernutrition in Childhood

The most common first step in detecting undernutrition in childhood is through anthropometric measurements of height, weight, or circumferential measures of the arm or head. To understand whether children are growing well, these measurements need to be compared to a reference population. In 2006, the World Health Organization (WHO) released a set of growth standards that are now used in nearly every country worldwide. The WHO Multicentre Growth Reference Study collected data on approximately 8,500 children representing the regions of North America (United States), Latin America (Brazil), Africa (Ghana), Asia (India), Europe (Norway), and the Middle East (Oman) (WHO Multicentre Growth Reference Study Group, 2006a). Among the resulting standard's advantages are that it reflects patterns of growth among children exposed to healthy environments, such as having been reared in nonsmoking households, provided with appropriate breastfeeding and complementary feeding practices, and rendered standard pediatric care, such as immunizations and care during illness. As such, this standard is considered prescriptive, revealing how child populations should grow, and it firmly establishes the breastfed infant as the model for normative growth and development.

Using the WHO growth standard, weight-for-age, height- or length-for-age, weight-for-length or height, and body mass index (BMI)-for-age growth charts have been created for monitoring children's growth and for comparing populations. For an individual child, measurements can be plotted on the growth chart to monitor trends over time and risk of undernutrition. As examples, three girls are shown in the charts in **FIGURE 7-1**. Child A's weight and length

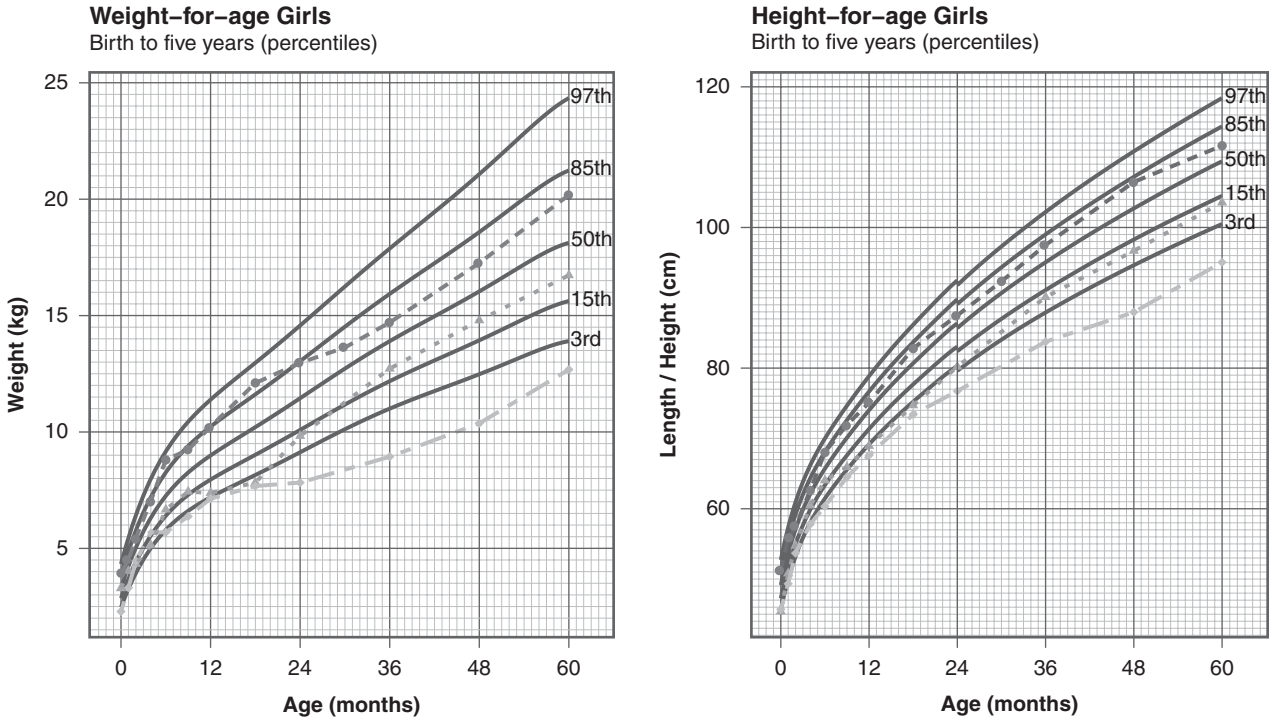


FIGURE 7-1 Weight-for-age and length-for-age growth charts illustrating the growth patterns for three girls. Child A is indicated with a dashed line and circles, Child B with dotted lines and triangles, and Child C with dashed lines and diamonds.

gains roughly track the centile lines, suggesting that she is growing well. In contrast, Child B experienced a period of weight loss at approximately 18 months of age followed by catch-up weight gain due to a short-term event, such as an episode of illness or a period of food deprivation, from which she has recovered. Her length gain over the period was largely unaffected and appears to have stayed on track. Child C, by comparison, has experienced a longer-term problem that is resulting in slower growth in length and weight. Her growth trajectory is crossing the lower centile lines, indicating that she is not growing at the rate that would be expected for her age and sex.

At the population level, *z*-scores are commonly calculated and used for cross-population comparisons or to monitor trends over time. Children who fall below the defined cutoffs of the age- and sex-specific WHO growth standard median would be classified as stunted (low height- or length-for-age), underweight (low weight-for-age), or wasted (low weight-for-length). The degrees of stunting, underweight, or wasting are shown in **TABLE 7-1**. Mid-upper arm circumference (MUAC) is also used as an indicator of wasting, often when access to scales might be limited. Measurements less than 11.5 cm or between 11.5 and 12.5 cm indicate severe or moderate wasting, respectively; these standards are most applicable to children 1 through 4 years, an age when circumferential growth is slow. The WHO Multicentre Growth Study firmly established that populations of children of diverse ethnic and geographic backgrounds with adequate access to health services and nutrition have the same distributions of height and weight (WHO Multicentre Growth Reference Study Group, 2006b). In a healthy population, therefore, we would expect that the median *z*-score of the indicator would be zero and that approximately 2% of children would fall below the –2 cutoff, representing an expected proportion of normal, small children. When the median falls below zero and when population has a high proportion of stunting, wasting, or underweight, that finding is indicative of a significant public health concern. In comparing regional trends in child growth, however, it becomes apparent that sub-Saharan Africa and Asia, and to a lesser degree, Latin America, have evidence of significant population growth faltering (**FIGURE 7-2**) (Victora, de Onis, Hallal, Blossner, & Shrimpton, 2010). These data also reflect that there is increasing severity of stunting with age. Irrespective of region, two key observations are the steadfast decline in attained length through the second year of life, occurring at a rate of approximately –0.10 *z*-score

TABLE 7-1 Interpretation of Anthropometric Indicators in Children

Indicator	Cutoff	Interpretation
Height- or length-for-age *z*-score (HAZ or LAZ)	< –3.0 cm	Severely stunted
	–3.0 < –2.0 cm	Moderately stunted
	–2.0 < –1.0 cm	Mildly stunted
Weight-for-age *z*-score (WAZ)	< –3.0 cm	Severely underweight
	–3.0 < –2.0 cm	Moderately underweight
	–2.0 < –1.0 cm	Mildly underweight
Weight-for-height or length *z*-score (WHZ or WLZ)	< –3.0 cm	Severely wasted
	–3.0 < –2.0 cm	Moderately wasted
	–2.0 < –1.0 cm	Mildly wasted
Mid-upper arm circumference (MUAC)	< 11.5 cm	Severe wasting
	11.5 < – 12.5 cm	Moderate wasting

Data from WHO Multicentre Growth Reference Study Group, 2006a.

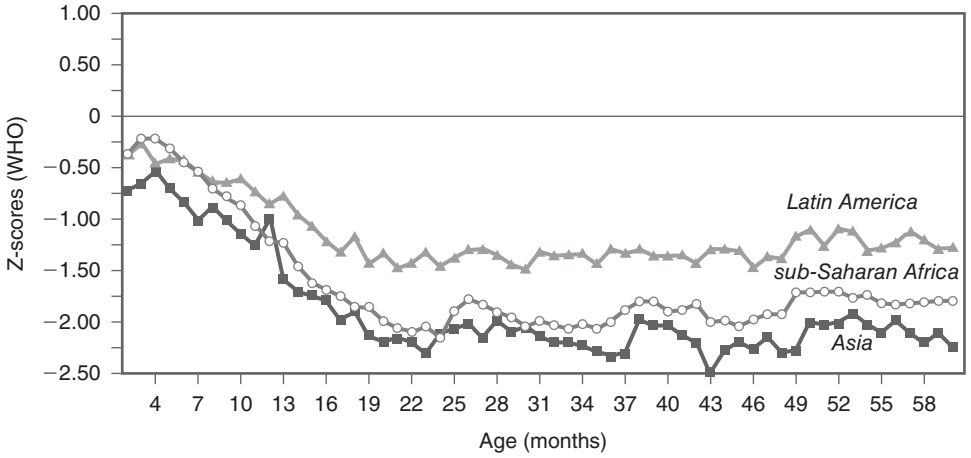


FIGURE 7-2 Mean height-for-age *z*-scores for children 1–59 months relative to the WHO growth standards. Geographic groupings are based on the WHO regions for Latin America and the Caribbean, South Asia, and sub-Saharan Africa.

Data from Victora, C. G., de Onis, M., Hallal, P. C., Blossner, M., & Shrimpton, R. (2010). Worldwide timing of growth faltering: Revisiting implications for interventions. *Pediatrics, 125*(3), e473–480.

per month, and the lack of statural catch-up thereafter. This pattern emphasizes the importance of promoting adequate fetal and postnatal linear growth in children through 2 years of age as critical to resolving lifelong, intergenerational stunting in the developing world.

Although milder stages of undernutrition are often clinically inapparent (**FIGURE 7-3**, left child), severe acute malnutrition (SAM) refers to two clinically distinct presentations, marasmus and kwashiorkor, and a mixed form, marasmic kwashiorkor. *Marasmus*, a term originating from the Greek word *marasmos*, which means "withering," is marked by severe wasting, grossly clinically evident accompanied by a weight-for-height usually less than −3 *z*-scores. In a preschool child, it is diagnosed based on a MUAC less than 11.5 cm, from loss of muscle and adipose; a resultant "baggy" appearance to the skin; moderate to severe stunting from a near cessation of linear growth; soft, sparse hair; absence of edema; alertness; and hunger (Reddy, 1991) (Figure 7-3, right child). *Kwashiorkor* is a term from the Ghanaian language, Ga, for a condition that was initially observed to have developed when an older child was displaced from the breast (Williams, 1933). As implied by the term, children 2 to 3 years of age are at highest risk of this condition. Evidence of kwashiorkor includes edema (an essential feature), milder wasting, reddish hair changes, enlarged liver, frequent dermatosis (flaky paint rash), and a state of misery and disinterest in food (Reddy, 1991). Children with either marasmus or kwashiorkor are at high risk for corneal xerophthalmia due to severe vitamin A deficiency (Sommer & West, 1996). While both conditions are considered to result from severe deficiencies in energy and protein, their distinct clinical and biochemical profiles (Di Giovanni et al., 2016) suggest that this description is an oversimplification, with differences in other nutrient deficiencies and oxidative stress offering potential explanations for their occurrence (Golden, 2002). The case fatality rate for severe acute malnutrition is high (Brown et al., 1981).

FIGURE 7-3 Mild underweight (clinically normal appearance) and marasmus (evident by extreme wasting of limbs and torso) in 1-year-old Bangladeshi fraternal twins. The child on the left is a boy, and the child on the right is a girl.

West, K. P. Jr., Caballero, B., & Black, R. E. (2001). Nutrition. In Merson, M. H., Black, R.E., Mills, A. J. (Eds.), "*International Public Health*" (Chapter 5, Figure 5–7, pp. 229). Gaithersburg, MD: Aspen Publishers, Inc. Photo: Keith West.

Undernutrition During Pregnancy

When it occurs during pregnancy, undernutrition has significant consequences for both the mother and her developing fetus. Intrauterine growth restriction (IUGR) may develop in the fetus if the mother begins her pregnancy malnourished or does not gain a sufficient amount of weight during gestation (Institute of Medicine [IOM], 2003). Other factors might also cause IUGR, such as smoking or pregnancy-induced hypertension. IUGR is assessed in a variety of ways. Historically, the primary criteria was low birth weight (LBW; less than 2,500 g), although a lower weight at birth may be due to either IUGR or to earlier gestational age at delivery (Kramer, 1987). Infants born preterm (less than 37 weeks' gestation) may have lower weight at birth, but might not be growth restricted. A child with a weight or length below the 10th percentile of the sex- and gestational age-specific growth reference would be considered small for gestational age (SGA) (Alexander, Himes, Kaufman, Mor, & Kogan, 1996).

Unlike with the child growth standards, there is some debate over whether country- or ethnic-specific fetal growth charts should be used. Two sets of fetal growth standards for international use have been released in recent years. In 2014, an international consortium (INTERGROWTH-21st) released a set of fetal growth standards based on data from 60,000 pregnancies in Europe (Turin, Italy, and Oxford, United Kingdom), Asia (Muscat, Oman; Beijing, China; and Nagpur, India), North America (Seattle, Washington), South America (Pelotas, Brazil), and Africa (Nairobi, Kenya) (Villar et al., 2014). This study found relatively similar patterns of fetal growth across participating countries, supporting the idea that a single growth standard could be used globally. In 2017, however, WHO examined fetal growth patterns using data from 1,387 pregnancies in Europe (Copenhagen, Denmark; Paris, France; Hamburg-Eppendorf, Germany; Bergen, Norway), Asia (Delhi, India, and Khon Kaen, Thailand), South America, (Rosario, Argentina, and Campinas, Brazil), and Africa (Assiut, Egypt, and Kinshasa, Democratic Republic of Congo) (Kiserud et al., 2017). This study found significant heterogeneity in growth patterns between countries that could not be fully explained by differences in maternal height, health during pregnancy, or other factors.

Adequate maternal weight gain during pregnancy is essential for a healthy pregnancy. Weight gain recommendations vary, depending on the prepregnancy maternal BMI (**TABLE 7-2**). These recommendations are based on evidence that women who gain insufficient weight are at a greater risk of adverse birth outcomes, including stillbirth, SGA, and preterm birth (Goldstein et al., 2017; IOM, 2009). In contrast, women who gain too much weight are at risk of developing gestational diabetes mellitus (GDM) and hypertensive disorders during pregnancy, as well as retaining excess weight postpartum. Their infants are at elevated risk of requiring a cesarean delivery and developing obesity in later childhood. Recently, other research groups have evaluated whether these standards are appropriate for use in low-income country settings. A study of healthy pregnant women in Malawi who delivered live-born singleton, normal-weight, term infants reported lower weight gain during pregnancy than recommended by the IOM (Xu, Luntamo, Kulmala, Ashorn, & Cheung, 2014), drawing into question the usage of the North American standard for use in African populations.

TABLE 7-2 Weight Gain Recommendations for Pregnancy

Prepregnancy Weight Category	Body Mass Index (kg/m^2)	Recommended Range of Weight Gain (kg)	Recommended Rates of Weight Gain in the Second and Third Trimesters (range [kg/wk])
Underweight	Less than 18.5	12.5–18	0.44–0.58
Normal weight	18.5–24.9	11.5–16	0.35–0.50
Overweight	25–29.9	7–11.5	0.23–0.33
Obese	30 and greater	5–9	0.17–0.27

Modified from Institute of Medicine. (2009). *Weight gain during pregnancy: Reexamining the guidelines* (Table S-1, pp. 2). Washington, DC: National Academies Press. Copyright 2009 by the National Academy of Sciences.

Population Trends

Estimates of the prevalence of stunting, wasting, and underweight compiled by the United Nations Children's Fund (UNICEF), WHO, and the World Bank suggest that there were nearly 155 million children stunted (HAZ < −2) and 16.9 million severely wasted (WHZ < −3) in 2016 (**TABLE 7-3**). These conditions afflicted nearly 23% and 2.5% of children, respectively. Although the highest prevalence of stunting occurs in Eastern Africa, where 36.7% of children are stunted, the largest number of children who are stunted—61 million—live in South Asia. As an indicator of prenatal

TABLE 7-3 Estimated Prevalence (%) and Number (in millions) of Children Affected by Moderate to Severe Stunting (HAZ < −2) or Severe Wasting (WHZ < −3) by UN Region in 2016

	Stunting		Severe Wasting	
	Percentage Stunted	Number Stunted	Percentage Wasted	Number Wasted
Global	22.9	154.8	2.5	16.9
Developing regions	25.0	151.9	2.8	16.9
Africa	31.2	59.0	2.2	4.1
Eastern Africa	36.7	24.0	1.6	1.1
Middle Africa	32.5	8.9	2.1	0.6
Northern Africa	17.6	5.0	3.5	1.0
Southern Africa	28.1	1.8	1.0	0.1
Western Africa	31.4	19.2	2.3	1.4
Asia	23.9	86.5	3.5	12.6
Central Asia	12.5	1.0	1.4	0.1
Eastern Asia	5.5	4.9	0.4	0.4
Southern Asia	34.1	61.2	5.0	9.0
Southeastern Asia	25.8	15.1	4.7	2.7
Western Asia	15.7	4.4	1.1	0.3
Latin America and Caribbean	11.0	5.9	0.3	0.2
Caribbean	5.3	0.2	1.1	0.0
Central America	15.4	2.5	0.2	0.0
South America	9.5	3.2	0.2	0.1
Oceania	38.3	0.5	3.3	0.0

Data from UNICEF, WHO, & World Bank, 2017.

undernutrition, an estimated 15% to 20% of infants are born with low birth weight, although there is considerable uncertainty about this estimate because nearly half of infants are not weighed at birth (WHO, 2014a).

Over the past 30 years, the prevalence of most forms of undernutrition has been declining globally. Notably, the prevalence of stunting declined from 39.5% in 1990 to 22.9% in 2016 (UNICEF, WHO, & World Bank, 2017), resulting in nearly 100 million fewer children stunted. Despite these impressive gains, the average rate of reduction has not been on track to achieve the World Health Assembly targets (**EXHIBIT 7-1**). The declines also have been uneven across regions. The fastest declines have been in Asia, particularly East Asia, where there was a decline from 48% in 1990 to 24% in 2016 (**FIGURE 7-4**). By contrast, the corresponding decline in Africa was from 42% to 32%. This was offset by the overall population growth in the African region, however, which means that there were actually 12.5 million more stunted children in Africa in 2016 than in 1990.

EXHIBIT 7-1 Using Data to Track Progress Towards Achieving Nutrition Goals

Since the launch of the Millennium Development Goals in 2000 and the release of the Sustainable Development Goals in 2015, it has been increasingly recognized that the effective use of data can stimulate global development efforts. In 2012, the World Health Assembly endorsed a set of six global nutrition targets that it aimed to achieve by 2025:

- A 40% reduction in the number of children younger than age 5 who are stunted
- A 50% reduction in anemia among women of reproductive age
- A 30% reduction in low birth weight
- No increase in child overweight
- A 50% increase in the rate of exclusive breastfeeding in the first 6 months
- Maintenance or reduction in child wasting to less than 5% (WHO, 2014b)

At the time that these targets were adopted, few countries were routinely collecting data on all six of these indicators. The Global Nutrition Report, which was launched in 2014, was designed as a mechanism to track countries' progress and hold them accountable to achieving the targets, highlight success stories, and serve as means of advocating for change. In 2016, only 41 of 193 countries were considered to be on track toward meeting the stunting target, but 79 lacked sufficient data to track progress (International Food Policy Research Institute [IFPRI], 2016). Only 36 were on track to meet the exclusive breastfeeding target, but 110 lacked data on this issue. Moreover, despite more complete data with which to monitor trends, only 3 countries were on target to meet the anemia target. Routine data collection on the prevalence of low birth weight has been notably absent in many countries, rendering it impossible to track progress on achieving the goal. In many places, this may be due to a lack of skilled birth attendants or inadequate equipment available to weigh newborns. There is a need for a global data revolution to aid in nutrition monitoring efforts, to track progress at the country level, and to better target policies and programs to the most vulnerable populations that could stand to benefit most.

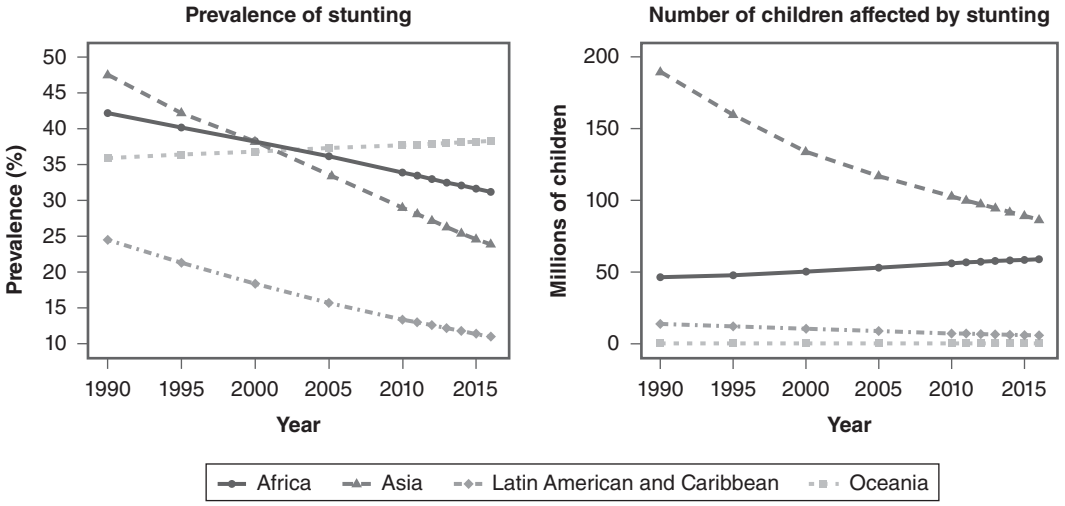


FIGURE 7-4 Trends in the prevalence and number of children affected by stunting from 1990 to 2016.

Data from UNICEF, WHO, & World Bank, 2017.

Health Consequences

Undernutrition is a risk factor for infectious morbidity and associated case fatality among children in low-income countries. Among neonates, given that gestational age and age-adjusted size at birth are often correlated with nutritional status during early childhood, it is appropriate to consider these factors as part of a continuum of risk between nutrition and infection. Nevertheless, studies have generally examined separately the risk of mortality related to status at birth in the neonatal or infant period and the risk related to nutritional status throughout childhood, at least up to age 5 years.

Low Birth Weight: Morbidity and Mortality Risk. To examine the role of nutritional risk factors for morbidity and mortality early in life, it is important to evaluate births that have been subjected to fetal growth restriction, as characterized by their being small for gestational age. There is a larger body of literature on the risks of LBW that does not distinguish whether this condition was due to SGA, to preterm birth, or to both. Although recent analyses do provide the specific effects of SGA, the effects of LBW will also be considered because this metric remains a common measure of birth status when the gestational age of the newborn is difficult to determine.

LBW babies have the largest increase in risk during the neonatal period, and they continue to have additional risk in the post-neonatal period of infancy. In high-income countries, birth weights of 3,500 to 4,500 grams (7 lb, 11 oz to 9 lb, 14 oz) are associated with the lowest risk of neonatal mortality. The relative risk of neonatal mortality increases with birth weights less than 2,500 grams and increases even more for very LBW babies of less than 1,500 grams (3 lb, 3 oz) (Ashworth, 1998). In low-income countries, the data are more limited due to the difficulty of obtaining accurate birth weight measurements for deliveries, which occur predominantly in the home. In these settings, neonatal mortality rates typically range as high as 50 deaths per 1,000 live births, and increase with decreasing birth weight. A pooled analysis of 20 birth cohorts with data on more than 2 million live births from Africa, Asia, and Latin America provides the relative risk of mortality associated with SGA (Katz et al., 2013). The pooled relative risk for babies who were SGA (birth weight in the lowest 10th percentile compared with a reference population) was 1.83 (95% confidence interval [CI], 1.34–2.50) for neonatal mortality and 1.90 (95% CI, 1.32–2.73) for post-neonatal mortality.

Because of the importance of diarrhea and pneumonia as causes of death in children in low-income countries, studies have evaluated whether LBW confers additional risk for deaths from these two causes. In three studies conducted in low-income countries, the increased risk of diarrheal deaths during infancy was 2.5- to 2.8-fold for LBW babies (Ashworth & Feachem, 1985). Likewise, the risk of death from acute lower respiratory infection was increased, but with more variability; the relative risk ranged from 1.6 to 8.0 (Victora et al., 1999).

As might be expected from these studies of infant mortality, LBW has also been shown to be a risk factor for diarrheal and respiratory morbidity (Ashworth & Feachem, 1985; Victora et al., 1999). Studies in Papua New Guinea (Bukenya, Barnes, & Nwokolo, 1991), Thailand (Ittiravivongs, Songchitratna, Ratthapalo, & Pattara-Arechachai, 1991), and Brazil (Victora, Barros, Kirkwood, & Vaughan, 1990) found a relative risk associated with LBW, adjusted for other possible determinants of illness, ranging from 1.6 to 3.9 for acute diarrhea. Studies in China (Chen, Yu, & Li, 1988), Argentina (Cerquerio, Murtagh, Halac, Avila, & Weissenbacher, 1990), and Brazil (Victora et al., 1990; Victora, Smith, Barros, Vaughan, & Fuchs, 1989) found a relative risk ranging from 1.4 to 2.2 for acute lower respiratory infection or pneumonia hospitalization.

Beyond its immediate effects on survival and morbidity, LBW is associated with stunting in later childhood (Christian et al., 2013). Likewise, it is associated with increased risk of chronic diseases in adulthood (Uauy, Kain, & Corvalan, 2011).

Anthropometric Deficits in Children: Morbidity and Mortality Risk. Anthropometric status in early childhood is related to the risk of mortality and subsequent infectious disease. Severe malnutrition clearly carries a high risk of death, but it is the interaction of undernutrition with common infectious diseases in low-income countries that is of greater importance for mortality. This relationship reflects the synergy between undernutrition and infectious diseases in which the combined conditions result in much greater mortality than either would alone (Scrimshaw, Taylor, & Gordon, 1968). Compromised host defenses and more severe infectious diseases in undernourished children may lead to higher case fatality rates, although an increased incidence of infectious diseases may play some role as well.

A review of the observational population studies in a large number of countries clearly reveals a monotonic increase in the risk of mortality at each progressively worse level of anthropometric status compared to the WHO growth standards or other reference populations (Caulfield, de Onis, Blossner, & Black, 2004; Olofin et al., 2013; Pelletier, Frongillo, & Habicht,

1993). In the most recent pooled analysis of 10 prospective studies from Africa, Asia, and South America, all degrees of anthropometric deficits increased the hazards of dying (**TABLE 7-4**) (Olofin et al., 2013).

It is important to note that even mild to moderate undernutrition puts a child at increased risk of mortality. Because the greatest number of undernourished children have mild to moderate undernutrition rather than severe malnutrition, most of the excess risk of death is attributable to the less severe forms of undernutrition. Also important is the observation that with worsening levels of anthropometric status, the mortality rate increases logarithmically. An estimated 14.7% and 12.6% of all childhood deaths have been attributed to stunting and wasting, respectively (Black et al., 2013).

Diarrheal diseases and pneumonia are the two most important infectious causes of death in children

TABLE 7-4 Hazard Ratio (HR)[a] Estimates for Child Mortality Using WHO 2006 Standards

	All Deaths	Mortality from Respiratory Tract Infections	Mortality from Diarrheal Disease	Mortality from Other Infectious Causes[b]	Mortality from Malaria	Mortality from Measles
	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Weight-for-Age z-Score						
< –3	9.40 (8.02, 11.03)	10.10 (6.53, 15.64)	11.56 (8.63, 15.48)	8.28 (4.32, 15.89)	1.29 (0.39, 4.29)	7.73 (4.15, 14.39)
–3 to < –2	2.63 (2.20, 3.14)	3.11 (1.93, 5.02)	2.86 (2.03, 4.03)	1.58 (0.73, 3.45)	1.65 (0.77, 3.53)	3.12 (1.67, 5.80)
–2 to < –1	1.52 (1.28, 1.81)	1.85 (1.17, 2.91)	1.73 (1.24, 2.40)	1.54 (0.78, 3.03)	1.26 (0.66, 2.39)	1.00 (0.49, 2.03)
≥ –1	Ref	Ref	Ref	Ref	Ref	Ref
Height/Length-for-Age z-Score						
< –3	5.48 (4.62, 6.50)	6.39 (4.19, 9.75)	6.33 (4.64, 8.65)	3.01 (1.55, 5.82)	1.92 (0.89, 4.11)	6.01 (3.00, 12.07)
–3 to < –2	2.28 (1.91, 2.72)	2.18 (1.39, 3.43)	2.38 (1.71, 3.31)	1.86 (0.97, 3.57)	1.06 (0.48, 2.32)	2.79 (1.40, 5.56)
–2 to < –1	1.46 (1.23, 1.74)	1.55 (1.02, 2.37)	1.67 (1.20, 2.30)	0.95 (0.48, 1.87)	0.74 (0.35, 1.56)	1.25 (0.61, 2.58)
≥ –1	Ref	Ref	Ref	Ref	Ref	Ref
Weight-for-Length/Height z-Score						
< –3	11.63 (9.84, 13.76)	9.68 (6.07, 15.43)	12.33 (9.18, 16.57)	11.21(5.91, 21.27)	1.24 (0.17, 9.29)	9.63 (5.15, 18.01)
–3 to < –2	3.38 (2.86, 3.98)	4.66 (3.07, 7.09)	3.41 (2.52, 4.63)	2.73 (1.35, 5.54)	1.43 (0.52, 3.94)	2.58 (1.32, 5.06)
–2 to < –1	1.62 (1.41, 1.87)	1.92 (1.31, 2.84)	1.60 (1.23, 2.11)	1.65 (0.98, 2.79)	0.86 (0.39, 1.90)	1.02 (0.56, 1.85)
≥ –1	Ref	Ref	Ref	Ref	Ref	Ref

Abbreviations: CI, 95% confidence interval.

[a] Adjusted for age (as the time scale, in weeks), child's sex, cohort characteristics, and assigned treatment (in randomized trials).

[b] Septicemia, unspecified febrile illness, tuberculosis, meningitis, hepatitis, or cellulitis.

Reproduced from Olofin, I., et al. (2013). Associations of suboptimal growth with all-cause and cause-specific mortality in children under five years: A pooled analysis of ten prospective studies. *PLoS One, 8*(5): e64636.

younger than 5 years of age in low-income countries (Liu et al., 2016). These two conditions have a higher case fatality rate in undernourished children compared to those who are well nourished. For example, malnourished children who were discharged from a hospital after treatment for diarrheal illness in Bangladesh had a 14-fold greater risk of dying compared with better-nourished controls (Roy, Chowdhury, & Rahaman, 1983). In a community-based study in rural India, severely wasted children had a 24-fold higher diarrheal case fatality rate compared with better-nourished children (Bhandari, Bhan, & Sazawal, 1992). In Mexico, undernourished children had an 8-fold greater risk of death with severe diarrhea (Tome, Reyes, Rodriguez, Guiscafre, & Gutierrez, 1996). Studies in the Philippines (Tupasi et al., 1990), Papua New Guinea (Shann, Barker, & Poore, 1989), Bangladesh (Rahman et al., 1990), and Argentina (Weissenbacher et al., 1990) of acute lower respiratory infection and pneumonia found a 2- to 3-fold increase in the case fatality rate in undernourished versus better-nourished children. While measles fatality also rises with worsening nutritional status, fatality from malaria does not show this clear trend (Olofin et al., 2013).

Nutritional status has been assessed widely as a risk factor for the severity and incidence of diarrheal diseases. As might be expected with the reported higher case fatality rate in undernourished children, nutritional status has an association with disease severity. One measure of this relationship is the duration of the illness. In studies in several countries, the duration of diarrhea in mildly to moderately malnourished children was shown to be up to 3-fold longer than for better-nourished children (Black, Brown, & Becker, 1984a). Undernutrition is also an important risk factor for the occurrence of persistent diarrhea (Baqui et al., 1993). This relationship has been demonstrated for specific types of diarrhea, such as episodes due to *Shigella* species and enterotoxigenic *Escherichia coli*, in which the illness duration in Bangladeshi children was 2.5-fold longer in undernourished children (Black, Brown, & Becker, 1984b). Poorer nutritional status has also been documented to be associated with an increased rate of stool output in children, leading to an increased risk of dehydration (Black et al., 1984a). These mechanisms likely explain the increased risk for hospitalization and length of hospital stay due to diarrhea in undernourished children (Man et al., 1998; Victora et al., 1990).

The effect of undernutrition on diarrheal incidence has been more variable in different settings. Some studies have not found any increased risk of overall incidence in undernourished children, whereas others have found a 30% to 70% increased risk of incident diarrhea in such children (Black et al., 1984a; Checkley et al., 2002; El Samani, Willett, & Ware, 1986; Guerrant, Schorling, McAuliffe, & de Souza, 1992). Some enteropathogens causing diarrhea, such as *Cryptosporidium parvum*, may be particularly selected as a cause of infection in undernourished children (Checkley et al., 1998).

The role of undernutrition as a risk factor for severity or incidence of acute lower respiratory infections or pneumonia has also been the subject of extensive study. Evidence supports an association between undernutrition and severity of respiratory infection, with such relationships being found in both hospital-based and community-based studies. Undernutrition has been found to increase the likelihood that a child will have bacteremia, pleural effusion, and other complications (Johnson, Aderele, & Gbadero, 1992). Studies in the Philippines, Costa Rica (James, 1972), and Brazil (Fonseca et al., 1996; Victora et al., 1989) found a modestly increased incidence of acute lower respiratory infections in undernourished children. In contrast, studies in Papua New Guinea (Smith, Lehman, Coakley, Spooner, & Alpers, 1991), Uruguay (Selwyn, 1990), and Guatemala (Cruz et al., 1990) did not find any such relationship. A meta-analysis found that underweight status (less than −2 weight-for-age *z*-scores) confers an increased relative risk of pneumonia of 1.86 (95% CI, 1.06–3.28) (Fishman et al., 2004). One study in Gambia found that the development of pneumococcal infection was associated with a history of poor weight gain prior to the illness compared with other children in the community (O'Dempsey et al., 1996).

Early observations reported that malaria was more frequent in those persons with inadequate diets or with undernutrition (Garnham, 1954). Later studies appeared to show that malnutrition could actually be protective against malaria. This contention was based on observational studies (Hendrickse, Hasan, Olumide, & Akinkunmi, 1971) and on interventions such as the refeeding of famine victims (Murray, Murray, Murray, & Murray, 1975) in which an exacerbation of malaria appeared to occur. Animal studies seemed to support the suppressive effects of a poor diet on malaria, leading to the belief that undernourished children are less susceptible to the infection and consequences of malaria. If refeeding a malaria-infected, starved host can reactivate the low-grade infection and lead to more severe disease, antimalarial measures should be included during nutritional rehabilitation of famine victims who are likely to have malaria infection.

In addition to its association with increased risk of infection, chronic undernutrition has insidious effects on brain development and cognitive function, which in turn hinders an individual's capacity to perform well in school and seek economic opportunities in

adulthood (Black et al., 2017). In a systematic review of numerous observational studies, it was estimated that each standard deviation (SD) increase in height-for-age *z*-score among children younger than 2 years is associated with an approximate 0.2 SD increase in cognitive abilities in early childhood—an effect size that largely persists into later childhood (Sudfeld et al., 2015). Longitudinal studies of malnourished children have also found long-term effects on occupational status, economic productivity, and wages (McGovern, Krishna, Aguayo, & Subramanian, 2017). Precise estimates of these effects are difficult to develop because of variations in the measurement of the exposure and outcome in each of the studies, but cross-sectional studies of adults have suggested that each 1 cm increase in stature is associated with a 4% increase in wages among men and a 6% increase in wages among women (McGovern et al., 2017).

Micronutrient Deficiencies

Overview

The term "micronutrients" refers to vitamins and minerals that must be provided in minute amounts in the human diet to support essential cellular and tissue functions of growth, development, maturation, and homeostasis at each stage of life (Ames, 2010; Fenech, 2010; Gernand, Schulze, Stewart, West, & Christian, 2016). Discovered through animal experiments and clinical studies over the past century, micronutrients known to be required from the diet, always or conditionally, to maintain health include four lipid-soluble vitamins (A, D, E, and K) and nine water-soluble vitamins (thiamin, riboflavin, niacin, pyridoxine, folate, and cobalamin [collectively termed "B-complex"], plus vitamin C, biotin, and choline), plus 12 inorganic elements, including iron, zinc, selenium, and copper (IOM, 2005). Guided by the Dietary Reference Intakes (IOM, 2005), which provide age-, sex-, and pregnancy and lactation–specific estimates of nutrient intake adequacy, a diet diverse in vegetables, fruits, whole grains, and animal-source foods (meat, poultry, fish, eggs, milk, and cheese) is usually adequate in micronutrients. Carefully planned vegetarian diets can also meet most vitamin and mineral requirements, except for vitamin B_{12}, for which animal foods are major sources. Micronutrient requirements are assumed to be largely met in the first 6 months of life through adequate breastfeeding by well-nourished mothers (IOM, 2005).

Micronutrient intakes at either extreme may adversely affect health. Deficient intakes are unable to meet normal requirements, perturbing genomic, translational, energetic, and stress response pathways that may adversely affect multiple immune, neural, and metabolic functions and disease risk (Ames, 2010). Consequences of deficiencies range from nonspecific effects (e.g., impaired cognition or growth) to nutrient-specific pathological conditions (e.g., xerophthalmia from vitamin A or goiter from iodine deficiencies), depending on the nutrients involved and their interactions, the duration and severity of the deficiency, life stage, and exacerbating conditions such as inflammation from infection, chronic disease, or toxic exposures (Kutlesic et al., 2017). Evidence is emerging that the gut microbiome, depending on its community structure and interactions, may affect host micronutrient nutriture by enhancing or inhibiting nutrient uptake, synthesizing vitamin intermediates, or otherwise affecting micronutrient metabolic pathways (Biesalski, 2016). Gene mutations that lead to a defect in nutrient uptake or metabolism (e.g., a polymorphism in a beta-carotene cleavage enzyme in the intestine required for conversion to vitamin A [Shete & Quadro, 2013]) are being discovered and may explain deficiencies in some population groups. In contrast, with expanding, overlapping, or poorly regulated micronutrient interventions, the potential exists for some nutrients (e.g., vitamin A) to be excessively consumed, highlighting the importance of both effective and safe prevention. These many factors emphasize the importance of preventing micronutrient deficiencies in the context of local or regional diets, disease risks, and environs.

The global public health community has historically focused on assessing and preventing vitamin A, zinc, iron, and iodine deficiencies due to their known high prevalence, health consequences, and preventability—for example, through supplementation, food fortification, and dietary diversification. In truth, this narrow focus also reflects lack of data on prevalence, health risks, and resources needed to prevent other micronutrient deficiencies. For this reason, micronutrient deficiencies are colloquially known as "hidden hunger" (Ruel-Burgeron et al., 2015), emphasizing the need for more comprehensive and readily available approaches to assessing micronutrient deficiencies in populations (Cole et al., 2013; Lu et al., 2017).

Vitamin A

Vitamin A deficiency is a nutritional burden of impoverishment with public health consequences, especially among preschool-age children, infants, and women of reproductive age in LMICs (Sommer & West, 1996). Estimating its global burden is challenging, as national surveys are infrequent, leaving estimates imprecise (Stevens et al., 2015; Wirth, Petry, et al., 2017); employ a wide range of status indicators,

making direct comparisons difficult; and are conducted against multiple-intervention landscapes (Wirth, Petry, et al., 2017). A study modeling trends in preschool child vitamin A deficiency (serum retinol < 0.70 μmol/L), based on national surveys from 1991–2013, suggests a prevalence of 29%, with the highest rates in South Asia (44%) and sub-Saharan Africa (48%) (**FIGURE 7-5**) (Stevens et al., 2015). These rates are approximately 10% lower than the global, South Asian, and sub-Saharan African rates of 33%, 50%, and 44%, respectively, obtained by WHO for the period 1995–2005 that led, at the time, to a global estimate of 190 million affected children (WHO, 2009). Thus, while still high, the vitamin A deficiency burden may be gradually lessening, possibly resulting from intensified prophylaxis through fortification, biofortification, and vitamin A supplementation under way in more than 80 countries (Wirth et al., 2017).

Data on vitamin A deficiency among older children, adolescents, and women remain sparse. One estimate suggests that 23% of children 5–15 years of age in Southeast Asia are vitamin A deficient, and 2.6% have mild xerophthalmia (Singh & West, 2004). Among pregnant women in LMICs, 8% (nearly 10 million) are estimated to experience night blindness, while some 19 million may be vitamin A deficient (WHO, 2009). Inexpensive, valid indicators of vitamin A status are needed to greatly improve estimation of this burden (Tanumihardjo et al., 2016).

Function, Requirements, and Assessment. Vitamin A comprises a group of molecules called retinoids, known to regulate numerous cell functions. As an aldehyde (retinal) acting with the protein opsin in photoreceptors of the eye, vitamin A enables light detection and transduction in a process known as the visual cycle (Saari, 2016); as a ligand in the cell nucleus, retinoic acid acts via specific receptors to regulate transcription of hundreds of genes (Blomhoff & Blomhoff, 2006), whose protein products mediate cellular differentiation, proliferation, apoptosis, and plasticity (Al Tanoury, Piskunov, & Rochette-Egly, 2013) plus extracellular matrix composition and signaling functions (Barber, Esteban-Pretel, Marin, & Timoneda, 2014). As retinol, vitamin A regulates enzymatic reactions and cellular bioenergetics (Brossaud, Pallet, & Corcuff, 2017). Through these diverse mechanisms, vitamin A nutriture affects the development, shape, and function of tissues beginning in embryonic life, with deficiency exerting adverse effects on vision, innate and adaptive immune mechanisms, neurotransmission, and bone metabolism and growth.

Vitamin A is provided through the diet as preformed esters in foods such as fish or animal liver, milk, dairy products, eggs, and fortified foods (e.g., vegetable oils, flour, or sugar), and as provitamin A carotenoids, comprising pigments found in yellow-orange vegetables (e.g., yellow sweet potato, carrot) and fruits (ripe mango or papaya) and in dark green leaves (e.g., spinach) (IOM, 2005). Certain indigenous foods may

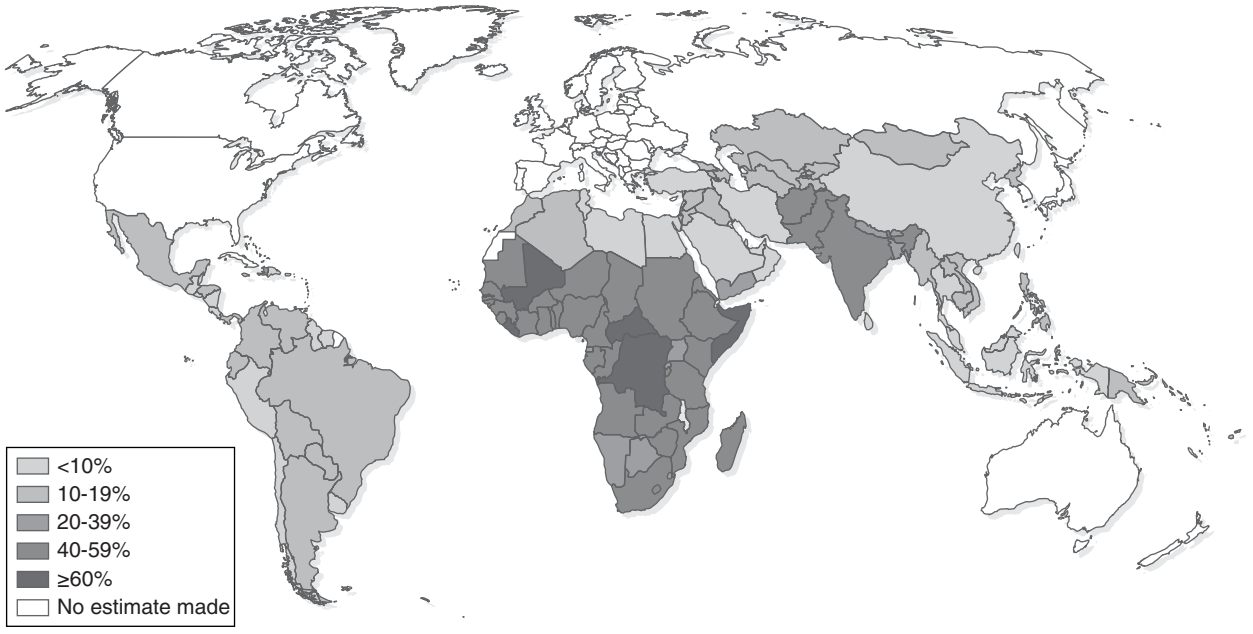


FIGURE 7-5 Estimated prevalence map of vitamin A deficiency among children 6–59 months of age from modeled trends based on serum retinol surveys of 138 countries between 1991 and 2013.

contribute provitamin A to the diet, as illustrated by gac (Aoki, Kieu, Kuze, Tomisaka, & Van Chuyen, 2002), buriti (Mariath, Lima, & Santos, 1989), and pandanus (Engleberger et al., 2006) fruits in Vietnam, Brazil, and the Western Pacific, respectively. Breast milk, which is affected by maternal vitamin A nutriture, provides a critical, protective dietary source of the nutrient to infants through the first 2 years of life (Haskell & Brown, 1999). After 6 months of age, complementary feeding must contribute dietary sources of vitamin A to protect against xerophthalmia and other deficiency disorders (Sommer & West, 1996). As vitamin A is stored in the liver and released into circulation, deficiency typically requires a prolonged period of inadequate intake, during which hepatic stores gradually decrease to low levels; in turn, this leads to low serum (or plasma) retinol concentrations, altered tissue function, and increased risk of impaired health (Tanumihardjo et al., 2016).

Assuring dietary adequacy of vitamin A in children and women of reproductive age is a public health priority, evaluated by comparing intakes, as retinol activity equivalents (RAEs) that combine contributions from preformed and provitamin A food sources (assuming a beta-carotene-to-retinol conversion of 12:1), to age-sex specific Estimated Average Requirements (EARs) and Recommended Dietary Allowances (RDAs) (IOM, 2005) (**TABLE 7-5**). These indices specify intakes at which 50% and 97.5%, (i.e., +2 SD above the EAR), respectively, of a healthy population can meet requirements for vitamin A. When an EAR (and, therefore, RDA) is not computable for an age group, such as infancy, an Adequate Intake (AI) has been derived based on observed intakes in healthy

TABLE 7-5 Dietary Reference Intakes: Estimated Average Requirement and Recommended Dietary Allowances for Vitamin A by Life Stage Group

		Dietary Reference Intakes (IOM, 2006)	
	Sex	EAR[a]	RDA[a]
Life Stage Group		RAE/d[b]	
≤ 6 months	Both male and female		400 (AI)[a]
7–12 months	Both male and female		500 (AI)[a]
1–3 years	Both male and female	210	300
4–8 years	Both male and female	275	400
9–13 years	Male/female	445/420	600/600
14–18 years	Male/female	630/485	900/700
19–70 years or older	Male/female	625/500	900/700
Pregnancy			
≤ 18 years	Female	530	750
19–50 years	Female	550	770
Lactation	Female	900	1,300

[a] Components of the Dietary Reference Intakes: EAR = Estimated Average Requirement; RDA = Recommended Dietary Allowance; AI = Adequate Intake (IOM, 2006).
[b] Retinol Activity Equivalent, where the beta-carotene-to-retinol conversion ratio is assumed to be 12:1.
Data from IOM. (2006). *Dietary reference intakes: Essential guide to nutrient requirements* (Vitamin A Chapter, Table 1, pp. 170). J. J. Otten, J. P. Hellwig, & L. D. Meyers (Eds.). Washington, DC: National Academies Press. Copyright 2006 by the National Academy of Sciences.

populations. The RDA for vitamin A is 300 to 600 μg RAE/day for children up to 13 years of age, and 700 to 900 μg RAE/day thereafter for both sexes, including for women during pregnancy, but increases to 1,300 μg/day during lactation. Intakes that are chronically below the EAR by age and sex can be expected to place populations at risk of vitamin A deficiency.

Vitamin A status is commonly assessed by clinical, biochemical, and functional indicators. Serum (plasma) retinol concentration remains the most commonly measured biochemical indicator, with a concentration less than 0.70 μmol/L being indicative of deficiency (Tanumihardjo et al., 2016) and prevalences of 10% or greater and 20% or greater reflecting WHO cut-offs marking moderate public health concerns and severe public health concerns, respectively (WHO, 2011a).

Although usually maintained within a normal range when adequately nourished, serum retinol falls progressively with depletion of body stores. It also declines in response to inflammation—for example, with infection. This latter response, which is attributed largely to retinol binding protein (RBP; the major transport protein) being sequestered in the liver, suggests that serum retinol may be stratified (or adjusted) by biomarkers of inflammation, such as C-reactive protein and α_1-acid glycoprotein, in affected populations. Because these biomarkers respond to severity and duration of inflammation, their use allows fractions of low serum retinol to be potentially explained by stages of infection versus nutritional deficiency (Thurnham, Mburu, Mwaniki, & De Wagt, 2005).

Although serum retinol remains the most common population indicator of vitamin A status, other approaches include measurements of serum RBP, breast milk and dried blood spot retinol concentrations, relative response tests following administration of a small oral dose of vitamin A (approximating liver store relative adequacy), stable isotope dilution (to estimate total body vitamin A stores), and functional tests of dark adaptation (Tanumihardjo et al., 2016). Each of these approaches has its own strengths, limitations, and costs that guide its use.

Health Consequences

Xerophthalmia. Ocular manifestations of vitamin A deficiency include night blindness and Bitot's spots (**FIGURE 7-6A**), each of which may affect 1% to 5% of preschoolers in a high-risk population. A far less common manifestation is the potentially blinding keratomalacia ("softening of the cornea"; **FIGURE 7-6B**).

Night blindness (designated XN), the mildest symptom, results from a depletion of vitamin A in the rod photoreceptors of the eye and inability to see under low-light conditions (Sommer, 1995). Night blindness is often detectable by history, using a local term that may refer to *twilight or evening blindness* or *chicken eyes* (chickens' eyes lack rods, so these animals are night blind). Pregnant women may become night blind in the latter half of pregnancy due to gestational demands in the context of an inadequate diet (Christian, 2002). XN usually disappears in both children and mothers within 24–48 hours following vitamin A treatment per WHO guidelines (Sommer, 1995).

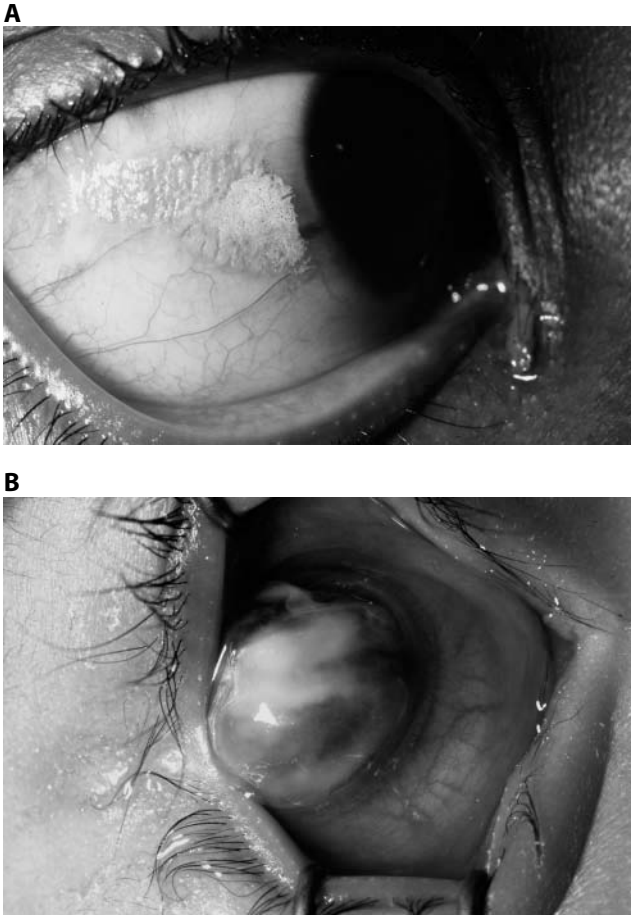


FIGURE 7-6 A. Bitot's spot. **B.** Keratomalacia.

Reprinted with permission of Oxford University Press, Inc. Photos courtesy of Randy Whitfield and Alfred Sommer.

Bitot's spots (X1B), the next most severe clinical stage, appear on the conjunctiva as cheesy or foamy patches of keratinized, sloughed cells. X1B does not affect vision and regresses following vitamin A treatment over a period of weeks, although it may not always disappear. While representing mild ocular disease, XN and X1B reflect moderate-to-severe systemic vitamin A deficiency, evident by their positive association with morbidity and mortality in young children (Sommer, Hussaini, Tarwotjo, & Susanto, 1983) and (for XN) pregnant women (Christian et al., 2000).

Keratomalacia (X3) represents a blinding stage of xerophthalmia. It is associated with a high risk of fatality, with death usually precipitated by measles or other severe infections. Vitamin A therapy will heal the cornea in patients with X3, but may not preserve vision, depending on the extent and severity of corneal destruction (Sommer & West, 1996).

Morbidity. Vitamin A supplementation has been shown to reduce severity of infections, including diarrhea or dysentery (Stevens et al., 2015), measles (Fawzi, Chalmers, Herrera, & Mosteller, 1993), malaria (Shankar et al., 1999), and other infectious illnesses. In addition, it can reduce hearing loss associated with purulent ear discharge (Schmitz et al., 2012), an effect likely explained by effects of the vitamin in maintaining normal epithelial barrier (Sommer & West, 1996) and by innate and adaptive immune function (Beijer, Kraal, & den Haan, 2013), including tolerance and the capacity to resolve infections (Mora & Iwata, 2015). While animal studies reveal vitamin A deficiency may also compromise lung defenses and, in human studies, be associated with an increased risk of respiratory infection (Sommer & West, 1996; Thornton, Mora-Plazas, Marin, & Villamor, 2014), vitamin A supplementation does not appear to reduce risk of lower respiratory infections (Nacul et al., 1998; Vitamin A and Pneumonia Working Group, 1995), for reasons that remain unclear.

Mortality. Vitamin A deficiency increases the risk of mortality among preschool-age children, including infants, in underserved and undernourished regions. Large community trials in South Asia and Africa, designed and implemented to test the efficacy of supplementation, have shown that vitamin A, delivered as a high oral dose every 4–6 months (200,000 IU for children 12 months or older, half-dose at age 6–11 months), or as a weekly small dose (15,000 IU), or provided as a fraction of an RDA daily via food fortification, can reduce preschool child mortality by 23% to 34% (Beaton et al., 1993; Fawzi et al., 1993; Sommer & West, 1996) (**FIGURE 7-7**). Embedded in this decline in all-cause mortality are likely strong effects of vitamin A in reducing the case fatality of measles, which has been documented in both community (Rahmathullah et al., 1990; West et al., 1991) and hospital-based trials (Barclay, Foster, & Sommer, 1987; Coutsoudis, Broughton, & Coovadia, 1992; Hussey & Klein, 1990), and in decreasing the severity of diarrhea and dysentery (Arthur et al., 1992; Banajeh, 2003; Stevens et al., 2015) and *Plasmodium falciparum* malaria (Shankar et al., 1999). Evaluations of vitamin A programs may yield estimates of impact that are higher, if accompanied by other interventions that enhance child survival (Thapa, Choe, & Retherford, 2005). Conversely, such estimates may be lower if design, implementation, and resource constraints affect the quality of either the program or its evaluation (Awasthi et al., 2013; West, Sommer, Palmer, Schultink, & Habicht, 2015).

Early neonatal vitamin A supplementation (oral 50,000 IU dose) has been evaluated for efficacy in reducing infant mortality. An early, positive trial from Indonesia (Humphrey et al., 1996) was followed by 10 additional trials across the Southeast Asian, Eastern Mediterranean, and African regions to assess the effect on mortality in the first 6 months of life (**TABLE 7-6**). While no effects were observed in Africa (Malaba

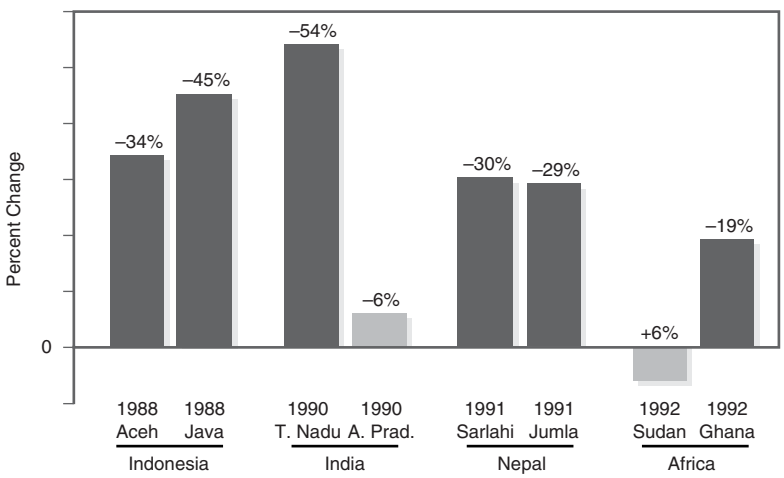


FIGURE 7-7 Summary of percent change in all-cause mortality of children (approximately 6 months or older to 72 months) receiving vitamin A versus no vitamin A (controls) in eight field intervention trials. Black bars represent statistically significant reductions; white bars indicate nonsignificant change. Total *N* (across trials) >165,000 children. Meta-analyses showed overall 23% to 34% reductions in mortality with vitamin A.

TABLE 7-6 Summary of Randomized Trials Evaluating Impact of Early Neonatal Vitamin A Supplementation on Infant Mortality

WHO Region Country First author, year	Vitamin A Dosage (IU)	Follow-up Period (months)	Mortality Rate (per 1,000 Live Births)						Relative Risk (95% CI)
			Vitamin A			Placebo			
			n Deaths	*N*	Rate	*n* Deaths	*N*	Rate	
Southeast Asia									
Indonesia Humphrey et al., 1996	50,000	12	7	970	7.2[1]	19	957	19.8[1]	0.36 (0.16–0.87)
Bangladesh Klemm et al., 2008	50,000	6	306	7,953	38.5	360	7,984	45.1	0.85 (0.73–1.00)
India Rahmathullah et al., 2003	48,000	6	146	2,713	53.8[1]	188	2,719	69.1[1]	0.78 (0.63–0.96)
India Mazumder et al., 2015	50,000	6	656	22,493	29.2	726	22,491	32.3	0.90 (0.81–1.00)
Eastern Mediterranean									
Pakistan Soofi et al., 2017	50,000	6	128	5,648	23	115	5,380	21	1.06[2] (0.82–1.37)
Africa									
Zimbabwe Malaba et al., 2005	50,000	12	88	4,195	21[1]	82	4,239	19.3[1]	1.08 (0.80–1.46)
Guinea Bissau Benn et al., 2008	50,000	12	88	1,795	49.0[1]	86	1,884	45.6[1]	1.07 (0.79–1.44)
Guinea Bissau Benn et al., 2010	25,000	12	83	757	110[1]	78	762	102[1]	1.08 (0.79–1.47)
Guinea Bissau Benn et al., 2014	25,000 or 50,000	12	115	2,744	42[1]	45	1,377	33[1]	1.28 (0.91–1.81)
Ghana Edmond et al., 2014	50,000	6	278	11,345	24.5	248	11,353	21.8	1.12 (0.95–1.33)
Tanzania Masanja et al., 2014	50,000	6	407	15,428	26.4	372	15,464	24.1	1.10 (0.95–1.26)

[1] Mortality rate ratios reported per 1,000 infant-years

[2] Reported value is an odds ratio

et al., 2005; Benn et al., 2008, 2010, 2014; Edmond et al., 2014; Masanja et al., 2014), or Eastern Mediterranean (Soofi et al., 2017) regions, 10% to 64% reductions in all-cause mortality (relative risks of 0.90 to 0.36) were reported from trials in Southeast Asia (Humphrey et al., 1996; Klemm et al., 2008; Rahmathullah et al., 2003; Mazumder et al., 2014). The stark regional differences may be partly explained by more pronounced gestational vitamin A deficiency in Southeast Asia (Rotondi & Khobzi, 2010). In South India, where morbidity was monitored, fatality related to diarrhea and fever was reduced 45% and 37%, respectively (Tielsch et al., 2007). Animal vitamin A repletion experiments in the puerperium support strengthened intestinal mucosal immune responses (Liu et al., 2014) and less colitic inflammation (Reifen et al., 2002) in vitamin A–replete versus vitamin A–deficient offspring as plausible mechanisms of action. Newborn vitamin A receipt was also associated with delayed pneumococcal colonization of the respiratory tract at 2–4 months of age in India (Coles et al., 2001), suggesting an additional possible pathway of effect.

Effects of antenatal vitamin A supplementation on pregnancy outcome and maternal mortality have been examined in South Asia and West Africa. In Nepal, where maternal night blindness was associated with increased morbidity, anemia (Christian et al., 1998), and mortality through 2 years postpartum (Christian et al., 2000), weekly, maternal low-dose vitamin A or beta-carotene supplementation reduced pregnancy-related mortality by approximately 40% (West et al., 1999), but had no overall effect on infant mortality (Katz et al., 2000). However, offspring of vitamin A–supplemented mothers followed to pubescence exhibited greater lung (Checkley et al., 2010) and innate immune (Palmer, Schulze, Khatry, De Luca, & West, 2015) capacity. In rural Bangladesh, a similar weekly vitamin A or beta-carotene regimen revealed no effect on materno-infant mortality (West et al., 2011), although risk of maternal bacterial vaginosis was reduced (Christian et al., 2011). In Ghana, where maternal night blindness is rare, weekly vitamin A during pregnancy also had no effect on mortality (Kirkwood et al., 2010), suggesting that levels of maternal vitamin A deficiency may influence health outcomes of this intervention.

Poor Growth and Anemia. Vitamin A is required for mammalian growth. Children with mild xerophthalmia (XN or X1B) are more likely to be stunted than children who do not have xerophthalmia (Sommer & West, 1996). Nevertheless, growth responses to population-based vitamin A supplementation have revealed mixed results, with most studies finding either no effect (Ramakrishnan, Aburto, McCabe, & Martorell, 2004) or seasonal, subgroup, sex-specific, or anthropometry-specific effects (Sommer & West, 1996).

Vitamin A regulates erythropoiesis, modulates iron metabolism and mobilization, and influences immune functions, providing plausible pathways by which vitamin A deficiency could lead to an increased risk of anemia (Christian et al., 1998; Semba & Bloem, 2002). Where deficiency is endemic, there is moderate evidence from meta-analyses that vitamin A supplementation during pregnancy may reduce maternal anemia by one-third (odds ratio [OR], 0.64; 95% CI, 0.43–0.94) (McCauley, Van Den Broek, Dou, & Othman, 2015).

Iron Deficiency and Anemia

Iron is essential in the body for oxygen transport and cellular respiration—functions that are especially critical in red blood cells, brain, and muscle (Beard, 2001). Iron deficiency is considered the most common micronutrient deficiency in the world; anemia, characterized by abnormally low blood hemoglobin concentration, is its major clinical manifestation. Although anemia is not specific to iron deficiency, the two conditions are closely linked in most malnourished populations, making anemia the most frequently reported clinical index of iron deficiency in low-income countries (WHO, 2007).

Anemia has both nutritional and non-nutritional causes. Major non-nutritional causes of red cell mass loss or destruction and consequent anemia include hookworm, schistosomiasis, trichuris, malaria, or HIV infections (Thurnham & Northrop-Clewes, 2007) as well as inherited hemoglobin disorders such as sickle cell or thalassemia (Engle-Stone et al., 2017). The term *nutritional anemia* goes beyond iron deficiency and includes the anemia due to deficiencies in other nutrients, particularly vitamin A, folate, riboflavin, and vitamin B_{12} (Fishman, Christian, & West, 2000), and trace elements that participate in erythropoiesis (Olivares, Hertrampf, & Uauy, 2007; Semba, 2007). Approximately half of the global burden of anemia appears to be responsive to iron supplementation (Gera, Sachdev, Nestel, & Sachdev, 2007). At the same time, iron deficiency may occur without the development of anemia (Engle-Stone et al., 2017; Wirth et al., 2017). The relative degree to which the distributions of anemia and iron deficiency overlap varies by the underlying burden of infection and by the prevalence of nutritional deficiencies. In many low-income populations, where the burden of infectious disease is high and micronutrient deficiencies are common, it is difficult to determine which of these factors may be the primary cause of anemia; in turn, multiple strategies might be needed to address these conditions.

An estimated 800 million children and women of reproductive age are anemic and 153 countries are classified as having anemia prevalence of moderate to severe public health significance (WHO, 2015a). In the period between 1995 and 2011, there was a modest reduction in the prevalence of anemia worldwide; however, due to population growth, the absolute numbers of women and children affected have increased (Stevens et al., 2013). Pregnant women, infants, and young children are at highest risk of anemia. On average, an even higher percentage of preschool-age children are anemic than pregnant women or other women of reproductive age (**TABLE 7-7**). The condition affects 42% of children, or an estimated 273 million worldwide. Africa and Southeast Asia have the highest burden of anemia, and presumably iron deficiency: An estimated 54% to 62% of young children and 46% to 48% of pregnant women in those regions have anemia. However, the largest number of people with anemia live in Southeast Asia, where roughly 97 million children, 11 million pregnant women, and 191 million nonpregnant women are estimated to suffer from anemia.

Function, Requirements, and Assessment. In the body, iron is found in metabolically active "functional" and "storage" pools, accounting for approximately 75% and 25% of total body iron, respectively (Lönnerdal & Hernell, 2013). Approximately 80% of functional iron complexes with hemoglobin during erythropoiesis, where it plays a central role in oxygen transport to cells. Ten percent of this metabolic iron pool is incorporated into intracellular myoglobin, where oxygen is stored for use during muscle respiration and contraction. In addition, iron serves as a cofactor in more than 200 heme and nonheme enzymes involved in cellular respiration, division, neurotransmission, immunity, and growth (Beard, 2001). It plays a catalytic role in numerous metabolic oxidation–reduction (redox) reactions due to the ability of ferrous (Fe^{2+}) and ferric (Fe^{3+}) iron to donate and accept electrons, respectively. However, excess iron can promote the formation of reactive oxygen species that can cause damage to DNA, proteins, or lipids within cells (Coffey & Ganz, 2017).

During pregnancy, there is an additional iron requirement for maternal red blood cell expansion and fetal growth and development (Cao & O'Brien, 2013). During the fetal and infant periods, iron is essential for the developing brain through its role in neurogenesis and differentiation, neurotransmitter synthesis, and energy metabolism in the brain (Beard, 2008; Lozoff et al., 2006). Iron also plays an important role in the human innate immune system, such that deficiency of this mineral may result in thymic atrophy, a depression

TABLE 7-7 Prevalence of Anemia by Life Stage[a] and Geographic Region of the World, 2011

	Preschool-Age Children		Pregnant Women		Nonpregnant Women	
WHO Region	Prevalence (%)	Number Affected (millions)	Prevalence (%)	Number Affected (millions)	Prevalence (%)	Number Affected (millions)
Africa	62.3	84.5	46.3	9.2	37.8	69.9
Americas	22.3	17.1	24.9	2.4	16.5	38.1
Southeast Asia	53.8	96.7	48.7	11.5	41.5	190.6
Europe	22.9	12.7	25.8	1.8	22.5	48.4
Eastern Mediterranean	48.6	35.8	38.9	3.9	37.7	55.2
Western Pacific	21.9	25.7	24.3	3.6	19.8	92.6
Global	42.6	273.2	38.2	32.4	29.0	496.3

[a] Preschool-aged children (0–5 years), pregnant women (15–49 years), and nonpregnant women (15–49 years).

Reprinted from World Health Organization (WHO). (2015). *The global prevalence of anaemia in 2011*. Retrieved from http://apps.who.int/iris/bitstream/10665/177094/1/9789241564960_eng.pdf

of T-lymphocyte production and decreased neutrophil function (Jonker, Te Poel, Bates, & Boele van Hensbroek, 2017). These ubiquitous functions highlight the importance of adequate iron nutriture in achieving health benchmarks as diverse as normal physical performance, pregnancy, and motor and cognitive function.

The metabolic and functional pathways of iron in the body provide the basis for the indicators used to assess iron status. The total body iron balance is largely regulated at the point of absorption because humans lack a mechanism for controlled iron excretion (Coffey & Ganz, 2017). Human diets contain iron in two forms: (1) heme iron, derived mainly from hemoglobin and myoglobin found in meat, poultry, and seafood; and (2) various forms of inorganic non-heme iron, typically from plant sources. A higher proportion of heme iron is absorbed, making this the more bioavailable source of iron, yet non-heme iron is the predominant form of iron in most diets. When iron status is adequate, the hormone hepcidin will block the release of iron from enterocytes in the small intestine, causing excess iron to be lost when enterocytes are sloughed off into the intestinal lumen (Coffey & Ganz, 2017). During periods of iron deficiency, hepcidin expression is reduced, thereby allowing a greater proportion of iron to be transported into the plasma.

Newly absorbed iron as well as endogenous iron released from normal degradation of senescent red blood cells is transported to tissues by plasma transferrin. Although almost all cells require iron, the majority of transferrin-bound iron is delivered to bone marrow for red blood cell production—a process mediated by expression of a specific transferrin receptor that reflects tissue iron need (Coffey & Ganz, 2017; Knutson, 2017). Transferrin will also serve to deliver iron to ferritin, the major body iron storage protein, found primarily in the liver (Gibson, 2005).

During prolonged dietary deficit, iron is released from intracellular ferritin into the circulation via transferrin to support hematopoiesis, which has the effect of decreasing iron stores. As iron depletion progresses, transferrin carries and delivers a diminishing supply of iron to the marrow and other tissues. This stage of iron deficiency without anemia is reflected by increased transferrin receptor expression on cell surfaces, a low level of transferrin being saturated with iron, and increased amounts of circulating erythrocyte protoporphyrin, a protein that accumulates in red blood cells lacking iron. If the iron deficit continues, iron-deficiency anemia develops as hemoglobin concentration falls and red cells become smaller in size (microcytic) and more pale in color (hypochromic) (Gibson, 2005).

Several indicators may be used to track these changes in iron status, only a few of which are commonly used. Assessment of hemoglobin or hematocrit concentrations and evaluation of their distributions against conventional cut-offs by age, life stage, and gender (**TABLE 7-8**) is the standard approach for diagnosing and estimating the prevalence of anemia (WHO, 2011b). Hemoglobin concentrations are affected by altitude and smoking status, so adjustments should be made to account for those factors (Gibson, 2005). Simple diagnostic tools available to assess anemia in primary care settings include the clinical diagnosis of pallor (palmar, tarsal conjunctival, and nail bed), which can be used to identify severe anemia (hemoglobin less than 50 to 70 g/L or hematocrit less than 15%) with approximately 10% to 50% sensitivity and 90% to 100% specificity (Kalter et al., 1997; Luby et al., 1995). If facilities allow, assessment of additional hematologic measures is recommended, such as the mean cell volume (MCV), mean cell hemoglobin (MCH), and mean cell hemoglobin concentration (MCHC), all of which will be low in the presence of iron-deficiency anemia (Gibson, 2005). Assessing hemoglobin response to supplementation provides another, more accurate, but complex approach to estimating the extent of iron-deficiency anemia in populations (Gera et al., 2007).

TABLE 7-8 Commonly Used Indicators of Anemia in Various Life Stage Groups

	Anemia	
	Hemoglobin[a] (g/dL)	Hematocrit (%)
Children 0–5 years	<11.0	<33.0
Children 5–11 years	<11.5	<34.0
Nonpregnant women	<12.0	<36.0
Pregnant women		
First trimester	<11.0	<33.0
Second trimester	<10.5	<32.0
Third trimester	<11.0	<33.0
Men	<13.0	<39.0

[a] Hemoglobin cut-offs are for individuals residing at sea level.
Data from Cao & O'Brien, 2013; Gibson, 2005; WHO, 2011b.

Serum ferritin and transferrin receptors are more specific indicators of iron deficiency. Although the majority of ferritin is intracellular, serum ferritin concentrations are reflective of iron stores and will decline as the individual becomes deficient in iron. Concentrations less than 12 μg/L are indicative of iron deficiency. In contrast, serum transferrin receptors will be expressed in greater concentrations as cells signal the need for more iron, with concentrations greater than 8.3 mg/L being indicative of iron-deficient erythropoiesis (Rohner et al., 2017). Serum ferritin and, to a lesser degree, transferrin receptors are both sensitive to inflammation, making interpretation of their values in the presence of infection challenging. Various methods have been proposed to adjust ferritin and transferrin receptor estimates by measuring two inflammatory cytokines, C-reactive protein (CRP) and alpha-1 acid glycoprotein (AGP) (Namaste et al., 2017; Rohner et al., 2017; Thurnham et al., 2010). New multiplex, rapid immunoassays are being developed that include concurrent measurements of ferritin, CRP, and AGP, along with other micronutrients (Lu et al., 2017).

Other indicators of iron status that are useful, albeit less commonly used in resource-limited settings, include the measurement of zinc protoporphyrin, serum iron, total iron-binding capacity, and transferrin saturation. The latter three are useful for differentiating between iron-deficiency anemia and anemia due to chronic infections, inflammation, or neoplastic diseases (Gibson, 2005). For individual assessment and when resources allow, it is recommended that multiple iron status and red blood cell indices be measured to more specifically determine the likely cause of anemia or degree of iron deficiency (Gibson, 2005).

Health Consequences. The highest-risk groups, in terms of both probability of becoming anemic from iron deficiency and suffering its consequences, are women of reproductive age, especially during pregnancy; infants; and young children.

Pregnancy-Related Outcomes. Women are at high risk of anemia due to periodic menstrual blood loss and because of increased requirements during pregnancy due to expanded red blood cell mass and accretion of iron in fetal tissue and the placenta. In a multicountry study of severe adverse pregnancy outcomes, the most common risk factor for maternal death was severe anemia, which was present in 50% of cases (Lumbiganon et al., 2014). While no experimental trials have been undertaken to demonstrate that iron supplementation reduces maternal mortality, plausible causes associated with moderate to severe anemia include puerperal cardiac failure or hemorrhage (Allen, 1997). A meta-analysis suggests there is a continuous reduction in maternal mortality risk, reflected by a protective odds ratio of 0.75 (95% CI, 0.62–0.89), with each 1 g/dL increase in blood hemoglobin concentration (Stoltzfus, Mullany, & Black, 2004).

Risks of preterm delivery, low birth weight, fetal malformations, and fetal deaths have been found to follow a U-shaped curve with respect to maternal hemoglobin levels measured early in pregnancy, with an elevated hemoglobin level presenting as much or more risk of an adverse outcome as a lower hemoglobin concentration (Cao & O'Brien, 2013). Inadequate plasma volume expansion may explain some of the excess risk among those pregnant women with high hemoglobin concentrations, but the causal mechanisms yielding these risks are not well understood.

Maternal iron-deficiency anemia may also place newborns at risk of low iron stores during infancy (Cao & O'Brien, 2013). Iron supplementation during pregnancy has been found to reduce the risk of anemia at term and increase maternal hemoglobin levels at 6 weeks postpartum, although the protective effects on low birth weight and preterm delivery were somewhat weak and not statistically significant (LBW: relative risk [RR], 0.84; 95% CI, 0.69–1.03; preterm birth: RR, 0.93; 95% CI, 0.84–1.03) (Pena-Rosas, De-Regil, Garcia-Casal, & Dowswell, 2015).

Child Health Outcomes. Late infancy and early childhood is a high-risk period for iron deficiency and anemia because of the high iron supplies needed to support rapid growth coupled with low dietary intake of often poorly bioavailable iron (Dewey, 2013). Although some studies have noted an association between iron-deficiency anemia and stunted growth (Chwang, Soemantri, & Pollitt, 1988), meta-analyses of iron supplementation trials have failed to find an overall effect on child growth (Pasricha, Hayes, Kalumba, & Biggs, 2013; Thompson, Biggs, & Pasricha, 2013).

A meta-analysis of epidemiologic studies in Africa revealed that for each 1 g/dL increase in hemoglobin, there was an associated 24% reduction in the odds of child mortality (OR, 0.76; 95% CI, 0.62–0.93) (Scott, Chen-Edinboro, Caulfield, & Murray-Kolb, 2014). There is a complex relationship between iron status and morbidity, however. As described earlier, iron is required for many pathways in cellular immunity. Yet, iron is also required by most pathogenic bacterial, viral, and parasitic organisms, which sets the stage for a competition between the host and the pathogen for access to iron (Drakesmith & Prentice, 2012). For this reason, there may be some benefit for the host to create a state of functional iron deficiency, by reducing iron absorption and sequestering iron

from circulation during periods of infection (Jonker et al., 2017). Iron supplementation during a period of acute infection, such as with malaria, may provide limited benefit because the individual cannot absorb iron efficiently during that time, so the intervention may actually increase the risk and severity of infection (Jaramillo et al., 2017; Sazawal et al., 2006).

Iron-deficiency anemia is also associated with poorer cognitive outcomes and school performance in children (Prado & Dewey, 2014), with associated long-term consequences on economic opportunity (Horton & Ross, 2007). Iron is essential for neurogenesis and differentiation of brain cells and regions. Deficiency likely disturbs the developing brain via at least three iron-dependent pathways: (1) disruption of oligodendrocytes, the cells responsible for producing myelin; (2) disruption of neurotransmitter synthesis and metabolism; and (3) regulation of energy metabolism in the brain (Beard, 2008). Longitudinal studies of children who had been anemic prior to age 2 years have consistently demonstrated long-term deficits in cognition and school achievement from 4 to 19 years of age (Lozoff et al., 2006), although observational studies may be confounded other risk factors for poor child development common among children living in poverty (Black et al., 2017; Prado & Dewey, 2014). A number of randomized controlled trials of iron supplementation for infants or children have examined developmental outcomes. There is evidence of weak, but positive effects on motor development, cognitive/language development, and socio-emotional development, although the few long-term follow-up studies that have been conducted have not demonstrated persistent effects into later childhood (Pasricha et al., 2013; Prado & Dewey, 2014; Thompson et al., 2013).

Iodine Deficiency

Iodine is an essential component of thyroid hormones that control cellular metabolism and neuromuscular tissue growth and development. Deficiency in iodine and consequent thyroid hormone production during critical periods of organogenesis can damage the brain and nervous tissue, causing irreversible mental retardation and other developmental abnormalities. The spectrum of mild through severe health consequences causally linked to iodine deficiency at different stages of life is collectively known as iodine-deficiency disorders (IDD) (Delange, 1994; Li & Eastman, 2012), a concept that has led to greater understanding of the multiple health and societal consequences of this nutrient deficiency. Informative treatises exist on the histories of goiter and cretinism—the two most notable clinical syndromes of iodine deficiency (Hetzel, 1989).

Iodine adequacy of plant and animal sources of food grown in a general locale depends on adequacy of the nutrient in soil, which can vary by region of the world. Most iodine is found in the world's oceans, such that the highest soil concentrations have historically been found in coastal areas (Eastman & Zimmermann, 2017). More recently, recent epidemiologic studies of IDD have revealed the problem to extend beyond inland and mountainous regions to include island nations (Li & Eastman, 2012). In some cultures, insufficient iodine intake is augmented by routine consumption of goitrogenic substances, such as linamarin, a cyanide-containing compound found in the root of cassava (Delange, 1994; Rohner et al., 2014). Thiocyanates, which result from detoxification of linamarin in the liver, decrease iodine uptake by the thyroid gland and suppress circulating thyroid hormone, leading to secondary iodine deficiency.

As of the 1990s, an estimated 1.6 billion persons, or nearly 30% of the world's population, were thought to be at risk of iodine deficiency (WHO, 1993), based on documented prevalence and sizes of affected populations living in iodine-depleted regions. Despite successes in improving iodine status via salt iodization, iodine deficiency remains a significant public health problem in 32 countries (Andersson, Karumbunathan, & Zimmermann, 2012). Revised estimates of the extent of iodine deficiency obtained from WHO's global database for the 193 member-states of the United Nations suggest that nearly 2 billion people, including 30% of school-age children, have iodine intakes below the required amounts (**TABLE 7-9**), including 76 million children in Southeast Asia and nearly 58 million children in Africa. However, the highest proportions of children with inadequate iodine intake live in the European (43%) and Eastern Mediterranean (39%) regions. In contrast, iodine intakes may be excessive in 47 countries, emphasizing the importance of regular monitoring of national IDD prevention initiatives.

Function, Requirements, and Assessment. Iodine, ingested as either iodide or iodate, is an essential constituent of thyroid hormones. Once in circulation, approximately 80% of total body iodine is stored in the thyroid gland, where it is oxidized, bound to tyrosine amino acids, and catalyzed by an iron-containing peroxidase to form the thyroid hormones, T_3 (triiodothyronine), and T_4 (tetraiodothyronine, or thyroxine) (Li & Eastman, 2012). T_3 and T_4 are stored in the thyroid gland in association with the glycoprotein thyroglobulin. During periods of low dietary intake, thyroid-stimulating hormone (TSH) is secreted from the anterior pituitary gland, which leads to reduced renal clearance of iodine and increased breakdown of

TABLE 7-9 Number of Countries, Proportion of Population, and Number of Individuals with Insufficient Iodine Intake in School-Age Children (6–12 Years) and in the General Population (All Age Groups), by WHO Region in 2011

		Insufficient Iodine Intake (UI < 100 µg/L)			
		School-Age Children		General Population	
WHO Region[a]	Countries (number)	Proportion (%)	Total Number (millions)[b]	Proportion (%)	Total Number (millions)[b]
Africa	10	39.3	57.9	40.0	321.1
Americas	2	13.7	14.6	13.7	125.7
Southeast Asia	0	31.8	76.0	31.6	541.3
Europe	11	43.9	30.5	44.2	393.3
Eastern Mediterranean	4	38.6	30.7	37.4	199.2
Western Pacific	5	18.6	31.2	17.3	300.8
Total	32	29.8	240.9	28.5	1881.2

[a] Comprising 193 WHO member-states.
[b] Based on population estimates in the year 2010.
Reproduced from Andersson, M., Karumbunathan, V., & Zimmermann, M. B. (2012). Global iodine status in 2011 and trends over the past decade. *Journal of Nutrition*, *142*(4), 744–750, by permission of Oxford University Press.

thyroglobulin (Li & Eastman, 2012). In this process, T_3 hormones are secreted and T_4 is preferentially converted to T_3 in peripheral tissues. As iodine concentrations become progressively depleted, hypothyroidism develops and hyperplasia of the thyroid epithelial cells may lead to the development of goiter.

During gestation, both maternal- and fetal-derived T_3 and T_4 contribute to the regulatory thyroxine pool that, in concert with other hormones (e.g., growth hormone, insulin), directs fetal tissue function, growth, and development. In particular, thyroid hormones are required for normal neuronal migration and myelination of the brain during fetal and early postnatal life and play a role in regulating fetal and infant growth (Eastman & Zimmermann, 2017; Rohner et al., 2014). Paradoxically, excess buildup of iodine in the thyroid gland can suppress the release of T_3 and T_4 and cause goiter, particularly if there is an overcorrection for prior deficiency (Eastman & Zimmermann, 2017). High iodine intake after prolonged deficiency may also lead to iodine-induced hyperthyroidism.

Iodine status can be assessed through clinical and biochemical means. Urinary iodine concentration, goiter rate, serum TSH, and serum thyroglobulin are four methods generally recommended for assessment (Rohner et al., 2014). They provide a complementary picture of iodine status, with urinary iodine representing a sensitive indicator of recent iodine intake, thyroglobulin responding to a change in status over a period of weeks to months, and the goiter rate usually reflective of longer-term iodine nutrition over a period of months to years (Zimmermann & Andersson, 2012). Indicators with suggested cut-offs, target populations for assessment, and criteria related to severity of iodine deficiency as a public health problem have been published by WHO, UNICEF, and the International Council for the Control of Iodine Deficiency Disorders (ICCIDD) (2007) (**TABLE 7-10**).

Virtually all goiter occurring in iodine-deficient areas can be attributed to iodine deficiency; thus, goiter prevalence can serve as a useful population indicator of risk. The size of the thyroid gland changes inversely in response to change in iodine intake. Nevertheless, there are practical limitations for assessing goiter where iodine deficiency is mild or in very young age groups because of the challenges in palpating the smaller thyroid gland. High-resolution ultrasonography is a feasible option for assessing goiter, but has limited use due to its expense and training requirements (Zimmermann & Andersson, 2012).

Urinary iodine (UI) excretion over 24 hours and measurement of UI concentration, alone or preferably

TABLE 7-10 Epidemiologic Criteria for Assessing the Severity of Iodine-Deficiency Disorders Based on Urinary Iodine or the Total Goiter Rate in School-Age Children and Pregnant Women

	Urinary Iodine		Total Goiter Rate	
Population Group	**Median Concentration (μg/L)**	**Interpretation**	**Proportion (%)**	**Interpretation**
School-age children	<20	Insufficient: severe iodine deficiency	>30	Severe
	20–49	Insufficient: moderate iodine deficiency	20.0–29.9	Moderate
	50–99	Insufficient: mild iodine deficiency	5.0–19.9	Mild
	100–199	Adequate	0.0–4.9	None
	200–299	Above requirements	5.0–19.9	Mild
	≥300	Excessive		
Pregnant women	<150	Insufficient		
	150–249	Adequate		
	250–599	Above requirements		
	≥500	Excessive		

Reprinted from World Health Organization (WHO), UNICEF, & International Council for Control of Iodine Deficiency Disorders (ICCIDD). (2007). *Assessment of iodine deficiency disorders and monitoring their elimination: A guide for program managers,* 3rd ed. Retrieved from http://apps.who.int/iris/bitstream/10665/43781/1/9789241595827_eng.pdf

in relation to creatinine excretion, serve as conventional, biochemical means of determining the current iodine intake and status of a population. However, because collecting 24-hour samples can be impractical (Zimmermann & Andersson, 2012) and measuring creatinine is challenging, UI concentration (in μg/L) is commonly reported. With a large number of samples, day-to-day variations in hydration and intake average out so that, across population groups, median concentrations correlate well with medians from 24-hour collections and creatinine-corrected concentrations (Zimmermann & Andersson, 2012). A median UI value of more than 100 μg/L urine is considered to reflect a normal (average) iodine intake of 150 μg or more per day) and adequate status in the community (WHO, UNICEF, & ICCIDD, 2007).

Additional iodine status indicators include serum, whole blood, or whole blood spot TSH and serum thyroglobulin concentrations. Although TSH concentrations among adults are not sensitive to deficiency, TSH measurement is recommended for screening neonates for hypothyroidism (Eastman & Zimmermann, 2017). Thyroglobulin concentration, reflecting turnover of thyroid cells, is reflective of iodine nutrition over a period of months or years. A dried blood spot assay, suitable for field assessment, now exists for thyroglobulin, making this indicator easier to measure in field settings (Zimmermann & Andersson, 2012).

Health Consequences. IDDs range from severe and life-threatening to mild disorders of lifelong consequence. Fetal and neonatal hypothyroidism, due to gestational iodine depletion, permanently alters the structure and function of the brain as well as other nervous tissues, giving rise to permanent neurologic and developmental abnormalities (Hetzel, 1994; Stein, 1994). Severe maternal iodine deficiency can also result in spontaneous abortion, stillbirth, and congenital abnormalities (Eastman & Zimmermann, 2017). Postnatal iodine deficiency perpetuates thyroid failure that, depending on its duration and severity, leads to hypothyroidism, growth

retardation, sexual immaturity, and impaired cognition and motor development (Halpern, 1994).

Cretinism (**FIGURE 7-8**) represents the most severe clinical spectrum of the IDD, usually manifested as severe mental and growth retardation, paraplegia, rigidity, deaf–mutism, and facial disturbances (Hetzel, 1994). Central nervous system defects of cretinism can be linked to severe iodine deficiency in the second trimester of pregnancy (DeLong et al., 1994; Halpern, 1994), a period when the cerebral cortex, basal ganglia, and cochlea undergo rapid growth and development. Normal amounts of thyroid hormone are needed for neuronal migration and myelination of the fetal brain—processes that are irreversibly damaged if inadequate iodine is available (Zimmermann, 2012). Severe fetal hypothyroidism gives rise to neurologic cretinism, whereas severe postnatal iodine deficiency can lead to myxedematous cretinism, marked by severe growth retardation, neurologic damage, sexual delay, and musculoskeletal deformity (Boyages, 1994). For these reasons, iodine deficiency is known as the most common preventable cause of mental retardation (Walker et al., 2007).

FIGURE 7-8 Cretinism.

Photo: Copyright John Dunn.

Mild, biochemical, or noncretinous hypothyroidism is a major public health concern due to its frequency in infancy and early childhood. Children living in iodine-deficient regions, who usually exhibit one or more IDDs, tend to have lower intelligence quotients (IQs) and perform more poorly in cognition, motor function, and school achievement tests than their iodine-sufficient peers. For example, a meta-analysis of findings from 37 Chinese publications found that children growing up in iodine-deficient areas had average IQs 12.5 points below those of children being raised in iodine-sufficient areas (Qian et al., 2005). Another meta-analysis of studies among children younger than 5 years of age found that iodine deficiency during pregnancy or childhood was associated with a reduction in IQ of 6.9 to 10.2 points (Bougma, Aboud, Harding, & Marquis, 2013).

Recognition and quantification of this subclinical "base of the IDD iceberg" have been key in motivating programs to prevent iodine deficiency (Hetzel, 1989). Globally, it has been estimated that IDDs account for 2.6 million disability-adjusted life years (DALYs) lost, primarily due to cognitive and motor impairment and hearing loss associated with deficiency (Black et al., 2008).

Goiter is an enlarged thyroid gland and is the most commonly observed clinical manifestation of iodine deficiency. Goiter size can range from barely palpable with the neck extended to grotesquely visible from a distance. The health risks posed by an enlarged thyroid due to iodine deficiency have not been well delineated. Nevertheless, hypothyroidism associated with goiter may cause lethargy, cold sensitivity, weight gain, weakness, depression or memory impairments (Chaker, Bianco, Jonklaas, & Peeters, 2017).

Zinc Deficiency

Zinc is essential for many metabolic functions, growth, and survival; in turn, zinc deficiency has many serious consequences (Hambidge & Krebs, 1999; Keen & Gershwin, 1990; Solomons, 1999). Even so, zinc deficiency has been one of the least "visible" micronutrient deficiencies. Evidence of some of the clinical abnormalities due to human zinc deficiency became clear in the 1960s (Prasad, 1985), but the prevalence of this deficiency has proved difficult to ascertain due to absence of adequate, nationally representative data on zinc status (Administrative Committee on Coordination/Sub-Committee on Nutrition [ACC/SCN], 2000). This uncertainty, which is attributed to a lack of reliable indicators or lack of consensus with respect to their use, interpretation, and target groups

to assess (Black, 1997), coupled with the limited experience with prevention programs (Gibson & Ferguson, 1998a), may have been factors that led to the near exclusion of zinc on the global micronutrient agenda until recently (Trowbridge et al., 1993).

Extensive evidence accumulated in the last two decades, primarily from randomized, controlled trials of zinc supplementation in children in low-income countries, has conclusively demonstrated the importance of zinc deficiency in these settings in regard to growth, risk of infectious disease morbidity and mortality, and other outcomes. This evidence, coupled with new estimates indicating that zinc deficiency is highly prevalent in LMICs (Wessels & Brown, 2012), has resulted in a better appreciation of the magnitude of this problem.

Function, Requirements, and Assessment. Zinc, which is found in all cells, serves as a constituent of more than 200 enzymes and numerous transcription proteins (as a "zinc finger") that regulate nucleic acid synthesis; metabolism of proteins, lipids, and carbohydrate; and cell differentiation (Cousins & Hempe, 1990; Stipanuk, 2000). These functions confer on zinc important roles in organogenesis, tissue growth, functional development, and immunity (Shankar & Prasad, 1998; Stipanuk, 2000). Such broad involvement in metabolism virtually assures zinc an important role in maintaining health.

Zinc is absorbed both by passive diffusion across a concentration gradient and by energy-dependent processes when intake is low (Stipanuk, 2000). Specific transporter proteins may facilitate its absorption (McMahon & Cousins, 1998). In mixed diets, the efficiency of this absorption can vary widely, from practically nil to 40%, with the lowest absorption associated with high grain and plant consumption and the highest absorption with human milk and meat consumption. Absorption rates tend to be higher in individuals with zinc deficiency (Solomons, 1999). Uptake into tissues appears to be regulated in some fashion, although the mechanisms involved are poorly understood.

Approximately 85% of total body zinc resides in skeletal muscle, calcified bone, and marrow; it is mostly bound to the storage protein known as metallothionein. This leaves only a small exchangeable body pool that can respond to short-term variations in zinc intake. Less than 1% of body zinc is found in circulation (Cousins & Hempe, 1990). Although the body lacks reserves of zinc in the conventional sense, bone may serve as a passive reserve, with zinc being made available from this source during normal turnover.

Recommended Dietary Allowances for zinc in healthy individuals have been available from WHO and were updated with new values developed by the Institute of Medicine (2005) in the United States. Many experts considered the IOM values to be too low for grain-based, low-meat diets, leading to the development of different recommendations by the International Zinc Nutrition Consultative Group (Hotz & Brown, 2004). For children, these values range from 3 to 14 mg per day depending on age, weight, sex, and the bioavailability of dietary zinc. For adults, the recommended intakes range from 8 to 15 mg per day depending on sex and physiologic state (i.e., pregnancy or lactation). Because zinc is lost during diarrhea and may be used more rapidly during infections, such conditions may increase the need for dietary zinc.

Zinc status may be assessed by a combination of clinical, biochemical, test-response, and dietary methods (Gibson, 2005; Gibson & Huddle, 1998). Clinical signs of moderate to severe zinc deficiency, including marked growth retardation, dermatitis and other skin changes, poor appetite, and mental lethargy, are either rare or lack sufficient specificity to be useful for population assessment. Diagnosis of more prevalent mild zinc deficiency usually rests on determining serum or plasma zinc concentration. Although unreliable for individual assessment, the distribution of serum zinc concentrations or responsiveness of the lower end of the distribution to interventions can identify groups at risk (Hotz & Brown, 2004). Serum zinc concentration varies by time of day, largely related to food ingestion, and can be lowered by infection because of cytokine-mediated shifts into the liver. Extensive recent work has resulted in suggested lower cut-offs for the assessment of serum zinc concentration depending on time of day, fasting state, age, sex, and pregnancy (**TABLE 7-11**) (Hotz & Brown, 2004).

Dietary assessment can provide valuable insight with respect to the bioavailable dietary zinc from local food resulting from concurrent estimation of zinc and phytate content, but by itself this approach cannot determine the status of an individual or population (Gibson & Ferguson, 1998b). Extensive guidelines on conducting and interpreting dietary surveys to assess the adequacy of dietary zinc intake have been published (Hotz & Brown, 2004).

Other possible measures of zinc deficiency in a population are the functional responses to zinc supplementation, such as improved growth and reduced infections, which will be considered in the following "Health Consequences" section. Because of the association of zinc deficiency with growth stunting (low height-for-age) as well as dietary inadequacy, the International Zinc Nutrition Consultative Group has developed a composite index consisting of the percentage of preschool-age children who are stunted

TABLE 7-11 Suggested Lower Thresholds for Mean Serum Zinc Concentrations (μg/dL)[a] to Classify Populations at Risk of Zinc Deficiency

	Age Group		
		10 Years or Older	
Time of Day of Blood Sample	Younger Than 10 Years	Nonpregnant Females	Males
Morning, fasting	Unknown	70 (10.7)	74 (11.3)
Morning, other	65 (9.9)	66 (10.1)	70 (10.7)
Afternoon/evening	57 (8.7)	59 (9.0)	61 (9.3)

[a] Values in parentheses represent thresholds in μmol/L. 1 μmol/L= 6.54 μg/dL.

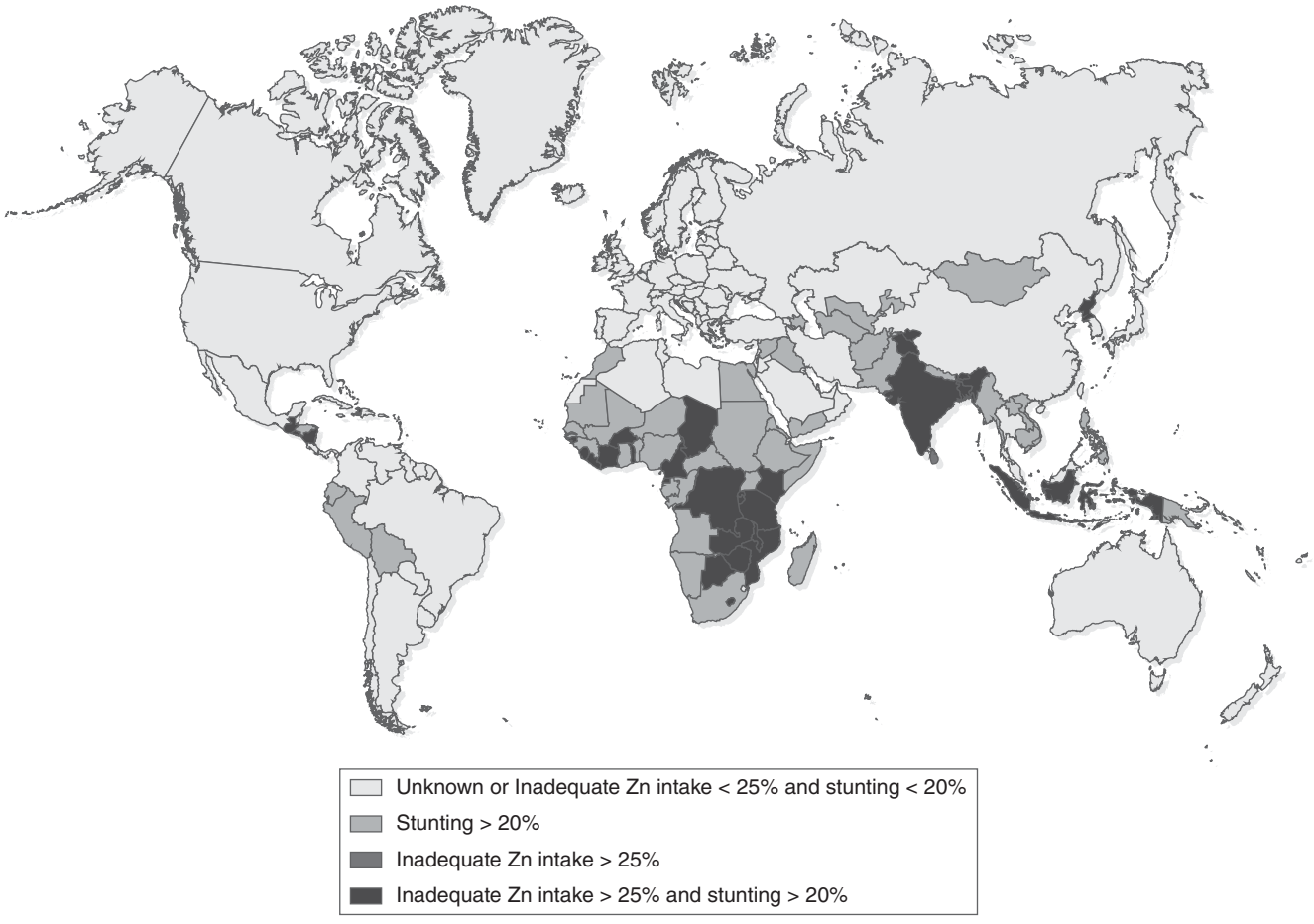


FIGURE 7-9 National risk of zinc deficiency based on the prevalence of childhood stunting and the estimated prevalence of inadequate zinc intake.

Reproduced from Wessells, K. R., & Brown, K. H. (2012). Estimating the global prevalence of zinc deficiency: Results based on zinc availability in national food supplies and the prevalence of stunting. *PLoS One, 7*(11), e50568. doi:10.1371/journal.pone.0050568

and the percentage of individuals at risk of inadequate zinc intake (from national food balance sheets) to classify countries based on their risk of zinc deficiency (**FIGURE 7-9**) (Wessels & Brown, 2012).

Health Consequences. Preschool-age children who exhibit low serum zinc levels are more likely to develop diarrhea and experience more severe episodes of diarrhea or acute respiratory infection than

children with adequate zinc status. The causality of this association has been examined by quantifying the impact of zinc supplementation on the incidence, duration, and severity of infectious diseases.

Diarrhea. Evidence for a role of zinc in reducing the incidence, severity, and duration of diarrhea is strong; 26 trials included 35 comparisons of zinc versus no zinc with 15,042 study subjects who had been provided with zinc as sulfate, gluconate, or methionate in daily or weekly supplements in doses generally of 5–15 mg per day for periods of weeks to more than a year to preschool children (Mayo-Wilson et al., 2014). Overall, there was a 13% reduction in the incidence of diarrhea (RR, 0.87; 95% CI, 0.85–0.89). There was no benefit (RR, 1.0; 95% CI, 0.96–1.05) in the subgroup of trials in which zinc was given with iron (compared to receipt of iron alone), but a significant benefit was found when zinc was given alone (compared to placebo) (RR, 0.82; 95% CI, 0.80–0.82). Zinc supplementation has also been found to be safe and to reduce the incidence of diarrhea in HIV-positive children in South Africa (Bobat et al., 2005; Siberry, Ruff, & Black, 2002).

Zinc has been successfully used in therapy of acute diarrhea. A systematic review identified 104 trials, including 18,882 study subjects, in which oral zinc was given in addition to fluids for therapy of acute watery diarrhea (Lamberti, Walker, Chan, Jian & Black, 2013). The meta-analysis of results from all trials found a 26% (95% CI, 20%–32%) reduction in the relative risk of the episode lasting more than 3 days. In those studies with pertinent data available, statistically significant benefits were found for average episode duration, duration of hospitalization, diarrhea lasting more than 7 days, and diarrhea stool frequency. Beneficial effects were similar for diarrhea associated with rotavirus infection and diarrhea for which the etiology was unknown. Of additional interest is the finding that zinc supplements given for 2 weeks during and following diarrhea have been shown to reduce the incidence of diarrhea in the subsequent 2 to 3 months (Zinc Investigators' Collaborative Group, 1999).

Respiratory Infection. A number of studies have evaluated the effects of zinc on prevention of acute lower respiratory infection. A meta-analysis of 12 zinc trials found a reduction in the incidence of lower respiratory infections (rate ratio, 0.92; 95% CI, 0.85–0.99) (Aggarwal, Sentz, & Miller, 2007). However, further analysis of these trials demonstrated that acute lower respiratory illnesses defined by specific clinical criteria were reduced to a greater degree (incidence rate ratio, 0.65; 95% CI, 0.52–0.82) (Roth, Richard, & Black, 2010).

Zinc has also been shown to have benefits as an adjunctive therapy given along with antibiotics for pneumonia. Studies in Bangladesh and India found that children with severe pneumonia who received zinc had a shorter duration of illness and a lower rate of failure for the initial antibiotic therapy (Brooks et al., 2004; Mahalanabis et al., 2004). Another study in India failed to find such a therapeutic benefit of zinc, however (Bose et al., 2006).

Malaria. Experimental zinc deficiency impairs host defenses against malarial infection (Shankar & Prasad, 1998). *Plasmodium falciparum* parasitemia has also been negatively associated with measures of zinc status or intake in Africa (Gibson & Huddle, 1998) and Southeast Asia (Gibson et al., 1991).

The public health impact of this association has been tested in four randomized, double-masked field trials (2,407 participants). A meta-analysis of these trials did not find a significant effect on the incidence of malaria (RR, 1.05; 95% CI, 0.95–1.15).

Mortality. Given the mounting body of evidence showing that adequate zinc nutriture can reduce the incidence and severity of infections, it is plausible that child mortality could be reduced by preventing even "mild" zinc deficiency through supplementation or dietary enhancement, including fortification. Several trials have provided evidence suggesting that zinc supplementation may lead to a large reduction in child mortality. One trial in full-term SGA Indian infants, who received daily zinc supplements from 1 to 9 months of age, found a reduction by two-thirds in deaths in the group compared to infants receiving a control supplement without zinc (Sazawal et al., 2001). Large randomized controlled trials of zinc supplementation in preschool-age children were also conducted in Pemba, Zanzibar, and Nepal (Sazawal et al., 2007; Tielsch et al., 2007). Although neither trial showed a benefit of zinc in children 1 to 11 months of age, both trials found a similar 18% reduction in mortality in older children (combined result RR, 0.82; 95% CI, 0.70–0.96).

In a large trial in Bangladesh, 50% fewer deaths were noted in preschool-age children who received a zinc supplement along with oral rehydration therapy for diarrhea compared to children receiving oral rehydration alone (Baqui et al., 2002). Another large effectiveness trial in India found that zinc given for diarrhea treatment reduced hospitalizations for diarrhea (OR, 0.69; 95% CI, 0.50–0.95) and pneumonia (OR, 0.29; 95% CI, 0.15–0.54) (Bhandari et al., 2008).

Poor Growth. A recent overview has clarified the extent and type of growth response that may occur when prepubescent children are given zinc on a daily basis. Thirty-three controlled trials were included in a meta-analysis that examined the effect of giving children younger than 12 years doses of 1 to 20 mg zinc

daily for periods ranging from 8 weeks to 15 months (mean of approximately 7 months) (Brown, Peerson, Rivera, & Allen, 2002). Differences in ponderal and linear growth were expressed as an "effect size" ([mean change in treatment group – mean change in control group] ÷ pooled standard deviation of the difference between groups), weighted by sample size, expressed as a standard deviation. The analysis revealed modest but statistically significant increases in weight (0.31 SD, $p < 0.001$) and height (0.35 SD, $p < 0.0001$), with larger effects seen for weight in children with lower weight-for-age and for height/length in more stunted children (Brown et al., 2002).

Reproductive Health. Zinc is required for normal maternal health, fetal growth, and development and parturition (Caulfield, Zavaleta, Shankar, & Merialdi, 1998). Experimental zinc deficiency leads to poor pregnancy outcomes (Apgar, 1985; Bunce, Lytton, Gunesekera, Vessal, & Kim, 1994). In a recent meta-analysis of 16 trials (7,637 participants) in which pregnant women either did or did not receive zinc supplements, there was a statistically significant reduction in preterm (less than 37 weeks' gestation) births (RR, 0.86; 95% CI, 0.76–0.97) (Ota et al., 2015). There was a similar, but not statistically significant, reduction in low birth weight (RR, 0.93; 95% CI, 0.78–1.12), but no reduction in stillbirths or births that were small for gestational age.

Zinc plays a known role in mammalian cell differentiation and turnover, ontogeny of mammalian systems, and thymic and other lymphoid tissue development and function (Cousins & Hempe, 1990). These effects could, for example, predispose individuals deprived of essential zinc *in utero* to permanent impairment in immunity and host resistance (Beach, Gershwin, & Hurley, 1982). Evidence from trials of zinc supplementation in pregnancy supports a beneficial effect on neonatal and infant infectious disease morbidity (Osendarp et al., 2001).

Micronutrient Frontiers

Public health frontiers exist in addressing micronutrient deficiencies. For example, current mapping of hidden hunger globally is modeled on indicators of only three nutrients (Ruel-Bergeron et al., 2015), signaling a need for more comprehensive, rapid, and affordable approaches to assess micronutrient status, establish population risk, and guide interventions (Cole et al., 2013). In particular, there are needs in the following areas:

- Articulate plausible pathways by which micronutrients protect health and prevent disease, illustrated by the apparent roles of periconceptional folate in one-carbon metabolism in preventing birth defects (Nazki, Sameer, & Ganaie, 2014).
- Clarify nutrient and disease interactions of public health importance, illustrated by the interplay of iron with pathogens affecting health (Mwangi, Prentice, & Verhoef, 2017) and the likely mediating effects of the gut microbiome in influencing status for some micronutrients (Biesalski, 2016).
- More clearly link micronutrient status to healthy eating patterns to guide dietary recommendations, requiring improvements in dietary data and regional food composition tables (World Food Programme [WFP], 2017).
- Conduct better-equipped and -informed epidemiologic studies and trials to establish public health importance of, to date, largely ignored micronutrients. For example, other B-complex vitamins in addition to folate, vitamins D (Roth et al., 2017) and E (Shamim et al., 2015), and selenium (Perkins & Vanderlie, 2016) are gaining public health recognition for their roles in materno-fetal health (Gernand et al., 2016), while supplemental antenatal intake of multiple micronutrients may affect cognitive development of offspring (Prado et al., 2012).

Obesity and Overweight

The following sections identify trends in overweight and obesity rates as well as the lifelong, detrimental effects caused by excess body weight and diet-related noncommunicable diseases.

Population Trends

Overweight and obesity are defined as abnormal or excessive fat accumulation on the body that may impair health outcomes (WHO, 2016c). Both conditions are rapidly increasing on a global scale: Worldwide, obesity has more than doubled since 1980. Measured using body mass index (BMI) cut-offs, a calculation of mass and height, an estimated 2 billion adults suffer from one or both of the conditions: 600 million adults are obese, while 462 million adults are overweight (Ng et al., 2014; WHO, 2017). Among children younger than age 5, undernutrition is still the most prevalent form of malnutrition. In contrast, among adults, overweight and obesity are the primary malnutrition burden. Global estimates indicate overweight and obesity are now responsible for more deaths among adults than underweight (WHO, 2016c).

In the last three decades, high-income countries such as the United States, the United Kingdom, and Australia have experienced rapid increases in obesity

among adult men and women (Ezzati et al., 2015). However, overweight and obesity are increasingly affecting both children and adults in LMICs as well, not only in urban centers but also in peri-urban and rural areas (Prentice, 2005). More than half of the world's obese population lives in just 10 countries, but the countries' income classifications vary widely. From the largest to smallest shares of the obese population, and classified by the World Bank as having either high-income (H), middle-income (M), or lower-middle-income (LM) economies, these countries include the United States (H), China (M), India (M), Russia (H), Brazil (M), Mexico (M), Egypt (LM), Germany (H), Pakistan (LM), and Indonesia (LM) (Ng et al., 2014).

Childhood obesity rates continue to rise in countries of all income levels (Ells et al., 2015; Kelly et al., 2013). In 2013, the prevalence of overweight and obese children and adolescents in high-income countries was 23.8% for boys and 22.6% for girls. In LMICs, this prevalence was 12.9% for boys and 13.4% for girls (Ng et al., 2014). Asian and African countries have almost half of the burden of all overweight children younger than age 5 (UNICEF et al., 2017).

Health Consequences

Data from 195 countries spanning 25 years show that excess BMI contributes to an estimated 4 million deaths (7.1% of all deaths) and accounts for 120 million DALYs (4.9% of all DALYs among adults) (Afshin et al., 2017).

Overweight and obesity are considered major risk factors for NCDs, a group of comorbidities that include cancer, cardiovascular disease (CVD), chronic respiratory diseases, and diabetes (Lozano et al., 2013; WHO, 2014b). NCDs (covered in the *Chronic Diseases and Risks* chapter, are currently the most common cause of death and disability worldwide, accounting for more than 60% of global mortality and two of every three deaths (Islam et al., 2014), with increased rates of NCD-related deaths being found in all regions of the world (Ezzati & Riboli, 2013). NCD-related deaths occur at younger ages in LMICs, with 30% of NCD-related deaths occurring before the age of 60 as compared to 13% of such deaths in high-income countries (Harikrishnan, Leeder, Huffman, Jeemon, & Prabhakaran, 2014; WHO, 2014b). CVD alone is a significant cause of premature death and the primary driver of morbidity for all NCDs, the largest burden of which occurs in LMICs (Zoghbi et al., 2014). Diabetes prevalence among adults increased 81% from 1980 to 2014, with this disease now affecting an estimated 422 million adults (NCD Risk Factor Collaboration, 2016; Korat et al., 2014).

Causes of Malnutrition

Poverty and Food Insecurity

Malnutrition is rooted in environments characterized by poverty and food insecurity. Food security, as defined by the Food and Agriculture Organization (FAO, 2017a), "exists when all people have physical, social, and economic access to sufficient, safe, and nutritious foods that meet their dietary needs and food preferences for an active and healthy life." There are four dimensions of food security: availability, economic and physical access, utilization, and stability over time. Food production, international trade, transport networks, and food storage conditions facilitate food availability within markets, while household incomes, poverty, and food prices influence household access to foods. Food utilization refers to how foods are prepared, used, and distributed within the home, and to how efficiently individuals are able to absorb and metabolize the nutrients from those foods. Individuals with frequent or chronic infections may have impaired absorption of nutrients, hindering their ability to efficiently utilize the foods that they consume. Finally, stability is influenced by environmental factors, such as weather or seasonal patterns; political factors, such as stable governance structures and trade agreements; and economic factors, such as unemployment or fluctuating food prices.

Tracking food insecurity is done in a variety of ways—through household surveys assessing access or food purchases over a given period, by monitoring food production and food prices, or through monitoring of population nutritional status. At the national level, the FAO (2017a) tracks the prevalence of undernourishment, which is defined as the proportion of the population whose dietary energy consumption is lower than that required to meet a basic need for light activity levels. This metric is constructed using national-level data from food balance sheets, which estimate the food availability within a country by summing the food production and imports and subtracting the exports, foods not used for human consumption, and food wastage. Using the population demographic structure and their known caloric requirements, one can estimate the degree to which the food energy availability within a country meets the needs of that country's population. Finally, household surveys are used to estimate the degree of inequity in access to foods within a population (FAO, 2017a).

The undernourishment metric is used for cross-country comparisons and for tracking trends over time. Since 1990, there have been substantial reductions in the prevalence of undernourishment

worldwide, which dropped from 18.6% in 1990 to 11% in 2016 (FAO, 2017a). This corresponds to a reduction from 991 million to 815 million people affected—a dramatic decrease considering the overall increase of 2 billion persons in the total global population over that same time period. Progress has been uneven between and even within regions, however, and 2017 marked the first time in 15 years in which the undernourished population has increased. The prevalence of undernourishment has declined rapidly in Latin America and Central, East, and Southeast Asia, but has remained intractable in South Asia and sub-Saharan Africa.

Food security is highly correlated with poverty, although economic growth is not sufficient by itself to eliminate food insecurity and hunger. Within many countries, inequalities exist between urban and rural areas, across regions, or between various ethnic groups. Although inequities in intra-household food allocation within a household can themselves lead to nutritional vulnerabilities (e.g., for female children or adult women in some cultures), variations in undernutrition can usually be traced to chronic food insecurity at the level of the household, the community, or larger aggregates of society. Broad economic development initiatives may fail to have an impact on food security because very poor households or communities may lack investment capital or the education or skills necessary to participate in new employment opportunities (FAO, 2017b). It is also apparent that financial crises that lead to sustained rises in food prices, losses of jobs, and lower real income initially create short-term adjustments and asset losses, which may then be followed by extended nutritional consequences (Alderman, 2010). Food insecurity is also significantly more likely in areas with acute or protracted crises arising from conflict or natural disasters (FAO, 2017b). These complex patterns of causation highlight the utility of conceptualizing food insecurity and malnutrition in present and intergenerational contexts.

Climate change poses an emerging threat to food security worldwide. Changes in the climate have already begun to negatively affect agricultural productivity. Wheat and maize yields have declined and cereal grain prices have increased following climate extremes in critical production regions (Intergovernmental Panel on Climate Change [IPCC, 2014]. In contrast, production of cassava, an important staple crop for many populations in Africa, is projected to increase, because this crop thrives in warmer temperatures and responds positively to CO_2 increases. Increased CO_2 will also affect the nutritional quality of major staple grains, with models suggesting a decrease in iron, zinc, and protein content in these foods with climate change (Myers et al., 2014, 2015, 2017). Less is known about other effects of climate change on horticultural crops or livestock production system. Marine fisheries have been impacted to varying degrees. Coral reef ecosystems provide food and other resources to 500 million people worldwide, but reef degradation has led to a decline in the abundance of fish and invertebrate species (IPCC, 2014).

The increasing incidence and frequency of extreme climate events will likely affect food security in unpredictable ways. The 2010 heat wave in Russia and floods in Pakistan led to an embargo on grain exports, creating fear of shortages and leading to subsequent price spikes (Nelson et al., 2010). Projections of the effects of climate change through 2050 suggest relatively modest global average declines on food production capacity, but after that point, grain yields may decline by 14% to 29% in the second half of the century (Nelson et al., 2010).

Conceptual Models

Two conceptual models in common use help to visualize the continuum from food insecurity to undernutrition, the causal factors involved, and the consequences of such failure to provide full nutrition and the ways in which health is affected.

One conceptual model (**FIGURE 7-10**), put forth in a series of papers in *The Lancet* focused on maternal and child nutrition, links the societal, underlying, and immediate contributors to optimal nutrition with its immediate and long-term health effects (Black et al., 2013). Building on an older undernutrition framework developed by UNICEF (1998), this model highlights that there is no single cause of optimal nutrition, but rather a set of contributing factors that vary in intensity, duration, specificity, and proximity to the undernourished individual. At the immediate level, optimal infant and young child feeding, care practices within the home, and low burden of infectious disease are immediate supporting factors. Underlying conditions of food security, caregiving resources, functioning health systems, and clean environments are necessary support systems. All of these assets are shaped by social and economic conditions, buoyed by leadership and financial resources, and backed by positive political governance structures. This conceptual model goes further than the UNICEF framework, however, by linking causes of undernutrition to potential programmatic actions that could be employed at multiple levels.

The second model illustrates the interconnectedness of nutrition across the various stages of life, from generation to generation (**FIGURE 7-11**). If individuals have access to adequate diets, reside in healthy environments, and have access to health care, they are able to achieve their health and developmental potential.

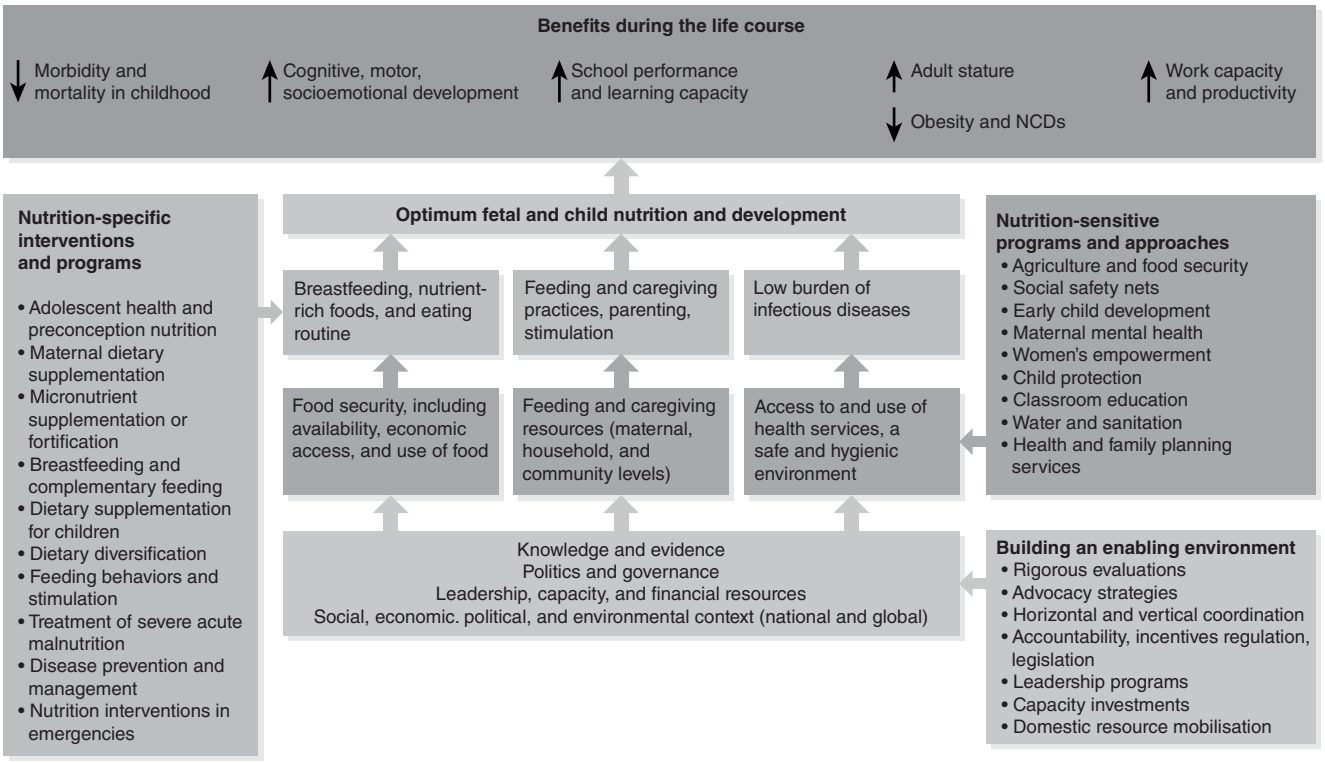


FIGURE 7-10 Framework for actions to achieve optimal fetal and child nutrition and development.

Reprinted from Black, R. E., et al. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet, 382*(9890), 427–451. Copyright 2013, with permission from Elsevier.

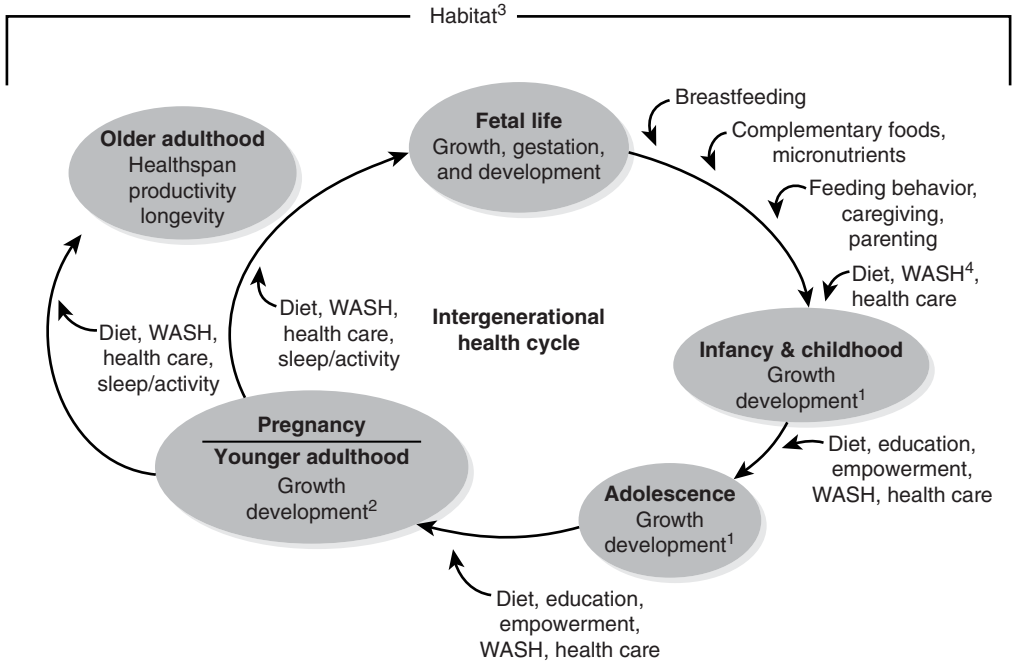


FIGURE 7-11 The intergenerational health cycle depicts behaviors, opportunities, and environmental exposures required to remain healthy and achieve genetic potential across stages of life. Ovals denote life stage status; short arrows follow activities and exposures.

[1] Physical, cognitive, motor, emotional, and maturational development.
[2] Phenotype with respect to size, maturity, strength, metabolic, mental, and reproductive capacity.
[3] Environment with respect to healthy ecosystems (air, water, biome), adequate food production and security, social equities and tolerance, good governance, and a supportive economy.
[4] WASH = household water, sanitation, and hygiene.

Modified from the Administrative Committee on Coordination/Sub-Committee on Nutrition (ACC/SCN). (2000). *Fourth report on the world nutrition situation: Nutrition throughout the life cycle.* Washington, DC: International Food Policy Research Institute. Retrieved from http://www.ifpri.org/publication/4th-report-world-nutrition-situation.

In the absence of this, infants may be growth restricted and developmentally delayed at birth, due to maternal undernutrition. Subsequently, they then face a sequence of nutritional, health, and developmental insults, mediated by socioeconomic constraints, which stunt their growth throughout early life. These factors may lead to an increased risk of infection and mortality in the short term as well as developmental delays that may affect the individual's potential to succeed in school and go on to achieve economic opportunities upon reaching adulthood. These exposures may continue through adolescence and the reproductive years. If they occur among women, they can affect both the women themselves and their offspring. Stunted, undernourished women are at greater risk of complications during pregnancy and delivery (Black et al., 2013). Their offspring are at greater risk of intrauterine growth restriction, preterm delivery, and later growth faltering (Kozuki et al., 2015; Martorell & Zongrone, 2012). Nutritional factors in childhood and young adult life also set the stage for long-term chronic health conditions, such as cardiovascular disease and type 2 diabetes (Barker, 1998). This model is meant to illustrate the value of nutrition and health inputs throughout the life course and suggest opportunities for effective and timely intervention.

Dietary Patterns Across the Life Course

A requisite for health throughout life is a diet that is adequate in quantity (calories), quality (nutrient density), diversity (variety of foods), and safety (hygienic) (Forouzanfar et al., 2015). Further, an adequate diet must be affordable (i.e., reasonably priced foods), accessible (i.e., in physical proximity; comprise foods that can be prepared and cooked), and culturally appropriate (i.e., considering local norms and traditions) (High Level Panel of Experts [HLPE], 2017). No one "ideal" diet exists for optimal health, nutrition, or environmental benefit, as these vary based on geographic location, season, and cultural context.

Infant and Young Child Diets

The infant-to-child feeding continuum should be viewed as a process, involving first breastfeeding and then gradual transitional feeding, that results in adequate nourishment to support normal growth, health, and development from birth through early childhood. The process of initiating and establishing breastfeeding, followed by phasing in complementary feeding and the eventual transition to family foods, is illustrated in **FIGURE 7-12**. Under optimal circumstances, a young infant will be fed only breast milk for a period of 6 months after birth. "Exclusive breastfeeding" is therefore defined as feeding only breast milk, excluding any solids or other liquids, including water (WHO, 2008a). Complementary feeding begins as infants are first introduced to other foods, typically soft or semi-solid transitional foods. Gradually, infants will gain the abilities to consume solid foods and eating more from the family pot. During this time, breastfeeding frequency will wane with concurrent declines in the proportion of calories and nutrients from breast milk until the young child fully transitions to family foods.

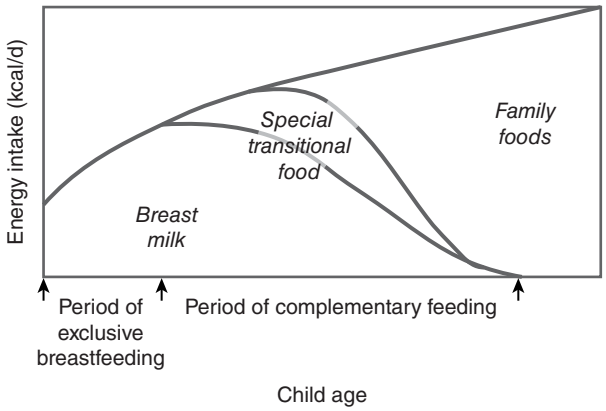


FIGURE 7-12 The sequential phases of early childhood feeding.

Reprinted from Brown, K. H., Dewey, K., & Allen, L. (1998). *Complementary feeding of the young children in developing countries: A review of current scientific knowledge.* Geneva, Switzerland: World Health Organization.

Striking differences in infant and young child feeding practices can occur both across and within regions. Patterns of breastfeeding initiation, introduction of complementary foods, and graduation to a household diet, as observed in four countries, are illustrated in **FIGURE 7-13** (WHO, 2008b). The data from Rwanda (Panel A) reveals a pattern in which more than 90% of infants exclusively breastfeed shortly after birth, with a high proportion continuing to exclusively receive human milk through age 6 months. Beyond that age, most infants begin to receive solid, semi-solid, or soft foods with human milk. In contrast, in Burkina Faso (Panel B), fewer than 20% of infants are exclusively breastfed at birth. In this culture, most young infants receive either plain water or other non-milk liquids in addition to regular breastfeeding during the first 6 months of life, beyond which solid, semi-solid, or soft foods are introduced while breastfeeding continues. South and Southeast Asian mothers tend to partially breastfeed their infants for a longer duration, indicated by a prevalence of any breastfeeding of 50% to 80% during the third year of life (Panel D), whereas a more rapid transition to the family diet is evident in Latin America (Panel C).

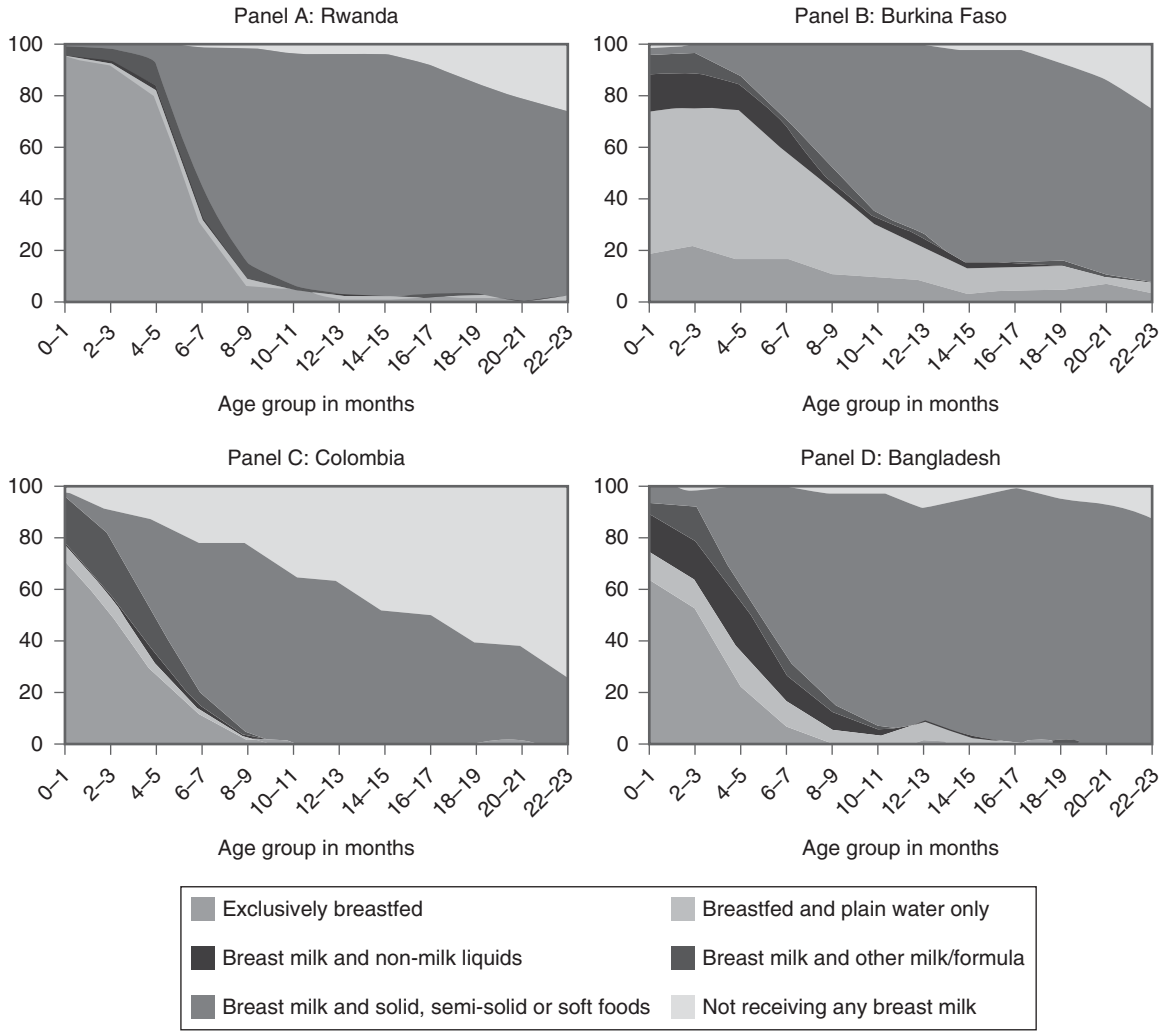


FIGURE 7-13 Distribution of breastfeeding and complementary feeding practices in Rwanda **(Panel A)**, Burkina Faso **(Panel B)**, Colombia **(Panel C)**, and Bangladesh **(Panel D)**.

Reprinted from World Health Organization (WHO). (2012). *Indicators for assessing infant and young child feeding practices: Part 3 country profiles*. Retrieved from http://apps.who.int/iris/bitstream/10665/44368/1/9789241599757_eng.pdf

Optimal breastfeeding and complementary feeding start with the initiation of breastfeeding, ideally within the first hour after birth (Debes, Kohli, Walker, Edmond, & Mullany, 2013). This ensures that the newborn will be exposed to the mother's colostrum, the thick, yellowish, immunologically rich first fluid produced by the mammary gland. While the amounts of calories or nutrients provided by the colostrum on the first 2–3 days of life are low, colostrum provides a form of an early inoculation for the infant, delivering maternal antibodies and immune factors such as secretory immunoglobulin A (IgA), lactoferrin, and leukocytes that protect the newborn as it transitions to life outside the womb (Ballard & Morrow, 2013). In a study in Nepal and Ghana, initiating breastfeeding more than 24 hours after birth was associated with a 2.4- to 3-fold increased risk of neonatal mortality compared to breastfeeding within 24 hours. The lowest risk of mortality was observed among infants breastfed within the first hour of birth (Debes et al., 2013). Globally, however, only about half of newborns are breastfed within the first hour (Victora et al., 2016).

Early initiation of breastfeeding also supports the establishment of healthy breastfeeding thereafter. WHO recommends that exclusive breastfeeding be practiced for the first 6 months of life, based on the consensus view that, on average, breastfeeding can provide adequate energy, nutrients, and fluid; protect against infection; permit normal growth for infants through this age; and reduce the risk of neonatal and infant mortality (WHO, 2013a). During the first 6 months of life, average human milk intakes for populations of infants have been noted to range between 700 and 800 mL per day in both low-income and higher-income countries (Brown, Dewey, & Allen, 1998; Dewey & Brown, 2003); this amount is adequate to meet the average child's nutritional needs.

Unlike infant formula, the composition of human milk is variable, varying within a feed, diurnally, over the course of lactation, between women, and across populations (Ballard & Morrow, 2013). The energy and protein content of human milk are relatively well conserved, even under conditions of maternal undernutrition (Prentice, 1995). The fat content of human milk is the most variable of the macronutrients, with 2–3 times higher concentrations of fat being found in the hindmilk, which is expressed later in a feed, as compared to the initially produced foremilk (Ballard & Morrow, 2013). The concentrations of long-chain polyunsaturated fatty acids, such as docosahexaenoic acid (DHA), are strongly influenced by the maternal diet (Smit, Koopmann, Boersma, & Muskiet, 2000; Urwin et al., 2013). Micronutrient quality of the breast milk is less well understood. Concentrations of certain nutrients, such as folate, calcium, iron, copper, and zinc, appear to be largely unaffected by maternal intake or nutritional status, while other nutrients—notably thiamin, riboflavin, vitamin B_6, vitamin B_{12}, choline, retinol, vitamin A, vitamin D, selenium, and iodine—are affected by maternal depletion or low intake (Allen, 2012).

Beyond its nutritive properties, human milk is a bioactive fluid that supports optimal infant health. It contains a variety of growth factors that promote tissue growth as well as maturation and repair of the intestinal mucosa (Ballard & Morrow, 2013). It also contains bioactive compounds that confer immunologic protection to the infant, as well as regulate the inflammatory response to infection. Human milk oligosaccharides (HMOs) support a healthy microbiota—the complex community of bacteria that live within the gut, promoting proliferation of healthier gut flora, and reducing susceptibility to pathogenic bacteria (Bode, 2015).

Despite the protections conferred by breast milk, young infants are at significantly elevated risk of morbidity and mortality when other liquids are introduced early in their diets. In many low-income settings, contaminated water sources serve as a primary route of exposure to enteric pathogens for young infants. Even predominant breastfeeding among infants younger than 6 months, wherein the infant receives primarily breast milk with only the addition of water, tea, or juice, is associated with a nearly 50% increased risk of mortality compared to exclusive breastfeeding (Sankar et al., 2015). By contrast, nonbreastfeeding among infants in this age is associated with a 14-fold increased risk of mortality, driven in large part by diarrhea-specific mortality.

The benefits of breastfeeding extend beyond child survival. This feeding practice has been associated with a reduction in the incidence of diarrhea and respiratory infections, increases in child intelligence scores in later life, and probable reduction in risk of overweight (Victora et al., 2016).

At roughly 6 months of age, the energy provided by breast milk is no longer adequate to meet the growing infant's needs. Unlike the strong, clear evidence that supports the benefits of breastfeeding, evidence about the benefits of specific complementary feeding practices have been mixed (Dewey & Adu-Afarwuah, 2008). Complementary feeding practices are, by contrast, a more complex set of behaviors, influenced by the diversity of locally available foods.

The total energy requirements for healthy, breastfed infants are roughly 600 kcal/day for 6- to 8-month-olds and increasing to 900 kcal/day for 12- to 23-month-olds (Dewey & Brown, 2003). While breastfeeding by itself is not sufficient to provide this amount of energy for children beyond 6 months of age, much of this energy need can be met by human milk. The added energy requirements from complementary foods for infants with "average" human milk intakes are 200 kcal/day at 6 to 8 months of age, rising to 550 kcal/day at 12 to 23 months of age (Dewey & Brown, 2003). The energy density of many complementary foods in low-income countries is low, however, because infants often consume watery porridges. To meet their energy requirements, a child needs to be fed at least 2–3 times per day from 6 to 12 months and 3–4 times per day from 12 to 24 months with food of sufficient energy density, at least 0.8 kcal/g. Protein requirements are estimated to be 1 g/kg of body weight per day for 6- to12-month-olds and 0.87 g/kg/day for 1- to 3-year-olds (IOM, 2005). Protein is rarely a limiting nutrient in infant diets, except when the overall intake of complementary foods is low (Arsenault & Brown, 2017).

Dietary diversity is associated with improved micronutrient quality of the diet and a greater likelihood of meeting the nutrient requirements for infants and young children in low-income countries (Khor, Tan, Tan, Chan, & Amarra, 2016; Moursi et al., 2008; Wondafrash, Huybregts, Lachat, Bouckaert, & Kolsteren, 2016). Animal-source foods, such as meat, fish, eggs, and dairy products, are particularly important sources of iron, zinc, calcium, and vitamin A in the diet. Dark green leafy vegetables as well as orange and yellow fruits and vegetables offer beta-carotene, a pro–vitamin A carotenoid. These nutrients tend to be limited in the diets of children in low-income settings, because their higher costs put these nutrient-rich foods out of reach for many poor families. However, even in dietary optimization studies, in

which theoretical optimal diets are constructed from locally available foods, infants would still not be able to meet their nutrient requirements from diet alone (Hlaing et al., 2016; Santika, Fahmida, & Ferguson, 2009; Tharrey, Olaya, Fewtrell, & Ferguson, 2017). Critical gaps in iron, zinc, calcium, and some vitamins still exist, even in these "optimal" diets. Micronutrient supplementation or food fortification approaches are needed to fill those gaps.

Diets of Adolescent Girls

Adolescent girls are at high risk of malnutrition due to rapid growth in body mass (skeletal, fat, and fat-free body mass) and nutrient loss through the onset of menstruation. To meet their increased nutrient needs, adequate amounts of iron, folate, calcium, and zinc are required in their diets (Salam & Bhutta, 2015).

Throughout the world, adolescent diets are often characterized by limited dietary diversity, with cereals accounting for the majority of the diet while fruit and vegetable intake is low (Ochola & Masibo, 2014). In many urban areas, adolescents have high consumption of energy-dense, nutrient-poor snacks and beverages (Ochola & Masibo, 2014). A systematic literature review of diets in LMICs found that more than 50% of adolescent girls 10 to 20 years old had inadequate levels of zinc, calcium, vitamin D, folate, thiamine, and riboflavin—micronutrients that are necessary to support growth and positive pregnancy outcomes for both mother and child (Elliot, Lutter, Lamstein, Koniz-Booher, & Caulfield, 2015).

A variety of social and cultural factors influence the diet adequacy of adolescent girls, spanning issues related to body image in some countries to early marriages and social status in others. Adolescence is a critical window of opportunity to form independent food habits (Salam & Bhutta, 2015).

Diets of Women

Menstruation, pregnancy, and lactation increase women's dietary needs throughout the life cycle. Women may consume enough calories for their daily energy needs (i.e., high-quantity diet), yet do not obtain adequate quantities of micronutrients such as vitamin A and iron (i.e., low-quality diet) (Arimond et al., 2010). In 24-hour food recall surveys in six sub-Saharan African countries, a majority of women reported consuming starchy staples, but fewer than half of those surveyed consumed legumes and nuts, vitamin A-rich fruits and vegetables, dairy products, or eggs (Kothari, Abderrahim, Colie, & Cheng, 2014).

The state of women's high-quantity/low-quality diets around the world is likely exacerbated by social inequities, whereby household resource allocation (i.e., food, water, sanitation, and health care) may limit women's access to nutritious foods or inhibit nutrient absorption. In some settings, women's and children's food intake—particularly animal-source foods—may reflect their social valuation to the community and economic contribution to the household (Gittelsohn & Vastine, 2003). Men often have greater access to fats, protein, and micronutrient-rich foods than do women (Messer, 1997). Given the domestic and work demands and constraints placed upon women, their social status can have intergenerational nutritional consequences.

How Diets Change

Substantial longitudinal data confirm that significant changes in consumer preferences, dietary habits, and micronutrient deficiency or excess have occurred throughout the last three decades (Mozaffarian, 2016). From 1990 to 2013, consumption of "healthy foods," or nutrient-rich foods with some evidence of health benefit, generally grew in all regions (Global Panel for Agriculture and Food Systems, 2016). However, some important differences in the types of food groups consumed are notable. Fruit consumption increased throughout the world, while vegetable consumption increased in only four of the seven regions. Consumption of whole grains rose substantially only in Southeast Asia, while consumption of omega-3 fatty acids declined in three of the seven regions.

By contrast, changes in consumption of "unhealthy foods," referring to nutrient-poor foods that should be consumed in moderation, have been varied. Red meat consumption declined throughout the world except in East Asia, where it rose by nearly 40%. Consumption of processed meat also increased in all regions, a trend that has been linked to certain cancers such as colorectal cancer (Micha, Wallace, & Mozaffarian, 2010). Changes in salt and sodium consumption have been minimal in all regions. Sugar-sweetened beverage consumption has risen in more than half of the regions, with the largest increase occurring in North America.

Changing dietary trends reflect a variety of shifting conditions. The nutrition transition refers to shifts in dietary patterns as populations undergo demographic transformations, including urbanization, globalization, shifting age distributions, and economic development. These shifts subsequently

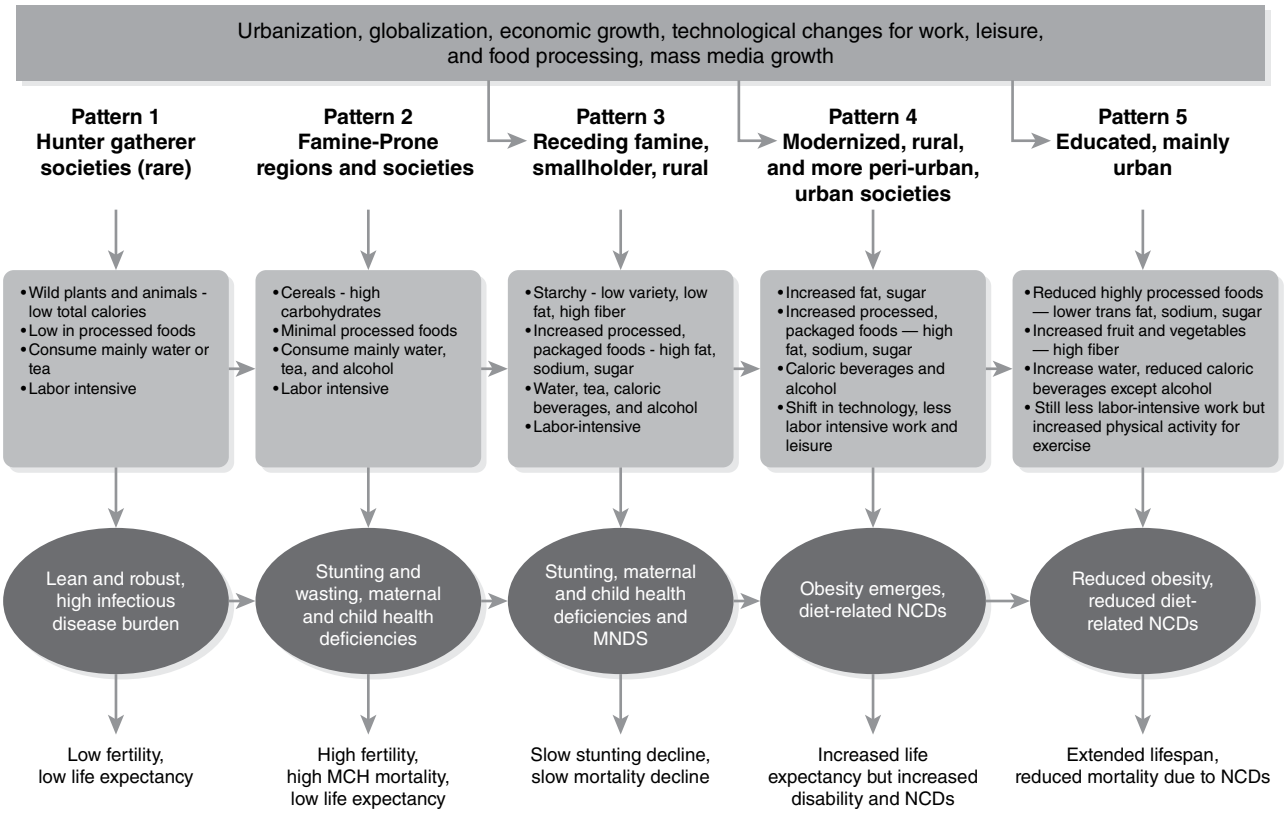


FIGURE 7-14 The five patterns of the nutrition transition.

Modified from Drewnowski, A., & Popkin, B. M. (1997). The nutrition transition: New trends in the global diet. *Nutrition Reviews, 55*(2), 31–43.

influence epidemiologic, dietary, and lifestyle patterns (Drewnowski & Popkin, 1997; Popkin, 2006). As illustrated in **FIGURE 7-14**, five nutrition transitions characterize changing food environments and food systems.

The first pattern in Figure 7-15, often called the *Paleolithic pattern* for its link to hunter–gather societies or subsistence rural communities, facilitates a local, somewhat nutritious but limited diet (Cordain et al., 2005). Individuals need to travel substantial distances to hunt game or gather wild foods, though diets contain little to no processed foods (Frassetto, Schloetter, Mietus-Synder, Morris, & Sebastian, 2009). They tend to be physically strong, but are vulnerable to infectious diseases, undernutrition, maternal and child mortality, and other health outcomes that result in a relatively short life expectancy. While some of these traditional societies still exist in the world, their livelihoods are largely threatened by globalization and climate change.

The second and third patterns reflect both the emergence of modern agriculture and periods of famine or "hunger seasons." These societies are often rural, producing food for their own consumption, though some subsistence farmers also produce food to sell. For some communities, local, traditional foods (e.g., leafy green vegetables, agroforestry foods, legumes and pulses, traditional grains and tubers) remain important components of the diet, but may be insufficient to fill nutrient gaps, initiating a state of high-quantity/low-quality diets. These populations often have a high prevalence of wasting and stunting, maternal and child mortality, and shorter life spans.

The fourth and fifth patterns represent the majority of the global population today. With urbanization, globalization, and trade liberalization, food supply chains are longer and more complex, offering access to new and diverse foods throughout the year, expanding food choices, and modifying dietary preferences. With 3 billion more people expected to enter the middle class by 2030, urbanization will affect income, gender relations, and technological development (Ranganathan et al., 2016; Seto & Ramankutty, 2016). These shifts also affect food preferences and dietary patterns. Urban consumers tend to consume more processed, convenient, street and fast foods (International Panel of Experts on Sustainable Food Systems, 2017). Highly processed foods have a longer shelf life, and are often more palatable, affordable, convenient, and easy to cook—but often lack key micronutrients and/or contain excess calories and sodium (Monteiro et al., 2017; Moreira et al., 2015; Moubarac, Batal, Louzada, Martinez Steele, & Monteiro, 2017; WHO, 2015b). Higher-income consumers in urban areas also demand greater quantities of animal source foods (Gaiha & Young, 1989; Tilman & Clark, 2014;

Timmer, Falcon, & Pearson, 1983). Changes in diet and physical activity patterns coincide with reduced rates of undernutrition and infectious diseases, but also initiate increased rates of overweight, obesity, and diet-related NCDs (Popkin, 2006).

In the fifth pattern, recognition of the poor health outcomes in the fourth pattern triggers behavioral change (e.g., reducing consumption of processed food, increasing physical activity) that can reverse the preceding negative health outcomes—although this is currently rare, even in high-income countries (Popkin, Adair, & Ng, 2012). Such behavioral changes may reflect increased education or higher-quality healthcare services.

Contributions of Infection to Undernutrition

Infectious morbidity—particularly diarrheal disease—has an important effect on physical growth and rates of malnutrition. The negative effect of diarrhea on weight and often height gain of children during and after an episode of acute diarrhea has been documented in diverse low-income country settings. Nevertheless, the magnitude of the effect on growth faltering has varied widely in these studies (Black, Brown, & Becker, 1983). Some studies, such as those conducted in Guatemala (Martorell, Yarbrough, Lechtig, Habicht, & Klein, 1975), Mexico (Condon-Paoloni, Joaquin, Johnston, deLicardi, & Scholl, 1977), and Bangladesh (Black et al., 1983) have reported that 10% to 24% of the growth faltering could be explained statistically by the prevalence of diarrhea. By comparison, studies in other countries such as Uganda (Cole & Parkin, 1977), Gambia (Rowland, Cole, & Whitehead, 1977), and Sudan (Zumrawi, Dimond, & Waterlow, 1987) have reported that diarrhea could explain as much as 40% to 80% of observed faltering. In nearly all instances, these percentages of growth faltering explained by diarrhea were higher than those explained by other infectious diseases, demonstrating the quantitative importance of the diarrhea–growth faltering relationship. In a pooled analysis of data from 9 community-based studies, Checkley et al. (2008) determined the effects of diarrhea prior to 24 months of age on stunting at 24 months: 25% of all stunting at 24 months of age was attributable to having at least 5 episodes of diarrhea in the first 2 years of life.

Infectious diseases can result in poorer weight gain or weight loss, which may be due to loss of appetite (Brown, Black, Robertson, & Becker, 1985) or increased metabolic demands related to inflammation. With diarrhea, there may also be intentional withholding of food, as well as malabsorption of ingested food in the damaged intestine (Behrens, Lunn, Northrop, Hanlon, & Neale, 1987; Lunn, Northrop-Clewes, & Downes, 1991). Variation in these factors may explain some of the differences in the magnitude of effect in various low-income country settings, but it also is necessary to consider factors such as the etiology of diarrhea, the age pattern of infection, the feeding pattern and dietary intake, treatment practices, and the length of the convalescent period.

To date, relatively few studies have examined the differential effect of diarrhea due to specific etiologic agents on growth. In Bangladesh, enterotoxigenic *Escherichia coli* and *Shigella* species were found to have the strongest effects on growth, whereas rotavirus and other enteropathogens, in part due to their lower prevalence, did not have a significant effect (Black et al., 1984b). In Peru, researchers had similar findings regarding these pathogens (Lee et al., 2014). The seasonal pattern of particular enteropathogens causing diarrhea may explain part of the seasonality in growth seen in some low-income country settings.

In regard to clinical syndromes, dysentery (bloody diarrhea) and persistent diarrhea (defined by WHO as an episode lasting 14 days or more) have particularly important adverse effects on the growth of children (Black, 1993). As one might expect, the magnitude of the weight deficit is inversely related to the duration of the diarrheal episode. Persistent diarrheal episodes usually occur in children who also have a higher burden of diarrhea, so that both the persistent episodes and the high prevalence of diarrhea adversely affect growth.

Both asymptomatic and symptomatic infections can have an adverse effect on growth. Infection with *C. parvum*, with or without illness, in Peruvian children was associated with a reduction in weight gain, after controlling for other variables (Checkley et al., 1998). Even though the effect size was smaller with asymptomatic infections than with symptomatic ones, because of their higher prevalence asymptomatic infections had a greater overall impact on growth compared to symptomatic infections. After experiencing an illness, children have the potential to grow more rapidly than they were growing previously; this phenomenon is known as "catch-up growth." Children with a *C. parvum* infection in the first 6 months of life did not have catch-up growth, such that this disease had a long-lasting adverse effect on linear growth, whereas children with infections at an older age did show some catch-up growth. A study in seven LMIC populations found that *Campylobacter* infection in early childhood was associated with growth shortfalls, increased intestinal permeability, and local and systemic inflammation (Amour et al., 2016).

Asymptomatic infection in the small intestine, as occurs in the disorder referred to as environmental enteric dysfunction or environmental enteropathy, may have a substantial effect on the growth of children (Keusch et al., 2014). This disorder is characterized by inflammation, increased intestinal permeability, and malabsorption. The continuous ingestion of large quantities of fecal bacteria in settings with poor sanitation and hygiene may be critical in the pathogenesis of tropical enteropathy. An accelerated immune response triggered by the increased small bowel permeability may divert nutrients to the production of proteins involved in the response to infection and other immune factors instead of supporting the child's growth (Campbell, Elia, & Lunn, 2003). Reduced nutrient absorption, along with a poor-quality diet, may also contribute to growth faltering in this disorder (Humphrey, 2009).

The feeding practices of a child at the time of illness can modify the effect of diarrhea on growth. Infants who are exclusively or predominantly breastfeeding tend to experience fewer adverse effects of diarrhea (Launer, Habicht, & Kardjati, 1990). This relationship may arise because breastfeeding ameliorates the severity of the illness or because breastfeeding generally seems to be continued without reduction in most circumstances during illness (Brown et al., 1985; Hoyle, Yunus, & Hen, 1980), whereas other foods may be reduced due to medical or cultural practices or anorexia because of the illness. Children in the first 6 months of life who are exclusively or predominantly breastfed may have less severe consequences of diarrhea or other infectious diseases (Khin-Maung-U et al., 1985; Launer et al., 1990; Rowland, Rowland, & Cole, 1988). Conversely, if very young children get diarrhea, they may experience long-term height deficits—effects that may be greater than if they developed diarrhea at an older age (Checkley, Epstein, Gilman, Cabrera, & Black, 2003).

Diarrhea has a lesser effect on growth among children whose usual dietary intake is greater or of better quality (Brown et al., 1988) and among children who receive food supplements, compared to children with poorer diets or those not receiving food supplements in the same setting (Lutter et al., 1989). Catch-up growth after a diarrheal episode is usually possible without specific supplementary feeding, due to the child's increased consumption of available food. However, the opportunity for catch-up growth may be limited by the duration of the healthy period between illnesses. Studies in Bangladesh (Black et al., 1983) and Zimbabwe (Moy, Marshall, Choto, McNeish, & Booth, 1994) have shown that, following a diarrhea episode, children take approximately 2 weeks to recover to their pre-illness weight and approximately 4 weeks to reach the weight that would have been expected if these children had continued their rate of growth prior to the illness. If another illness occurs during this month-long convalescent period, it may result in insufficient time for catch-up growth to occur and could add further nutritional insult (Richard, McCormick, Miller, Caulfield, & Checkley, 2014). The net effect in the long term is a reduction in both ponderal and linear growth.

Appropriate treatment of diarrheal illnesses may reduce the adverse effects of diarrhea on growth. The replacement of fluid and electrolytes with oral rehydration therapy may restore appetite and improve bowel function. Also, continued feeding during the illness results in improved weight gain in comparison to partial withholding of food during the acute phase (Brown et al., 1988). While antibiotics are not necessary for most cases of acute diarrhea, appropriate antibiotic treatment of dysentery would be expected to shorten the illness and, therefore, minimize the period of adverse effects on growth. The administration of zinc orally in treatment of diarrhea has been shown to reduce the severity and duration of diarrheal episodes, which would be expected to reduce the corresponding negative effects of diarrhea on growth (Sazawal et al., 1995). Zinc supplementation for 10 to 14 days in therapy of diarrhea is recommended by both WHO and UNICEF (WHO, 2006).

Acute respiratory infections—predominantly upper respiratory infections—have a high prevalence worldwide. Children in low-income countries, while having a similar prevalence of upper respiratory infections as children in higher-income countries, demonstrate a substantially higher rate of acute lower respiratory infections or pneumonia (Graham, 1990). Most of the studies of the effects of acute respiratory infections on growth have included both upper and lower respiratory infections and have generally not found that the illnesses alter growth patterns. In a Gambian study, children younger than 2 years of age with acute lower respiratory infections diagnosed by a pediatrician lost 14.7 g of weight per day of illness, an loss slightly, but not significantly, greater than the weight reduction observed with diarrheal diseases (Rowland et al., 1988). However, the prevalence of diarrhea was much higher than the prevalence of acute lower respiratory infections, so that diarrheal diseases explained half and respiratory infections only one-fourth of the observed weight deficit. Other studies in the Philippines (Adair et al., 1993), Papua New Guinea (Smith et al., 1991), Guatemala (Cruz

et al., 1990), and Brazil (Victora et al., 1990) have also shown that acute lower respiratory infections adversely affect growth.

Acute respiratory illnesses are associated with a 10% to 20% reduction in food intake, possibly due to a reduction in the child's appetite (Brown et al., 1985; Mata, Cromal, Urrutia, & Garcia, 1977). As with other illnesses, catabolism may also play a role. Further studies are required to assess the magnitude of the adverse effects of acute lower respiratory infections on growth and to document whether modifying factors, like those found with diarrheal diseases, exist.

A few studies have attempted to document whether malaria has an adverse effect on growth in children in low-income countries. In Gambia, malaria prevalence was shown to adversely affect weight gain but not linear growth (Rowland et al., 1977). Subsequent studies in Uganda (Cole & Parkin, 1977) and Gambia (Rowland et al., 1988) were not able to demonstrate any effect of malaria on the growth of children. A review of 12 studies that examined the relationship of malaria and stunting found mixed evidence among observational studies, but none of the randomized controlled trials of malaria interventions found an effect on stunting (Jackson & Black, 2017).

Although it has been believed for many years that measles causes a reduction in growth, this relationship has proved difficult to document on a population basis. This difficulty arises in part because of the low incidence of measles found in prospective studies, which in some cases was due to the administration of measles vaccine in the study cohort. Older studies suggest that children with measles lose weight or have reduced growth velocity (Reddy, 1991). Nevertheless, measles has been best recognized as an illness that precipitates severe clinical forms of malnutrition. In children with previous undernutrition, measles can precipitate kwashiorkor or marasmus (Morley, 1964) as well as xerophthalmia (Sommer, 1982).

Intestinal helminthic infections in children in LMICs have been associated with poor growth (Rousham & Mascie-Taylor, 1994). However, systematic reviews of trials of anti-helminth treatment of infected children or mass-deworming of infected populations have not found a consistent benefit for growth in children (Taylor-Robinson, Maayan, Soares-Weiser, Donegan, & Garner, 2015; Welch et al., 2017). This apparent inconsistency may be due to the effect of helminth infection on growth being limited to children with heavy worm infection, who constitute a very small proportion of these included in mass-deworming programs. (The *Infectious Diseases* chapter includes further information on such illnesses.)

Policies and Programs to Address Malnutrition in All Its Forms

As illustrated in the framework in Figure 7-11, policies and programs to address the determinants of malnutrition may operate at the immediate causal level, by addressing the quality of foods available and provided, care practices, or infectious disease risk. Alternatively, they may address underlying determinants of malnutrition, such as food insecurity, feeding and caregiving resources, or access to health services and water and sanitation systems. If effectively scaled and utilized, *nutrition-specific interventions* can significantly reduce stunting, wasting, micronutrient deficiencies, and even overweight and obesity. These interventions target the immediate causes of malnutrition. They largely focus on the most nutritionally vulnerable and underserved populations, such as women, and in particular, pregnant or lactating women and children younger than 2 years of age (Bhutta et al., 2013; Black et al., 2013). Examples of nutrition-specific interventions include micronutrient supplementation and exclusive breastfeeding counseling.

A growing area of research, *nutrition-sensitive approaches* address the underlying and basic determinants of malnutrition. Nutrition-sensitive approaches can serve as delivery platforms for nutrition-specific interventions. Because many of these interventions fit within multisectoral approaches, little research has identified the ways in which individual interventions contribute or interact with larger multisectoral collaborations, as well as the most efficient and systematic measurements when considered as a "package" (Reinhardt & Fanzo, 2014).

Nutrition-Specific Intervention Programs

Breastfeeding and Complementary Feeding Promotion

Although the benefits of breastfeeding have been recognized for decades, a high proportion of infants are not breastfed optimally in the first 2 years of life. Breastfeeding promotion and support programs have been shown to improve breastfeeding practices (Rollins et al., 2016). If these interventions could be scaled up where needed, it has been estimated that improved

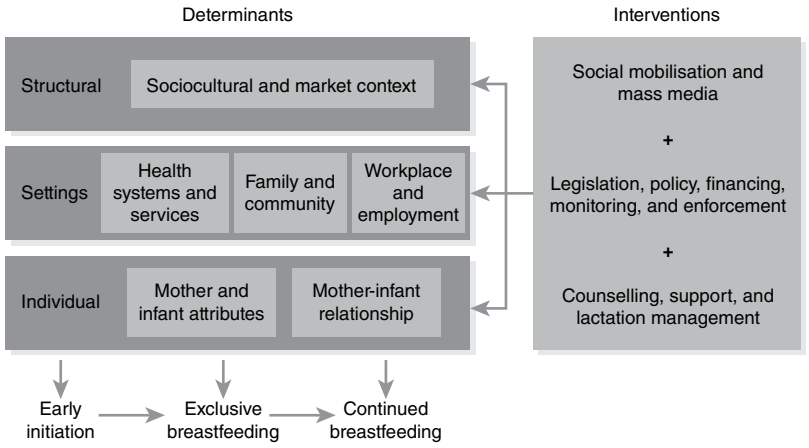


FIGURE 7-15 The components of an enabling environment for breastfeeding—a conceptual model.

Reproduced from Rollins, N. C., Bhandari, N., Hajeebhoy, N., Horton, S., Lutter, C. K., Martines, J. C., . . . *The Lancet* Breastfeeding Series Group. (2016). Why invest, and what it will take to improve breastfeeding practices? *The Lancet, 387*(10017), 491–504.

breastfeeding practices would prevent an estimated 823,000 deaths in children younger than 5 years and could protect against about half of all diarrheal infections and about one-third of all respiratory infections in young children (Victora et al., 2016). Interventions have been developed and evaluated in a variety of contexts to promote, protect, and support breastfeeding practices among new mothers (**FIGURE 7-15**). For example, they have included mass-media and social mobilization programs; legislation and policy reform; and counseling, support, and lactation management services for mothers.

Within the health system, the most widely implemented and evaluated program is the "Baby-Friendly Hospital Initiative," developed and promoted by WHO and UNICEF (2009). A hospital can be certified as "Baby Friendly" when it does not accept free or low-cost human milk substitutes, feeding bottles, or teats and has implemented the 10 steps to support successful breastfeeding (**EXHIBIT 7-2**). A meta-analysis of this program and its components found that the intervention increased exclusive breastfeeding by 49% (95% CI, 33%–68%) and any breastfeeding by 66% (95% CI, 34%–107%) (Rollins et al., 2016). Since the program was launched in 1991, more than 20,000 hospitals in 156 countries have been designated as "Baby Friendly" and are implementing the program (WHO & UNICEF, 2009).

A common reason for cessation of exclusive breastfeeding is the need for women to return to work after delivery of their infant and a lack of workplace protections to support lactating mothers (Rollins et al., 2016). Only half of all countries globally have maternity protection legislation mandating at least 14 weeks of maternity leave and, even among these, protections may not be available to women working in informal sectors (International Labor Organization, 2014).

EXHIBIT 7-2 Ten Steps to Successful Breastfeeding

Every facility providing maternity services and care for newborn infants should:

1. Have a written policy that is routinely communicated to all healthcare staff.
2. Train all healthcare staff in skills necessary to implement this policy.
3. Inform all pregnant women about the benefits and management of breastfeeding.
4. Help mothers initiate breastfeeding within a half-hour of birth.
5. Show mothers how to breastfeed, and how to maintain lactation even if they should be separated from their infants.
6. Give newborn infants no food or drink other than human milk, unless medically indicated.
7. Practice rooming-in (i.e., allowing mothers and infants to remain together 24 hours a day).
8. Encourage breastfeeding on demand.
9. Give no artificial teats or pacifiers to breastfeeding infants.
10. Foster the establishment of breastfeeding support groups and refer mothers to them on discharge from the hospital or clinic.

Reprinted from World Health Organization (WHO) & United Nations Children's Fund (UNICEF). (2009). *Baby-friendly hospital initiative: Revised*. Retrieved from http://www.who.int/nutrition/publications/infantfeeding/bfhi_trainingcourse/en/

Although only a few studies have evaluated workplace supports for breastfeeding, studies evaluating maternity leave policies have shown that provision of lactation rooms at work and paid break time for lactation are associated with significantly higher rates of exclusive and continued breastfeeding (Rollins et al., 2016).

In community-based settings, one-on-one program intervention methods have included peer counseling, community health worker interventions, lactation counselors, or other healthcare worker counseling methods. These types of programs have generally succeeded in improving exclusive breastfeeding rates by 48% (95% CI, 32%–66%). Group counseling or community-based social mobilization, with or without mass-media promotion, has increased exclusive breastfeeding rates by 20% (95% CI, 3%–39%). For example, an evaluation of two large-scale infant and young child feeding promotion programs in Bangladesh and Vietnam found that there were significant improvements in exclusive breastfeeding practice for infants residing in communities targeted with intensive interpersonal counseling coupled with mass-media promotion of breastfeeding (Menon et al., 2016). In Bangladesh, the prevalence of exclusive breastfeeding in intervention communities improved by 36 percentage points (95% CI, 21.0–51.5) compared to communities not exposed to the program. In Vietnam, the increase was nearly 28 percentage points (95% CI, 17.7–38.1). In India, a randomized evaluation of a program implemented by Anganwadi workers, traditional birth attendants, and other local providers reported a 17-fold increased odds of exclusive breastfeeding to 6 months of age and a 30% lower prevalence of diarrhea at 3 months of age in communities that received the intervention (Bhandari et al., 2003).

Marketing practices by breast milk substitute manufacturers significantly undermines breastfeeding promotion and support programs. Compelling case studies in the 1970s drew attention to these marketing practices and the associated risk of infants becoming sick, malnourished, or dying from contaminated or diluted breast milk substitutes (Muller, 1974). This recognition led to the adoption of the International Code of Marketing of Breastmilk Substitutes at the 34th World Health Assembly in 1981. This code outlines the responsibilities of governments, health systems, and companies in the marketing and manufacture of breast milk substitutes. As of 2016, 135 countries had enacted legislation containing all or many provisions of the International Code (WHO, UNICEF, & International Baby Food Action Network [IBFAN], 2016), which represents significant progress in recent years at building an enabling environment to support breastfeeding. However, enforcement of these policies remains a challenge in many countries, and only 6 countries reported having dedicated funding to support monitoring and enforcement.

Complementary feeding intervention programs target infants and young children between the ages of 6 and 24 months, at the time when they transition from exclusive breastfeeding to solid foods and their family diets. Interventions to improve complementary feeding practices have been diverse, differing in their approach to addressing key behaviors and types of foods or supplements promoted (Dewey & Adu-Afarwuah, 2008). Most intervention programs have implemented some or all of the key recommendations in the Guiding Principles of Complementary Feeding (**EXHIBIT 7-3**; Pan American Health Organization [PAHO], 2003), focusing on either providing education or behavior change recommendations, providing supplement packets that could be used for "home fortification" of complementary foods, or providing foods to improve the diversity or nutrient quality of the infant diets.

Education and counseling interventions have been found to have a modest effect on improving linear growth (mean effect size, 0.20; range, 0.04–0.64) and intake of iron-rich foods (Dewey & Adu-Afarwuah, 2008). Interventions with the largest effect sizes were those that stressed the importance of animal-source foods for young children (Guldan et al., 2000; Penny et al., 2005). More recently, large-scale evaluations of social and behavior change interventions in Bangladesh, Vietnam, and Ethiopia have been reported. While these programs have been largely effective at improving reported complementary feeding practice, notable improvements in child growth have not been observed (Kim et al., 2016; Menon et al., 2016; Rawat et al., 2017). The intervention in Bangladesh did, however, report significant improvements in child language and motor developmental outcomes (Frongillo et al., 2017). In general, these types of programs have not been as effective for food-insecure communities as the provision of food supplements or fortified food products (Dewey & Adu-Afarwuah, 2008), likely because of lack of access to high-quality, nutrient-dense foods among poorer households.

A few trials have examined the effects of interventions designed to improve responsive feeding behaviors. Responsive feeding is rooted in responsive parenting practices more broadly, reflecting the reciprocal interactions between parents and children (Black & Aboud, 2011). Interventions have focused on helping parents to navigate the progression of infant feeding development, responding appropriately to infant hunger and satiety cues, and creating a pleasant and supportive environment during feeding. In settings where underweight is a substantial public health concern, interventions have been found to improve

EXHIBIT 7-3 Guiding Principles of Complementary Feeding

1. Practice exclusive breastfeeding from birth to 6 months of age, and introduce complementary foods at 6 months of age while continuing to breastfeed.
2. Continue frequent, on-demand breastfeeding until 2 years of age or beyond.
3. Practice responsive feeding, applying the principles of psychosocial care.
4. Practice good hygiene and proper food handling by washing caregivers' and children's hands before food preparation and eating, storing foods safely and serving foods immediately after preparation, using clean cups and bowls when feeding children, and avoiding the use of feeding bottles.
5. Start at 6 months of age with small amounts of food and increase the quantity as the child gets older, while maintaining frequent breastfeeding.
6. Gradually increase food consistency and variety as the infant gets older, adapting to the infant's requirements and abilities.
7. Increase the number of times that the child is fed complementary foods as he or she gets older. For the average healthy breastfed infant, meals of complementary foods should be provided 2–3 times per day at 6–8 months of age and 3–4 times per day at 9–11 and 12–24 months of age, with additional nutritious snacks offered 1–2 times per day, as desired.
8. Feed a variety of foods to ensure that nutrient needs are met. Meat, poultry, fish, or eggs as well as vitamin A–rich fruits and vegetables should be eaten daily. Vegetarian diets cannot meet nutrient needs at this age unless nutrient supplements or fortified products are used.
9. Use fortified complementary foods or vitamin–mineral supplements for the infant, as needed.
10. Increase fluid intake during illness, including more frequent breastfeeding, and encourage the child to eat soft, varied, appetizing, favorite foods. After illness, give food more often than usual and encourage the child to eat more.

Reprinted from Pan American Health Organization (PAHO) & World Health Organization (WHO). (2003). *Guiding principles for complementary feeding of the breastfed child*. Retrieved from http://www.who.int/nutrition/publications/guiding_principles_compfeeding_breastfed.pdf

feeding practice, but effects on weight gain have been mixed (Black & Hurley, 2017). In settings where overweight is of concern, the interventions have similarly improved feeding practice and some studies (though not all) have found lower rates of weight gain in the intervention group (Black & Hurley, 2017).

Three randomized controlled trials have evaluated the impact of providing animal-source foods on young children's growth. A study conducted in Zambia, the Democratic Republic of Congo, Guatemala, and Pakistan evaluated the efficacy of daily consumption of meat in comparison to a micronutrient-fortified corn–soy blend for infants 6–18 months of age. The authors reported no significant differences between groups in growth, anemia, or iron status between groups (Krebs et al., 2012). A second study evaluating the efficacy of daily consumption of caterpillar cereal for infants 6–18 months of age also reported no significant benefits on growth, though the prevalence of anemia was significantly reduced (Bauserman et al., 2015). In contrast to the meat trial, the comparison group received their usual diet, which was unlikely to be fortified with micronutrients. A third efficacy trail in the highlands of Ecuador evaluated the impact of the daily provision of eggs for infants 6–15 months of age in comparison to a control group consuming their usual diet (Iannotti et al., 2017). In this trial, the researchers reported significant improvements in child growth and a nearly 50% reduction in the prevalence of stunting between groups. While these findings are promising, the study warrants replication in other settings to determine if the effects can be generalized more broadly.

Even with the best available foods, it is very challenging for families to meet the micronutrient requirements for infants and young children, such that nutrient gaps tend to persist, particularly in iron, zinc, calcium, and vitamin A (Arsenault, Hijmans, & Brown, 2015; Ferguson, Chege, Kimiywe, Wiesmann, & Hotz, 2015). Food supplementation interventions have often offered some type of fortified food products, such as a fortified cereal/legume blend, fortified milk powder, or a fortified lipid-based nutrient supplement. The effects of these types of interventions on growth have varied, but efficacy trials have reported generally modest effects on linear growth of approximately 0.3 standard deviation (Dewey & Adu-Afarwuah, 2008). The effects on anemia and iron status have been more consistently positive, with a mean reduction in the prevalence of anemia of 13 percentage points across all studies (Dewey & Adu-Afarwuah, 2008). Since the publication of a systematic review in 2008, many studies have sought to evaluate the efficacy of fortified lipid-based nutrient supplements (LNS) for the prevention of stunting.

While there is some variation in their composition, these supplements typically provide 100–200 kcal/day and are fortified with approximately 1 RDA of most vitamins and minerals. They are often made from a peanut or soy protein base, with added milk powder and oil (Arimond et al., 2015; Chaparro & Dewey, 2010). While some results are still pending, most studies have reported modest improvements in growth and anemia status (Abbeddou et al., 2017; Adu-Afarwuah et al., 2008; Christian et al., 2015; Dewey et al., 2017; Hess et al., 2015; Iannotti et al., 2014; Maleta et al., 2015).

Fortified cereal/legume blends and lipid-based nutrient supplements have also been evaluated as methods of preventing or treating moderate acute malnutrition, defined as a weight-for-length *z*-score between −2 and −3 (Lazzerini, Rubert, & Pani, 2013). If the problem is left unaddressed, children may progress to severe wasting, requiring more intensive treatment protocols (described in the next subsection). Acute malnutrition may occur in emergency situations, seasonally, and endemically, so it is likely that both short-term and long-term programs addressing the immediate needs as well as the underlying determinants of wasting, such as food insecurity and poor-quality water, sanitation, and hygiene, are needed (de Pee et al., 2015).

Management of Severe Acute Malnutrition

Although it affects only a small fraction of all undernourished children, severe acute malnutrition (SAM) has received programmatic attention because of its high case fatality rate without timely treatment. Severe wasting, the largest component of SAM, has an attributable fraction for child deaths of 7.4%, compared to 12.6% for moderate and severe wasting together and 44.7% for all undernutrition (Black et al., 2013).

For programmatic purposes, SAM is defined as a weight-for-height that is more than 3 standard deviations below the WHO growth standard median (considered severe wasting or marasmus), or bilateral pitting edema (edematous malnutrition or kwashiorkor), or a mid-upper arm circumference of less than 115 mm in children 6–59 months of age (WHO, 2013b). It is preferable to assess both weight-for-height and MUAC because these measurements may identify different children at risk. When screening populations of children for SAM, it is easier to use MUAC, since color-coded measurement tapes can be administered by workers with limited training. These two forms of SAM may occur together when a child with marasmus develops edema; marasmic kwashiorkor has the highest case fatality rate of the three forms of SAM. The proportion of SAM accounted for by kwashiorkor varies by world region, being high in southern Africa and lower elsewhere (Trehan & Manary, 2015).

In programs providing treatment for SAM, it is essential that the child be assessed for clinical complications that would require inpatient therapy—for example, dehydration, high fever, respiratory distress, lethargy, or severe anemia. In addition, it is important to assess anorexia by providing a test feeding of ready-to-use therapeutic food (RUTF). Specific diseases that need treatment, such as HIV/AIDS or tuberculosis, should also be considered.

For children with SAM with complications who are admitted to a health facility, WHO (1999, 2013b) has published a standard protocol for treatment. It includes initially treating infections, generally with intravenous antibiotics. A two-phase feeding approach with two therapeutic formulas, milk F75 and milk F100, is used, with the formulas gradually being replaced by solid foods. A blended food supplement, such as corn–soy or wheat–soy, or an RUTF can be used after the acute phase to continue the nutritional rehabilitation. Children can be transferred to outpatient care when their medical complications and edema are resolving and the child is alert and has a good appetite (WHO, 2013b). Vitamins and minerals should be provided in the food or as supplements. The availability of RUTF has allowed the shortening of the inpatient period from previously 30 days to only 5 to 10 days (Collins et al., 2006).

Children with SAM without complications can enter a feeding program based on the community management of acute malnutrition (CMAM) model. With active surveillance for SAM, a large proportion of cases will be identified early in the illness without complications and, therefore, will be eligible for community management. Children beginning treatment may benefit from oral antibiotics for 7 days (Trehan et al., 2013). CMAM is based on provision of RUTF, the most common form of which is a peanut paste fortified with macronutrients and micronutrients to meet standards set by WHO (WHO, World Food Programme, UN System Standing Committee on Nutrition, & UNICEF, 2007). Children can be provided with a 1- to 2-week supply of RUTF at a dose of about 150–200 kcal/kg/day (Trehan & Manary, 2015). Frequent evaluation of the child every 1 to 2 weeks is important. Treatment should continue until weight-for-height is at least −2 SD or MUAC is greater than 125 mm, depending on which enrollment criterion

was used. Although some children may meet these criteria sooner, treatment can take 10 to 12 weeks.

In well-managed programs that combine inpatient care if needed and CMAM with active surveillance for SAM, two-thirds to three-fourths of cases can be treated in the community, with the overall case fatality rate during treatment being kept as low as 5% (Collins et al., 2006). A systematic review of studies of treatment outcomes and a Delphi process was used to derive benchmarks for outcomes of treatment (Lenters, Wazny, Webb, Ahmed, & Bhutta, 2013). For inpatient treatment of complicated SAM, the estimated CFR was 14% (range, 5%–30%) and the recovery rate was 71% (range, 25%–95%). For CMAM, the CFR was estimated at 4% (range, 2%–7%) and the recovery rate at 80% (range, 50%–93%).

Because most SAM treatment programs follow children only until discharge criteria are met or for a short period thereafter, little is known about longer-term outcomes. The existing studies show substantial rates of relapse to SAM and of death, possibly due to return to impoverished and food-insecure households or to underlying disease, such as HIV/AIDS (Bhandari et al., 2016; Kerac et al., 2014). These findings emphasize the need to provide economic, nutritional, and medical support post discharge to the family through social protection and health programs.

Micronutrient-Deficiency Prevention Programs

As micronutrient deficiencies often occur where dietary quality is poor and lacking in variety, micronutrient deficiency prevention programs have historically utilized three complementary approaches: (1) supplementation; (2) fortification, extended to include biofortification of staple crops and micronutrient powders for home use (WHO, 2016b); and (3) dietary diversification, achieved by education, homestead food production, or other approaches, as detailed later in this section. The choice of strategy should be guided by the mix, extent, and severity of deficiencies, and the potential reach of programs to effectively cover those persons in need.

Supplementation. Supplementation programs are widely used to address micronutrient deficiencies. WHO recommends the supplement interventions listed in **TABLE 7-12**, based on reviews of the evidence demonstrating efficacy or effectiveness at preventing micronutrient deficiencies and reducing morbidity or mortality.

Pregnancy. In high-risk populations, current WHO guidelines suggest that women be supplemented with iron–folic acid (IFA) throughout pregnancy, as this intervention is estimated to reduce risks of anemia by 70% and iron deficiency by 57% (Pena-Rosas et al., 2015). While questions exist about whether to supplement women with iron in malaria-endemic areas, to date there is not strong evidence that IFA supplementation is associated with adverse risks in these settings (Mwangi et al., 2017), indicating it is prudent to both supplement and treat malaria as indicated during pregnancy. WHO also recommends use of calcium supplements during pregnancy. Preeclampsia and eclampsia are a leading cause of maternal mortality and preterm birth, and it has been estimated that calcium supplementation may reduce these risks by 20% and 24%, respectively (Hofmeyr, Lawrie, Atallah, Duley, & Torloni, 2014).

Many countries have adopted IFA and calcium supplementation recommendations as part of their national antenatal care guidelines, although low attendance, coverage, and stock-outs are frequent problems (Sununtnasuk, D'Agostino, & Fiedler, 2016). In such settings, facility-based IFA distribution has been complemented by community health workers who may reach women earlier in pregnancy and be less vulnerable than facilities to stock-outs (Kavle & Landry, 2017). There are also challenges to adding calcium supplements to IFA, which should be taken as three or four divided doses per day, preferably with food and apart from IFA supplements. Social support from family members and peers has been reported to improve IFA and calcium supplementation adherence (Martin et al., 2017), although more work needs to be done to improve adherence while delivering these interventions at scale.

Given that multiple micronutrient deficiencies typically coexist in poorly nourished populations, it is rational to consider replacing IFA with a multiple-micronutrient supplement during pregnancy that, although a higher-cost product, addresses a broader spectrum of nutritional deficiencies and utilizes the same delivery systems as IFA supplements. Multiple-micronutrient supplements have been compared to IFA in large field trials in Indonesia (SUMMIT Study Group, 2008) and Bangladesh (West et al., 2014), and in smaller trials in a dozen other countries. These studies reveal overall reductions of 8% to 15% in low birth weight, preterm birth, and stillbirth; some evidence of a reduction in risk of infant mortality; and a smaller but significant reduction in SGA infants (Smith et al., 2017). Benefits of multiple-micronutrient supplementation have been

TABLE 7-12 WHO Recommended Micronutrient Supplementation Interventions

Recommendation	Dosage	Rationale	Reference
Women of Reproductive Age			
Intermittent iron and folic acid supplementation	60 mg iron + 2,800 µg folic acid taken weekly	Prevention of anemia	(WHO, 2011c)
Pregnancy			
Iron–folic acid tablets	60 mg elemental iron + 400 µg folic acid taken daily	Prevention of anemia and iron deficiency	(WHO, 2016b)
Calcium	1,500–2,000 mg taken daily	Prevention of preeclampsia	(WHO, 2016b)
Infants and Children			
Semi-annual vitamin A supplementation	100,000 IU (6- to 11-month-olds) 200,000 IU (12- to 59-month-olds)	Prevention of vitamin A deficiency and child mortality	(WHO, 2011d)
Zinc supplementation	20 mg per day for 10–14 days (Treatment) 5–10 mg per day (Prevention)	Treatment and prevention of diarrhea	(WHO, 2006)
Multiple micronutrient powders for 6- to 23-month-olds and 2- to 12-year-olds	10–12.5 mg elemental iron + 300 µg retinol + 5 mg zinc + additional micronutrients to achieve 100% of the RNI taken at a target of 90 sachets over a 6-month period	Prevention of anemia and iron deficiency	(WHO, 2016a)
Iron supplements	10–12.5 mg elemental iron (6- to 23-month-olds) 30 mg elemental iron (24- to 59-month-olds) 30–60 mg iron (5- to 12-year-olds) Taken daily and in conjunction with malaria management strategies in malaria-endemic areas	Prevention of anemia and iron deficiency	(WHO, 2016a)

particularly noted when begun early in pregnancy (West et al., 2014), and among infants born to undernourished or anemic women (Smith et al., 2017).

Infancy and Childhood. High-potency vitamin A supplementation is the most common, direct strategy to prevent health consequences attributable to vitamin A deficiency in preschoolers. It is typically delivered as a 200,000-IU oral dose (half-dose to infants 6–11 months of age) in a gelatinous capsule or as an oily syrup every 4 to 6 months. While improving serum retinol concentrations for only a few months, this intervention's protection against xerophthalmia and child mortality can last for 6 to 12 months (Palmer, West, Dalmiya, & Schultink, 2012).

Scaled-up vitamin A supplementation programs, which were launched in the 1970s and expanded globally in the 1980s and 1990s, exist in more than 80 countries (Wirth et al., 2017). Supplementation typically exists as a stand-alone (vertical) activity or is packaged with other child health interventions,

such as deworming, immunizations, growth monitoring, and nutritional screenings. On average, such programs reach approximately 65% of the targeted beneficiaries (Klemm, Palmer, Greig, Engle-Stone, & Dalmiya, 2016). Supplementation should be shifted to other preventive strategies when hyporetinolemia (less than 0.70 μmol/L) is less than 5% in the target population, coupled with supportive evidence of an adequate dietary vitamin A intake in high-risk populations (Palmer et al., 2012).

Extensive evidence indicates that zinc supplementation started during diarrhea can reduce the severity and duration of illness (Lazzerini & Wanzira, 2016), supporting a recommendation that zinc supplements be provided with oral rehydration therapy for diarrhea (WHO, 2006). To date, zinc supplementation coverage has been low due to inconsistent zinc tablet supplies, inadequate healthcare provider knowledge, and lack of public demand (Lamberti, Fischer Walker, Taneja, Mazumder, & Black, 2015; Lamberti et al., 2015). These problems require greater focus on supply chain management, training of providers, and community-based advocacy to raise awareness, promote demand, and realize the public health impact of this efficacious intervention. Modest, consistent benefits of longer-term supplementation on child growth, morbidity, and survival have been observed (Mayo-Wilson et al., 2014), but have not led to recommendations for universal supplementation.

Iron supplements have long been recommended to prevent iron-deficiency anemia in children. However, a long-standing debate over whether prophylactic iron may be contraindicated where malaria is endemic was fueled by the findings of a randomized trial in a *falciparum* malaria-endemic area of Zanzibar; in that study, preschool-child iron supplementation significantly increased the risk of severe illness leading to hospitalization, death, and adverse clinical events due to malaria (Sazawal et al., 2006). Stratification in a subsample suggested that these adverse effects may have been restricted to non-anemic children. Given these findings, for a period of time, WHO had recommended that universal iron supplementation not be implemented in malaria-endemic areas without screening individuals for iron deficiency (World Health Organization Secretariat, 2007). Routine iron-deficiency screening proved challenging due to lack of simple, easily interpretable iron biomarkers for use in community settings. Thus, the net effect of this recommendation was that iron supplementation programs were halted for nearly 10 years. Updated guidance now encourages iron supplementation in both nonmalarial and malaria-endemic settings in tandem with the prevention and treatment of malaria (WHO, 2016a).

Like pregnant women, children in low-income settings are likely to have multiple micronutrient deficiencies. The previously mentioned supplement regimens target only a single nutrient deficiency, leaving unaddressed the challenge of multiple micronutrient deficiencies. Two supplemental approaches have been proposed: (1) multiple-micronutrient powders, which are single-serving sachets of a microencapsulated nutrient mix designed to be added to a child's meal (also known as point-of-use fortification), and (2) lipid-based nutrient supplements, also provided in single-serving sachets, usually combined in an approximately 100-kcal paste made from peanut, soy, and/or milk powder to which micronutrients are added as fortificants.

In several field trials in low-income countries, multiple-micronutrient powders have been found to reduce anemia and iron deficiency as effectively as iron syrups or drops (De-Regil, Suchdev, Vist, Walleser, & Pena-Rosas, 2011; Salam, MacPhail, Das, & Bhutta, 2013) and a multiple-micronutrient supplement (Adu-Afarwuah et al., 2008). Micronutrient powder delivery has been implemented in at least 22 countries, through health facilities and community distribution (UNICEF & Centers for Disease Control and Prevention [CDC], 2013). The sachets each cost approximately $0.03, or $2.70 for a 90-day course. Recent trials with LNS have also found that these lipid-based products offer significant benefits in preventing anemia and reducing iron deficiency (Abbeddou et al., 2017; Siega-Riz et al., 2014).

Fortification. Food fortification offers a potentially sustainable approach to combat micronutrient deficiencies. Foods that are fortifiable include sugar, salt, vegetable oil, flour, staple cereals, and condiments such as soy sauce, fish sauce, and bouillon cubes. Large-scale fortification has the potential to effectively replete populations with one or more deficient nutrients without significantly changing dietary behavior. Fortification is achieved by adding nutrient premixes to food items, typically at the point of production, in a controlled and regulated manner. An alternative approach, known as biofortification, involves selective breeding of staple grains, tubers, and beans to incorporate more micronutrients; it is increasingly becoming available as a means to increase micronutrient consumption and status (Bouis & Saltzman, 2017).

Globally, the oldest, most widespread, and most successful fortification program has been universal salt iodization. Salt iodization involves dry mixing or spraying food-grade salt with potassium iodate or iodide (Mannar & Dunn, 1995), targeting an iodine

concentration of 20 to 40 parts per million (ppm) to deliver, on average, 150 μg of iodine per person per day (WHO, UNICEF, & International Council for Control of Iodine Deficiency Disorders [ICCIDD], 2007). Prior to 1990, 130 countries were classified as iodine deficient, but by 2011 salt iodization programs had been scaled up to reach two-thirds of the world's population, reducing the number of deficient countries to 32 (Hodge, 2016). Despite successes, coverage remains sub-optimal or intermittent in many regions (Knowles et al., 2017), which can lead to a rapid increase in the prevalence of iodine deficiency (Zimmermann, 2004).

Vitamin A fortification has been under way in low-income countries since the 1970s, starting with an effective national program to fortify sugar in Guatemala, which was later expanded into Central America, Africa (Zambia), and Southeast Asia (Philippines). Fortification of monosodium glutamate (a flavor enhancer added to side dishes) with vitamin A was highly effective in reducing deficiency in Indonesia and the Philippines, but never scaled up due to a change in the retail product's color that raised marketing concerns with the producers (Sommer & West, 1996). The most extensive vitamin A fortification strategy to date has occurred with vegetable oil; it has reached an estimated 75% of the populations in eight countries in West Africa (Sablah et al., 2012) and is rapidly expanding into Southern Asia. Notwithstanding, effectiveness relies on numerous legislative, quality control, quality, and market coverage factors. Indeed, 97% of home-sampled vegetable oils were found to contain the prescribed 8 mg/kg of vitamin A in Abidjan (Cote d'Ivoire) (Rohner et al., 2016), but only 11% of tested sugar samples in a rural Zambian market met the statutory level of 10 mg of vitamin A per kg (Green, Kabaghe, Musonda, & Palmer, 2017).

Presently, wheat flour fortification is mandated in approximately 85 countries worldwide. (Hoogendoorn, Luthringer, Parvanta, & Garrett, 2016). While it is technically possible to fortify flours with several micronutrients, the most common fortificants are folic acid and iron, with the former intended to prevent occurrence of neural tube defects. A meta-analysis of eight population-based observational studies of folic acid–fortified flour programs concluded that these programs led to a 46% reduction in the incidence of neural tube defects, which translated into a 13% decrease in the number of neonatal deaths attributable to congenital abnormalities (Blencowe, Cousens, Modell, & Lawn, 2010). There have been few evaluations of the effectiveness of flour fortification on anemia or iron deficiency, but available data suggest this approach broadly improves iron status (Pachon, Spohrer, Mei, & Serdula, 2015).

Biofortification strategies are currently designed to enhance the iron, zinc, or vitamin A–precursor carotenoid content of staple crops such as sweet potato, cassava, maize, beans, rice, and pearl millet to provide 30% to 100% of the estimated average requirements for these nutrients (Bouis & Saltzman, 2017). Evaluation trials have demonstrated significant improvements in micronutrient status, reflected by improved iron status indicators following consumption of iron-fortified beans among women in Rwanda (Haas et al., 2016) and consumption of iron-fortified pearl millet by children in India (Finkelstein et al., 2015). Likewise, consumption of carotenoid-biofortified orange-fleshed sweet potatoes raised the vitamin A status of young children in Mozambique (Low et al., 2007). While controlled orange maize trials in Zambia have shown minimal effects of carotenoid fortification on serum retinol in children, intake was accompanied by improved (estimated) liver vitamin A reserves (Gannon et al., 2014) and function, indicated by dark-adapted pupillary responsiveness to light (Palmer et al., 2016), suggesting that consuming biofortified maize provided a broad dietary safety net in a population with marginal vitamin A status.

Regardless of the nutrients or vehicles used, fortifying a population's food supply to prevent micronutrient deficiencies requires legislation, political commitment, collaboration with industry, adherence to quality control, a plan for financial solvency, sustained market penetrance, societal engagement, and routine product monitoring under ambient conditions of use (Luthringer, Rowe, Vossenaar, & Garret, 2015).

Dietary Diversification. Micronutrient deficiencies emerge from a diet chronically lacking in nutritious food. In its quest to prevent chronic diseases across the course of life, to which micronutrient deficiencies may contribute, the WHO has long advocated a diet rich in diversity, including vegetables and fruits (Nishida, Uauy, Kumanyika, & Shetty, 2004). Population studies in low-resource settings over the past 30 years have consistently revealed low intakes of multiple micronutrients among women of reproductive age (Lee, Talegawkar, Merialdi, & Caulfield, 2012; Torheim, Ferguson, Penrose, & Arimond, 2010) and diets with low dietary diversity among children. Yet population evidence remains exceedingly sparse associating the usual dietary nutrient intakes with risk of multiple micronutrient deficiencies in women and children. Evidence exists that both (1) nutrition

counseling, social marketing, and behavior change efforts that promote nutrient-rich foods and (2) small-scale, homestead food production strategies that target poor, rural households can increase household availability or intakes of dark green leaves (Smitasiri, Attig, & Dhanamitta, 1992), eggs (de Pee et al., 1998), and other carotenoid-rich foods (Sommer & West, 1996; Talukder, Islam, Klemm, & Bloem, 1993). In terms of effects on status, while iron- and vitamin A–fortified complementary feeding in a small number of trials involving young children have generally reduced anemia and hyporetinolemia, respectively, zinc-fortified foods have not affected plasma zinc concentration (Dewy & Adu-Afarwuah, 2008). Nor have nutrition education trials generally lowered the risk of anemia, although in one setting (Burkina Faso), a 2-year integrated agriculture and behavior change intervention may have marginally improved hemoglobin levels in young children (Olney, Pedehombga, Ruel, & Dillon, 2015). Otherwise, virtually no trials exist that have tested the effects of dietary diversification on micronutrient status or risk of deficiencies in women or children, revealing both gaps in the literature and areas that are ripe for further research.

Choice of Interventions

Economic optimization models have been developed to help choose the most efficient strategy for a country, given regional variations in dietary intakes (Brown, Engle-Stone, Kagin, Rettig, & Vosti, 2015; Vosti, Kagin, Engle-Stone, & Brown, 2015). Population monitoring and evaluation are essential to improve program coverage of designees and assure efficient use of resources. There are also risks associated with the large number of intervention possibilities. Uncoordinated program delivery may lead to undesirably high intakes of certain micronutrients in some groups (Klemm et al., 2016). Nonetheless, intervention and monitoring options have never been greater, and will continue to increase, for optimizing micronutrient deficiency prevention in the future.

Obesity Prevention Programs

The NOURISHING framework for obesity reduction includes interventions in food environments, food systems, and behavioral change communication (BCC) (**TABLE 7-13**; Roberto et al., 2015). Most LMICs lack obesity-prevention policies due to the

TABLE 7-13 World Cancer Fund Research International NOURISHING Framework for Obesity Prevention

NOURIS	H	ING
Food Environment	**Food System**	**Behavior Change Communication**
Policy Area		
N	Nutrition label standards and regulations on the use of claims and implied claims on food	
O	Offer healthy food and set standards in public institutions and other specific settings	
U	Use economic tools to address food affordability and purchase incentives	
R	Restrict food advertising and other forms of commercial promotion	
I	Improve nutritional quality of the whole food supply	
S	Set incentives and rules to create a healthy retail and food service environment	
H	Harness food supply chain and actions across sectors to ensure coherence with health	
I	Inform people about food and nutrition through public awareness	
N	Nutrition advice and counseling in healthcare settings	
G	Give nutrition education and skills	

This material has been reproduced from the World Cancer Research Fund International NOURISHING framework and policy database. www.wcrf.org/NOURISHING

higher prioritization accorded to both treatment and prevention of undernutrition (Dietz et al., 2015). For instance, one review found that only 12% of LMICs (14 of 116) had proposed a policy that addressed all four of the WHO's NCD diet-related risk factors (salt consumption, fat consumption, fruit and vegetable intake, and physical activity), and 25% addressed only one risk factor (Lachat et al., 2013). Further, only one low-income country and one LMIC implemented at least one of five policy actions on the consumption of sugar sweetened beverages in 2014 (Popkin & Hawkes, 2016). As a result, much of the evidence base on obesity prevention policies and programs reflects high-income contexts (IFPRI, 2017).

Some studies suggest behavioral change communication (BCC) is viewed as the most valuable and acceptable policy action for obesity-reduction strategies (Holdsworth et al., 2013). As a result, most interventions to reduce childhood obesity are school-based programs targeting BCC among children 6 to 12 years old (Mead et al., 2017; Waters et al., 2011). In LMICs, school-based interventions have been associated with significant decreases in childhood BMI and fast-food eating behavior scores, along with increases in healthy food consumption and time spent being physically active (Verstraeten et al., 2012).

Several Cochrane reviews have examined obesity prevention and its impacts on different groups in different settings. One review (Waters et al., 2011) examined 55 studies (of which 50 were from high-income countries) on childhood obesity prevention in children age 6 to 12 years; the researchers found that, although there was significant heterogeneity of studies, some programs clearly had an impact on adiposity (measured as BMI). Program components that had beneficial effects included improving the nutritional quality of school meals, integrating healthy eating into school curriculums, promoting physical activity, providing parent support and home activities that improve diets and focus less on screen-based activities, and improving school food environments (Waters et al., 2011). Two additional Cochrane reviews summarized the results of 114 studies involving more than 13,000 children and young people in middle- and high-income countries/regions including Europe, the United States, Canada, New Zealand, Australia, Japan, and Malaysia. A combination of diet, physical activity, and behavioral change interventions was found to potentially reduce weight in children age 6 to 11 years and in adolescents age 12 to 17 years, though there were limitations in the studies and variation in the results. In children, these interventions (which were mainly school-based) may have small, short-term effects in reducing children's weight and BMI *z*-scores (Mead et al., 2017). In adolescents, some evidence suggested that the combined interventions could reduce adolescent weight by an average of 3.5 kg (Al-Khudairy et al., 2017).

While engaging with the food industry has produced only limited success in high-income countries (Roberto et al., 2015), LMICs continue to struggle to effectively coordinate systems interventions along the entire food supply chain. For instance, one review of NCDs in LMICs found that 23 of 26 studies focused solely on agricultural interventions, even though manufactured, prepackaged foods accounted for the majority of unhealthy food consumption (Pullar et al., 2016). With the rapid penetration of multinational food companies into LMIC markets, research suggests that nutrition labeling and incentives for reformulating foods for greater nutrient value will be critical to obesity prevention in these countries (Lobstein et al., 2015).

Nutrition-Sensitive Programs

Given the critical window of opportunity to influence child growth, undernutrition, and cognitive development, nutrition-sensitive approaches have indirect causal associations with reductions in malnutrition. These agriculture and food systems; water, sanitation, and hygiene (WASH); social protection; and early childhood development (ECD) programs have the potential to be brought to scale.

Agriculture and Food Systems

Several key reviews have examined the evidence to date on how agriculture can improve nutrition. While more rigorous, well-designed research is still needed, some lessons are emerging on which areas of agriculture can be constructively targeted to improve diets and nutrition (Girard, Self, McAullife, & Olude, 2012; Masset, Haddad, Cornelius, & Isaza-Castro, 2011). *Nutrition-sensitive agriculture* is a food-based approach to agricultural development that prioritizes nutrient-rich foods, dietary diversity, and food fortification at the center of the fight against malnutrition and micronutrient deficiencies (FAO, 2014). **FIGURE 7-16** illustrates the seven primary pathways that connect nutrition-sensitive agriculture interventions to nutrition outcomes. Two pathways, based on service expenditures and women's empowerment, are discussed in this section in terms of their

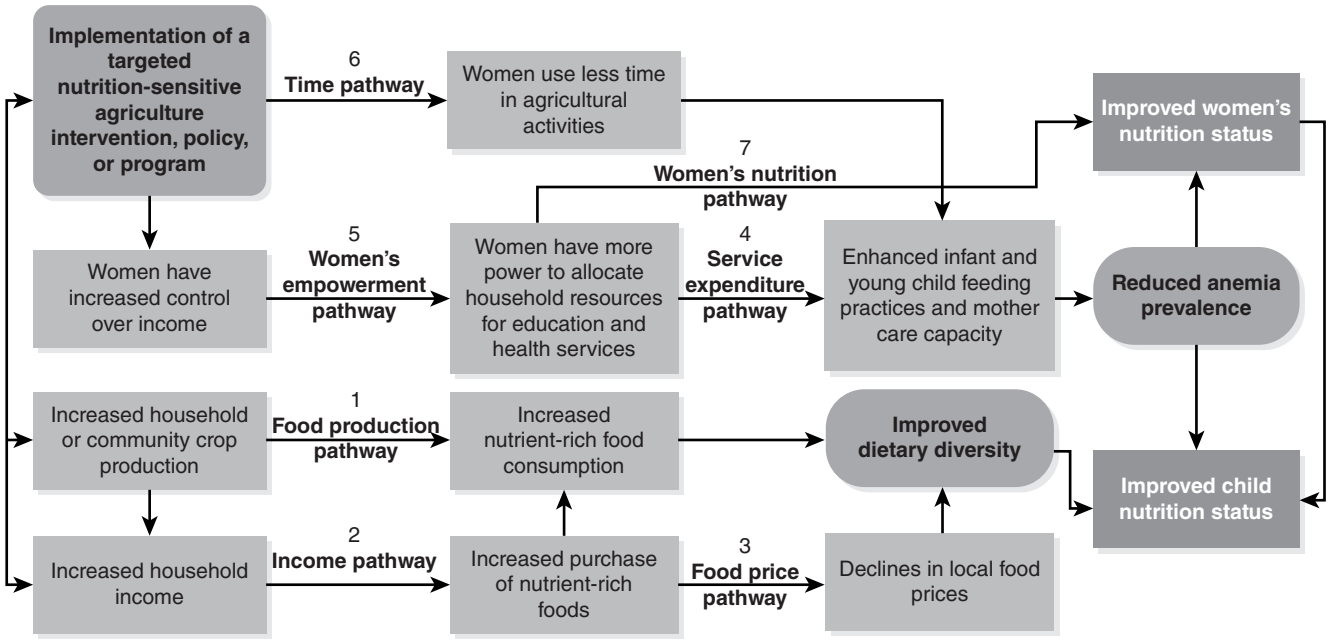


FIGURE 7-16 The pathways connecting nutrition-sensitive agricultural interventions and improvements in malnutrition.

Modified from Gillespie, S., Harris, J., & Kadiyala, S. (2012). *The agriculture-nutrition disconnect in India: What do we know?* Washington, DC: International Food Policy Research Institute.

opportunities to improve nutrition among women and adolescent girls.

Positive results from scaled-up programs have illustrated improvements in dietary patterns and intakes of specific micronutrients, including vitamin A (Girard et al., 2012; Reinhardt & Fanzo, 2014), but further research is needed to determine the direct causal linkages, if any, to childhood nutritional status (Ruel, Alderman, & Maternal and Child Nutrition Study Group, 2013). Bringing nutrition-sensitive agriculture projects to scale requires identifying context-specific enabling environments in relation to food systems, climate change and the physical environment, and coordination of investments and work with other development sectors (Hodge et al., 2015; van den Bold et al., 2015).

Agriculture is a critical component of food systems in which food supply chains and food environments can shape diets and nutrition outcomes (HLPE, 2017). The *food supply chain* encompasses all of the activities and actors that move food through production to consumption, concluding with waste disposal (Hawkes & Ruel, 2012). The supply chain is often referred to as the "farm to fork" transformation (Matopoulos, Vlachopoulou, Manthou, & Manos, 2007). Changes in decision making along this pathway can impact both food security and food's nutritional value, either positively (i.e., reductions of trans fats and sodium or increased micronutrient and macronutrient availability) or negatively (i.e., food loss and contamination). **FIGURE 7-17** illustrates the ways the food supply chain can enhance or inhibit a food's nutrient value.

The last arrow in Figure 7-18, consumption and food utilization, refers to food environments and consumer behavior. The *food environment* refers to the physical, economic, political, and sociocultural contexts that shape individual dietary preferences, food choices, and nutritional status (Global Panel on Agriculture and Food Systems for Nutrition [GLOPAN], 2017; Swinburn, Dominick, & Vandevijvere, 2014). Three elements of the food environment influence diets (HLPE, 2017):

- *Proximity and affordability*: Physical and economic access to food.
- *Food promotion, advertising, and information*: Broadcast, print, and digital advertising; packaging, labeling, and point-of-sale promotions; branding and sponsorships; merchandising and the use of licensed brand-based characters.
- *Food quality and safety*: Regulatory mechanisms to ensure quality control or safety standards; availability of cold storage; extent to which nutrient information and ingredients are readily available to consumers (Caspi, Sorensen, Subramanian, & Kawachi, 2012; Hawkes et al., 2015; Swinburn et al., 2014).

Food environments are considered "healthy" when they enable consumers to make nutritious dietary choices that may improve nutrition. Unfortunately,

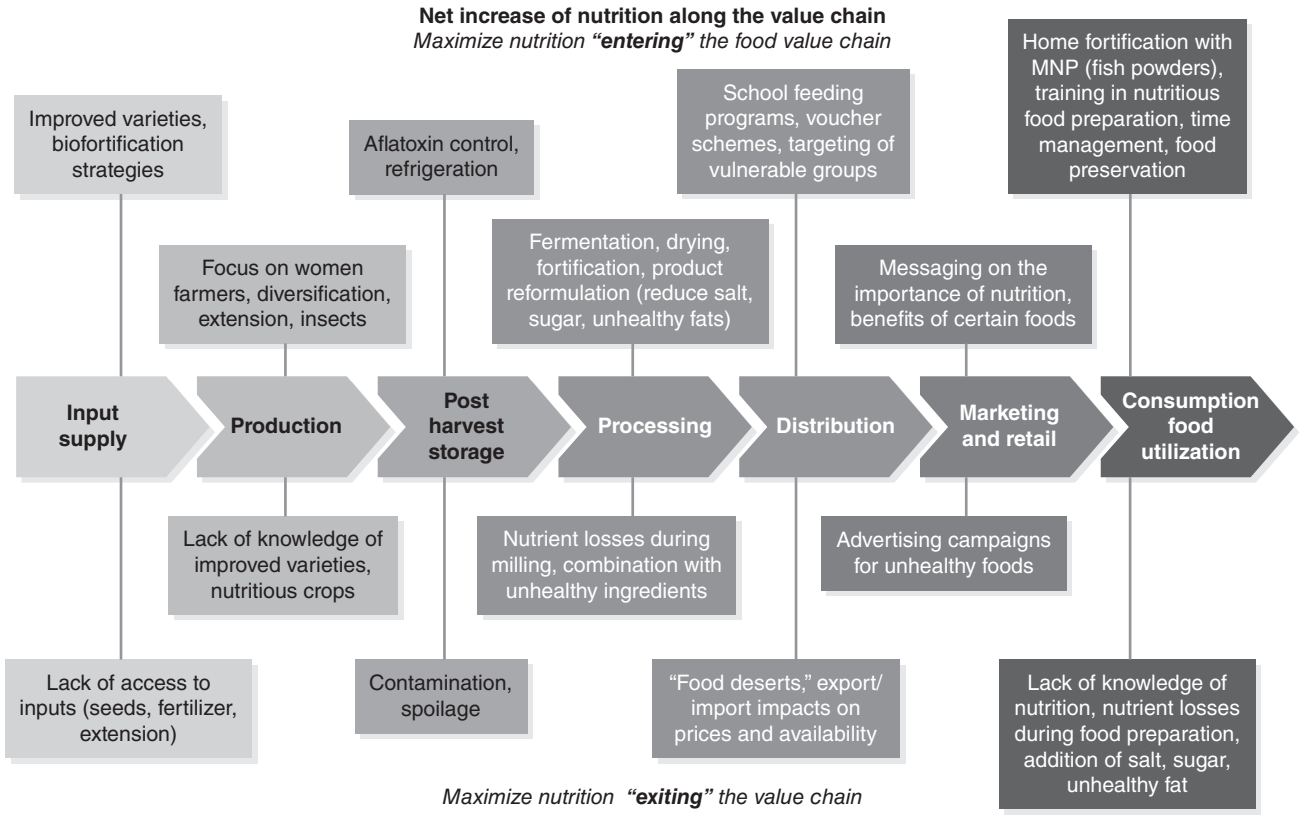


FIGURE 7-17 Maximizing and minimizing nutrition along the food value chain.

Reprinted by permission from Fanzo, J. C., Downs, S., Marshall, Q. E., de Pee, S., & Bloem, M. W. (2007). *Value chain focus on food and nutrition security.* Springer: Nutrition and Health in a Developing World.

many food environments are deemed "unhealthy" because they promote nutrient-poor dietary choices through misleading marketing (i.e., product placements and advertising), pricing policies, and deceptive labels on packaging (Baker & Friel, 2014; Malik, Willett, & Hu, 2013; Monteiro & Cannon, 2012; Moodie et al., 2013; PAHO, 2015).

Water, Sanitation, and Hygiene

Unsafe drinking water sources, inadequate sanitation, and poor hygienic practices are commonly found in LMICs, where they may contribute to malnutrition either as risk factors for diarrhea or through subclinical intestinal infections resulting in environmental enteric dysfunction (Figure 7-12). WASH interventions aim to reduce the exposure to infectious agents transmitted by the fecal–oral route (e.g., diarrhea pathogens) or through the skin (e.g., intestinal worms).

The most relevant recent systematic review examined the effects of a variety of WASH interventions on the nutritional status of children (Dangour et al., 2013). This review of 14 studies found an intervention benefit on height-for-age of small magnitude and borderline statistical significance, but no effect on weight-for-age or weight-for-height. It has been thought that limited effects of WASH interventions may reflect that previous efforts had low uptake or did not include sufficient attention to sanitation, or because comprehensive WASH interventions are needed.

Reducing open defecation may be of critical importance (Spears, Ghosh, & Cumming, 2013). One randomized trial of community-led total sanitation in Mali found a 14% reduction in the prevalence of stunting following implementation of this intervention (Pickering, Djebbari, Lopez, Coulibaly, & Alzua, 2015). However, another total sanitation trial in India did not find any effects on anthropometric measures in children younger than 5 years (Patil et al., 2014); the same was true in another trial of rural sanitation in India (Clasen et al., 2014). Furthermore, trials in Kenya, Bangladesh, and Zimbabwe that included comprehensive WASH interventions, including those related to water quality, sanitation, and hand washing, were unable to demonstrate a benefit in terms of the

nutritional status of children (Humphrey, Prendergast, Ntonzini, & Gladstone, 2017; Luby et al., 2018; Null et al., 2018).

Early Childhood Development

Poor nutrition during the first 1,000 days of life (from birth to age 2) can have lifelong consequences, including in areas such as a child's cognitive, nutritional, and physical development (Black, Gove, & Merseth, 2017). In addition, infectious diseases, maternal mental health, and environmental conditions can affect the young child's mental development.

Early child development (ECD) programs have been conceptualized as providing "nurturing care" for infants and young children (Black, Gove, & Merseth, 2017). This care includes nutrition, health, early learning, responsive caregiving, and safety, and is based on caregiver–child interactions that enable children to reach their developmental potential (Black, Gove, & Merseth, 2017). Evidence for positive benefits from such interventions on mental development have largely come from programs that supported psychosocial stimulation to infants and young children through a variety of approaches, including home visits, group sessions, health center appointments, links with conditional cash transfer programs, or a combination of these (Aboud & Yousefzai, 2016). Successful programs have used active learning to enable caregivers and paraprofessionals to adopt stimulation practices. Stimulation has also been integrated with programs to promote nutrition—a relationship that makes sense given the role of undernutrition as a determinant of poor development—but the expected synergistic effects have not been demonstrated (Aboud & Yousefzai, 2016).

While many ECD programs have focused on the first 1,000 days, the need for supporting a child's development continues after the first 2 years of life. The platforms that are used to reach children after this age include childcare settings, preprimary schools, and primary schools (Black et al., 2017). All of these are also settings where nutritional interventions can be implemented. In approaches targeting stunting, the emphasis has rightly been on the first 1,000 days of life, but there can be additional nutritional benefits from these programs, especially with the childcare and preprimary age groups. While school feeding programs have not been recommended for their nutritional benefits, they may enhance school attendance and learning (Bhutta et al., 2013).

Women's Empowerment

Inclusive of social status, mental health, and nutritional outcomes, women's empowerment influences when women first have children and their associated nutritional status at birth, as well as resources at their disposal and the care they can provide. Empowered and educated women are more likely to maintain healthy infant and young child feeding (IYCF) practices (Malapit & Quisumbing, 2015).

The socioeconomic and political status of women impact child and maternal health through three primary pathways: autonomy and control of household resources (i.e., education and healthcare expenditures), workload and time availability (i.e., IYCF practices), and social support networks (i.e., gender-equitable education attainment and food distribution) (Cunningham, Ruel, Ferguson, & Uauy, 2015). A substantial literature indicates that improvements in women's socioeconomic status have long-term benefits for child nutritional status (Kamiya, 2012), household functioning (Quisumbing & Maluccio, 2003), and education equity (Yoong, Rabinovich, & Diepeveen, 2012). Nutrition-sensitive interventions targeted toward improving women's empowerment must be culturally salient and sensitive to local contexts, using holistic measurements to assess changes women's socioeconomic status (Ghuman, Lee, & Smith, 2006; Yount, VanderEnde, Dodell, & Cheong, 2016). Women's empowerment programming often incorporates social protection programs (discussed later in this section) and family planning.

Maternal Health and Family Planning

Exemplifying the vicious cycle of undernutrition, women's nutritional status has direct effects on their children's nutrition at birth, and potentially throughout their lives. Three significant determinants of maternal and child nutrition are child marriage, the age of first pregnancy, and spacing of pregnancies (Fraser, Brockert, & Ward, 1995; Goli, Rammohan, & Singh, 2015). Recent evidence suggests children born to adolescent girls who were married before age 18 have a 25% higher risk for stunting than children born to women who married as adults (Efevbera, Farmer, Bhabha, & Fink, 2017). Further, these children are at an increased risk of low birth weight and have a 50% higher risk of perinatal death as compared to newborns of older mothers (Santhya, 2011; Viner et al., 2011). Lastly, pregnancy complications are the primary cause of death (complications of preterm delivery) for children younger than age 5 and the second

leading cause of mortality in adolescent girls worldwide (UNICEF, 2015; WHO, 2011c).

Effective interventions to reduce child marriage rates in LMICs include social protection programs conditional on school attendance and reproductive health education and services (Kalamar, Lee-Rife, & Hindin, 2016). Provision of low-cost or free contraception is cost-effective, ameliorating the high unmet demand for pregnancy prevention methods. Nevertheless, one analysis found that only 75% of women in 76 countries were able to access contraception, demonstrating a remaining gap in coverage (UN, 2015b). Family planning counseling can help women space out pregnancies to reduce child mortality (Ganatra & Faundes, 2016; Zakiyah, van Asselt, Roijmans, & Postma, 2016).

Health

An expansive body of literature has demonstrated that the public health sector is key for delivering high-impact and cost-effective nutrition-specific and nutrition-sensitive interventions (Bhutta et al., 2013), including interventions directed toward both underweight and overweight/obesity. To address malnutrition among children, targeted health interventions generally aim to manage and prevent nutrition-related diseases or provide counseling on nutrition through the healthcare system. Among underweight populations, immunization packages, antibiotics, and BCC messaging about both non-communicable and communicable diseases, such as pneumonia and diarrhea, intestinal worms, malaria, and tuberculosis, are often delivered through the health system (Boschi-Pinto, Dillip, & Costello, 2016; Buse & Tanaka, 2011). Among overweight children, interventions aim to teach healthy eating patterns, prevent and/or treat diabetes, and encourage weight loss and physical activity, but have not shown much success (Foster, Farragher, Parker, & Sosa, 2015). Recent innovations in delivery strategies, including mobile-health technologies and community-led platforms, represent avenues through which nutrition-sensitive interventions can be brought to scale (Kahn, Yang, & Kahn, 2010; WHO, 2010a).

Education

Children of mothers who no formal education are the most at risk of poor nutrition, regardless of socioeconomic status and rural or urban residence (Alderman & Headey, 2014; Hasan, Soares Magalhaes, Williams, & Mamun, 2016; Subramanyam, Kawachi, Berkman, & Subramanian, 2011). In turn, the strongest predictor of childhood nutrition status is maternal education (Alderman & Headey, 2014; Wamani, Tylleskär, Åstrøm, Tumwine, & Peterson, 2004). The pathways through which maternal education influences child nutrition include improved health-seeking IYCF, WASH behaviors, and promotion of diverse food consumption (Burchi, 2010; Casanovas et al., 2013; Saleem, Mahmud, Baig-Ansari, & Zaidi, 2014).

Nutrition-sensitive education interventions include both primary or secondary schooling and nutrition education. In more than 100 LMICs, public investment in primary and secondary education have been shown both to improve women's health outcomes and to reduce stunting among children (Headey, 2013). Recent evidence suggests the threshold for significant stunting reductions is 10 years of schooling (senior secondary level and above), but completion of primary school can significantly reduce wasting and underweight (Makoka & Masibo, 2015).

Nutrition education must be culturally sensitive and adapted to local contexts. Educated women who have greater access to appropriate nutrition information have more opportunities for employment, higher earning power, and healthier children (Semba et al., 2008).

Social Protection

Social protection programs aim to intervene at key leverage points during the lifespan, reducing vulnerability and investing in human capital among low-resource populations (HLPE, 2012; Tirivayi, Knowles, & Davis, 2016). The most cost-effective social protection programs linked to improvements in nutrition outcomes include school feeding programs and subsidies or vouchers for child education and health services (Akresh, de Walque, & Kazianga, 2013; Bundy & Alderman, 2012; Lagarde, Haines, & Palmer, 2007). Other effective social protection programs include conditional or unconditional cash transfers (CCT and UCT, respectively). Both types of programs provide supplementary income in an effort to stabilize household income and provide resilience against economic shocks, when families would otherwise skip meals or buy inexpensive, nutrient-poor foods (Fenton et al., 2016; FAO, 2015). Evidence suggests poverty alleviation programs have direct effects on individual nutrition outcomes, particularly child anthropometry (Headey, 2013; Leroy, Ruel, & Verhofstadt, 2009). Programs are most successful when accessible foods are both nutritious and diverse, intra-household allocation is equitable, and implementation is coordinated among multiple sectors (FAO, 2012; Ruel et al., 2013).

The Future of Nutrition-Sensitive Approaches

Despite the success of nutrition-sensitive interventions for a variety of nutritional, dietary, and development outcomes, many programs have yet to be adequately funded or brought to scale. Policy makers in LMICs face significant obstacles in implementing multisectoral programs. Inequities in public services provision, limited program reach, funding gaps, and substantial understaffing are all challenges that must be overcome in program implementation. Project responsibilities may overlap between the Ministries of Health, Education, and Agriculture, as well as the dispersed activities of nongovernmental organizations (NGO)s and civil society groups. To date, the limited indicators and metrics used to track success and monitoring mechanisms have failed to demonstrate their ability to effect improvements in health and nutrition outcomes (Gill & Stewart, 2011; van den Bold, Quisumbing, & Gillespie, 2013).

Enabling Environments for Improved Nutrition

Good governance is defined as the use of public authority, agencies, and funding to provide effective and transparent public services that can withstand changing political or socioeconomic contexts over time (Acosta Mejia & Fanzo, 2012; Solon, 2006). *Nutrition governance* is therefore the prioritization and funding of nutrition programming in national policy agendas (WHO, 2010b). Characteristics to assess the quality of nutrition governance in any given country include the adoption of a national nutrition policy or dietary guidelines, regular nutrition monitoring and surveillance, and the existence of a funding line for nutrition in annual budgets (FAO, 2015).

Enabling environments, which are embedded within the theory of good nutrition governance and the political economy of nutrition, are deliberate actions incorporated into political and policy processes that build and sustain momentum for effective program implementation and enforcement (Gillespie, Haddad, Mannar, Menon, & Nisbett, 2013; Gillespie, Menon, & Kennedy, 2015; Haddad, 2013). Five primary components are often used to gauge the enabling environment for nutrition (Acosta Mejia & Fanzo, 2012; Haddad, 2013):

- Policy makers must politically commit to nutrition, by ensuring that the issues of undernutrition as well as overweight and obesity are visible and prominent on the national development agenda.
- There must be substantive allocation of funding for both nutrition-sensitive and nutrition-specific interventions.
- Government agencies or ministries must effectively coordinate—both vertically from national to local levels and horizontally—the various stakeholders working in the country on evidence-based nutrition programs.
- Governments must invest in the systemic, institutional, and human capacity needed to implement and deliver nutrition programs and interventions.
- Good policy requires transparency, accountability, efficiency, and cost-effectiveness—all of which require solid monitoring and evaluation mechanisms with robust national data information systems (Arora, Chauhan, John, & Mukhopadhyay, 2011).

With enhanced cooperation between governments, NGOs, and the global nutrition community since the start of the twenty-first century, the enabling environment for nutrition has improved. Though undernutrition has yet to be fully eradicated, the timeline in **FIGURE 7-18** illustrates the increased visibility, funding, coordination, and accountability mechanisms over the last two decades.

How has nutrition governance been measured? One such measure is the Hunger and Nutrition Commitment Index (HANCI), developed in the early 2010s to rank 45 governments on their political prioritization of reducing hunger and malnutrition through legal frameworks, policies, and public expenditures (te Lintelo & Lakshman, 2015). Because most political commitments are voluntary, other, more expansive accountability mechanisms include the Global Nutrition Report and the Scaling Up Nutrition Movement.

Transparency and accountability have proved to be the most challenging components of nutrition governance in the twenty-first century. Five of the six primary challenges identified by (Balarajan & Reich, 2016) that impede progress in the political economy of nutrition also incorporate accountability and themes from the study of political economy, as illustrated in **FIGURE 7-19**.

First, while some nutrition outcomes may be immediate, many multisectoral interventions often require years to demonstrate their impact. Data are the most readily available means to hold nutrition actors accountable to their commitments, but results

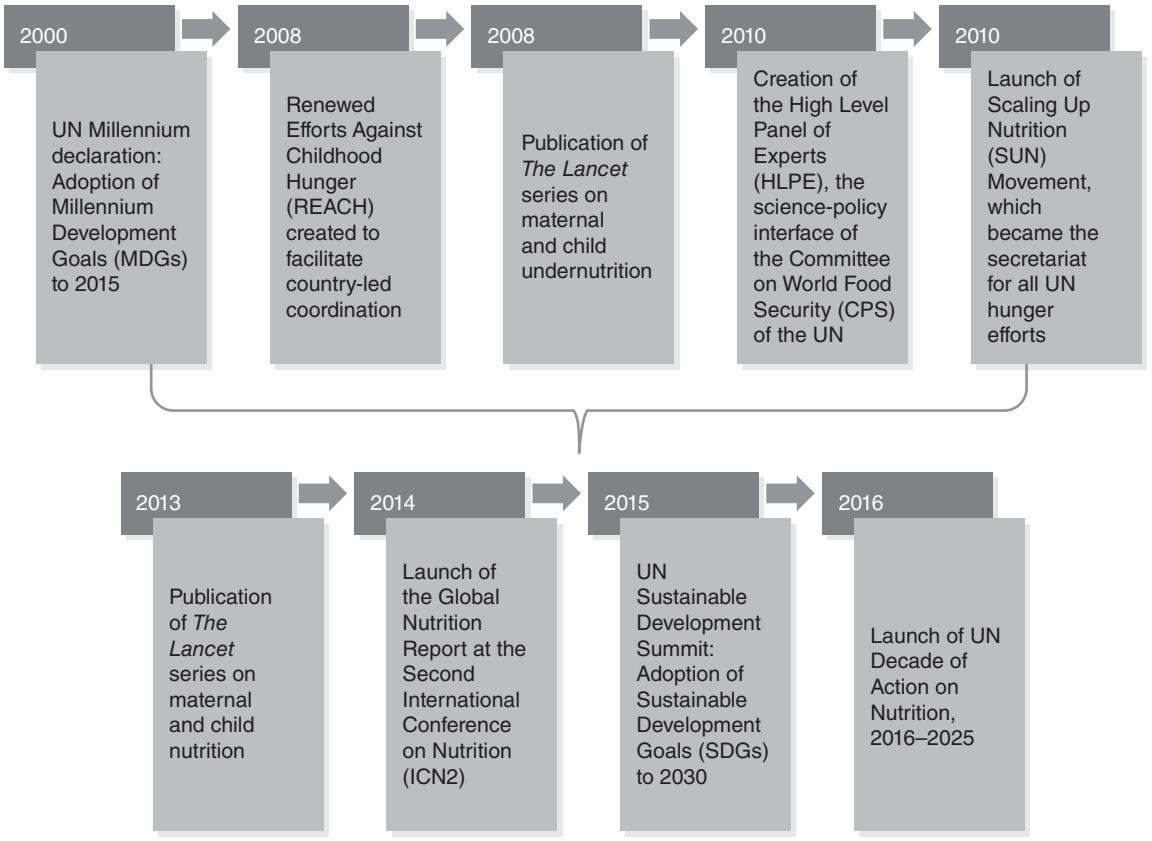


FIGURE 7-18 Timeline of global commitments, policies, and research priorities in the fight against undernutrition.

Modified from Gillepsie & Harris, 2015; Nisbett, Gillespie, Haddad, & Harris, 2014.

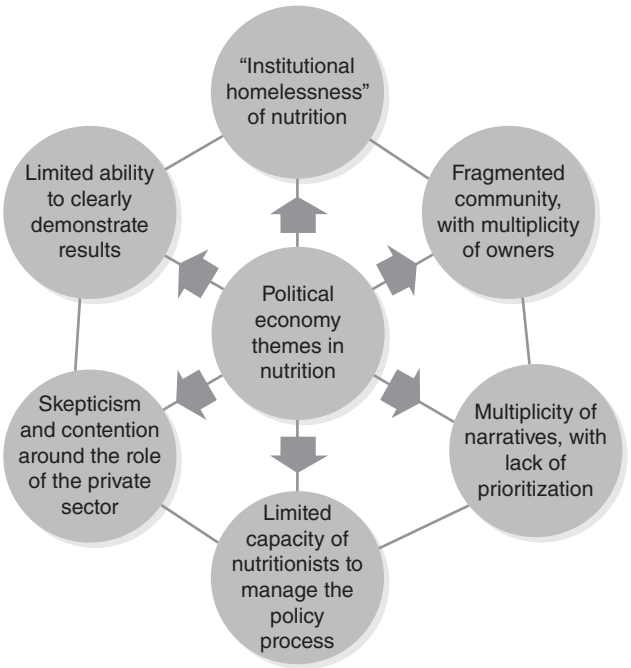


FIGURE 7-19 The six political economy factors of nutrition policy.

Reproduced from Balarajan, Y., & Reich, M. R. (2016). Political economy challenges in nutrition. *Globalization and Health, 12*(1), 70. Retrieved from http://creativecommons.org/licenses/by/4.0/

may not always be linear or easily quantifiable. Further, the variety and quantity of interventions make it difficult to compare results (Swinburn et al., 2015).

Second, "institutional homelessness" refers to the multisectoral nature of nutrition-sensitive interventions (Pinstrup-Anderson, 1993). Countries approach the design and responsibility for nutrition policy in a variety of ways; for example, nutrition responsibilities can be housed in the health or agriculture ministry, or may even be addressed by a stand-alone agency.

Third, and relatedly, fragmentation of ownership reflects nutrition policy's malleability (Morris, Cogill, & Uauy, 2008). Effective nutrition-sensitive interventions must be gender sensitive and culturally salient, targeted to a precise population (Ruel et al., 2013). In turn, opinions on the most relevant solution may differ widely among researchers, nutritionists, policy makers, community or religious leaders, and the population itself—and this discord impedes transparency and slows the policy process.

Fourth, nutritionists have low capacity, primarily working at the local level. Policy makers are generally not nutritionists, complicating the

EXHIBIT 7-4 Obstacles for Nutrition Governance in the Fight Against Obesity

In the last two decades, worldwide efforts to reduce obesity have been marked by poor progress. WHO's (2013) Global Action Plan for the Prevention and Control of NCDs set voluntary, international objectives through 2020 for improvements in health outcomes. The primary target for obesity—that is, to *halt* the rise in diabetes and obesity—recognizes the institutional, physical, psychosocial, and political complexities inherent in the political economy of obesity.

Researchers suggest obesity rates have plateaued or increased in most LMICs due to two conflicting core narratives over who is responsible for taking action to reduce those rates: The individual (consumers should make nutritious choices) or society (government should provide leadership through intervention) (Swinburn et al., 2015). Worldwide, most policies and programs have assumed the former, making it difficult to collect data, identify enabling environments, or hold powerful nutrition actors accountable.

The food industry, which devotes millions of dollars to lobbying for processed foods, has successfully opposed government efforts to create food-based obesity policies (Moodie et al., 2013; Popkin, 2013). Suggestions for enhancing accountability and transparency include managing private-sector conflicts of interest (i.e., food marketing to children, industry-sponsored nutrition studies), government monitoring, and improved engagement with civil society. The creation of data-sharing networks, such as the International Network for Food and Obesity/NCDs Research, Monitoring, and Action Support (INFORMAS), is effectively increasing collaboration between governments and NGOs in Europe and parts of Asia (Swinburn et al., 2013), but must be expanded to LMICs in the coming years to accelerate progress toward worldwide obesity-reduction targets.

Research has also contributed to the rise of awareness and advocacy to reduce the prevalence of obesity. Similar to the foundational *The Lancet* series on maternal undernutrition, two *The Lancet* series on obesity were published in 2011 and 2015. The first series focused on identifying the drivers of the obesity pandemic, economic and health burdens, and weight physiology. The second *The Lancet* series explored developments in effective policies, systems, and programs to reduce "obesogenic" food and built environments—the physical spaces, cultural traditions, and taste preferences that inhibit individuals from making nutritious diet choices and, therefore, contribute to overweight and obesity. Nutrition-sensitive interventions in obesogenic environments are ongoing, with proponents attempting to identify context-specific and culturally salient mechanisms with which to encourage nutritious food choices and initiate substantial reductions in obesity prevalence.

operationalization, accountability, and scaling-up of effective interventions from the local level to the national level.

Last, and relatedly, policy makers often scrutinize the private sector, unsure of their role in the fight against malnutrition. Despite this skepticism, NGOs and civil society organizations often work in tandem with public agencies to research and implement nutrition interventions, particularly in LMICs.

EXHIBIT 7-4 provides an overview of recent research that suggests strengthened accountability and transparency of systems is one of the key components in the continuing fight to eliminate overweight, obesity, and diet-related NCDs (Swinburn et al., 2015).

Discussion Questions

1. The same population may be affected by multiple micronutrient deficiencies, although people at different life stages and certain socioeconomic or geographic groups may be more vulnerable to individual nutrient deficiencies. Discuss ways in which single-micronutrient deficiency prevention strategies might be combined, integrated, or coordinated to achieve cost-effective control.
2. A lower-middle-income country may be afflicted with both high rates of childhood undernutrition and a rising epidemic of obesity among its lower socioeconomic groups while undergoing a nutrition transition. Discuss the challenges that this situation poses for the country's government in developing food and nutrition policies as well as national dietary guidelines for healthy eating.
3. The Ministry of Health of a lower-income country has decided to institute a national nutrition surveillance system to monitor the country's most pressing child and maternal nutrition problems. Discuss the kinds of nutritional problems, target groups, approaches to assessment, options for routine contact in the community, and types of agencies to organize into a surveillance system to guide the ministry.

References

Abbeddou, S., Yakes Jimenez, E., Some, J. W., Ouedraogo, J. B., Brown, K. H., & Hess, S. Y. (2017). Small-quantity lipid-based nutrient supplements containing different amounts of zinc along with diarrhea and malaria treatment increase iron and vitamin A status and reduce anemia prevalence, but do not affect zinc status in young Burkinabe children: A cluster-randomized trial. *BMC Pediatrics, 17*(1), 46.

Aboud, F. E., & Yousafzai, A. K. (2016). Very early childhood development. In R. E. Black, R. Laxminarayan, M. Temmerman, & N. Walker (Eds.), *Disease control priorities, reproductive, maternal, newborn, and child health: Child and adolescent health and development* (3rd ed., Vol. 2, Chapter 13). Washington, DC: World Bank.

Acosta Mejia, A., & Fanzo, J. (2012). *Analysing the political and institutional determinants of delivering a national multisectoral response in six countries.* Brighton, UK: Institute of Development Studies. Retrieved from https://www.ids.ac.uk/files/dmfile/DFID_ANG_Synthesis_April2012.pdf

Adair, L., Popkin, B. M., Van Derslice, J., Akin, J., Guilkey, D., Black, R., . . . Flieger, W. (1993). Growth dynamics during the first two years of life: A prospective study in the Philippines. *European Journal of Clinical Nutrition, 47*, 42–51.

Administrative Committee on Coordination/Sub-Committee on Nutrition (ACC/SCN). (2000). *Fourth report on the world nutrition situation: Nutrition throughout the life cycle.* Washington, DC: International Food Policy Research Institute. Retrieved from http://www.ifpri.org/publication/4th-report-world-nutrition-situation

Adu-Afarwuah, S., Lartey, A., Brown, K. H., Zlotkin, S., Briend, A., & Dewey, K. G. (2008). Home fortification of complementary foods with micronutrient supplements is well accepted and has positive effects on infant iron status in Ghana. *American Journal of Clinical Nutrition, 87*(4), 929–938.

Afshin, A., Reitsma, M. B., & Murray, C. J. (2017). Health Effects of Overweight and Obesity in 195 Countries. *The New England Journal of Medicine*, *377*(15), 1496–1497.

Aggarwal, R., Sentz, J., & Miller, M.A. (2007). Role of zinc administration in prevention of childhood diarrhea and respiratory illnesses: A meta-analysis. *Pediatrics, 119*, 1120–1130.

Akresh, R., de Walque, D., & Kazianga, H. (2013). Cash transfers and child schooling: Evidence from a randomized evaluation of the role of conditionality (No. 6340). Washington, DC: World Bank. Retrieved from http://documents.worldbank.org/curated/en/587731468005971189/pdf/wps6340.pdf

Alderman, H. (2010). Safety nets can help address the risks to nutrition from increasing climate variability. *Journal of Nutrition*, *140*(1), 148S–152S.

Alderman, H., & Headey, D. D. (2014). *The nutritional returns to parental education.* IFPRI Discussion Paper 01379. Washington, DC: International Food Policy Research Institute. Retrieved from http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/128434

Alexander, G. R., Himes, J. H., Kaufman, R. B., Mor, J., & Kogan, M. (1996). A United States national reference for fetal growth. *Obstetrics and Gynecology, 87*(2), 163–168.

Al-Khudairy, L., Loveman, E., Colquitt, J. L., Mead, E., Johnson, R. E., Fraser, H., . . . Rees, K. (2017). Diet, physical activity and behavioural interventions for the treatment of overweight or obese adolescents aged 12 to 17 years. *Cochrane Database of Systematic Reviews,* CD012691. doi:10.1002/14651858.CD012691.pub1

Allen, L. H. (1997). Pregnancy and iron deficiency: Unresolved issues. *Nutrition Reviews, 55*(4), 91–101.

Allen, L. H. (2012). B vitamins in breast milk: Relative importance of maternal status and intake, and effects on infant status and function. *Advances in Nutrition, 3*(3), 362–369.

Al Tanoury, Z., Piskunov, A., & Rochette-Egly, C. (2013) Vitamin A and retinoid signaling: Genomic and nongenomic effects. *Journal of Lipid Research, 54,* 1761–1775.

Ames, B. N. (2010). Optimal micronutrients delay mitochondrial decay and age-associated diseases. *Mechanisms of Ageing and Development, 131,* 473–479.

Amour, C., Gratz, J., Mduma, E., Svensen, E., Rogawski, E. T., McGrath, M., . . . Platts-Mills, J. A. (2016). Epidemiology and impact of *Campylobacter* infection in children in 8 low-resource settings: Results from the MAL-ED Study. *Clinical Infectious Diseases*, *63*(9), 1171–1179.

Andersson, M., Karumbunathan, V., & Zimmermann, M. B. (2012). Global iodine status in 2011 and trends over the past decade. *Journal of Nutrition, 142*(4), 744–750.

Aoki, H., Kieu, N. T., Kuze, N., Tomisaka, K., & Van Chuyen, N. (2002). Carotenoid pigments in GAC fruit (Momordica cochinchinensis SPRENG). *Bioscience, Biotechnology, and Biochemistry, 66*, 2479–2482.

Apgar, J. (1985). Zinc and reproduction. *Annual Review of Nutrition, 5*, 43–68.

Arimond, M., Wiesmann, D., Becquey E., Carriquiry, A., Daniels, M. C., Deitchler, M., . . . Torheim, L. E. (2010). Simple food group diversity indicators predict micronutrient adequacy of women's diets in 5 diverse, resource-poor settings. *Journal of Nutrition, 140*(11), 2059S. Retrieved from https://search.proquest.com/docview/762213708

Arimond, M., Zeilani, M., Jungjohann, S., Brown, K. H., Ashorn, P., Allen, L. H., & Dewey, K. G. (2015). Considerations in developing lipid-based nutrient supplements for prevention of undernutrition: Experience from the International Lipid-Based Nutrient Supplements (iLiNS) Project. *Maternal and Child Nutrition, 11*(suppl 4), 31–61.

Arora, M., Chauhan, K., John, S., & Mukhopadhyay, A. (2011). Multi-sectoral action for addressing social determinants of noncommunicable diseases and mainstreaming health promotion in national health programmes in India. *Indian Journal of Community Medicine, 36*(5), 43–49.

Arsenault, J. E., & Brown, K. H. (2017). Dietary protein intake in young children in selected low-income countries is generally adequate in relation to estimated requirements for healthy children, except when complementary food intake is low. *Journal of Nutrition, 147*(5), 932–939.

Arsenault, J. E., Hijmans, R. J., & Brown, K. H. (2015). Improving nutrition security through agriculture: An analytical framework based on national food balance sheets to estimate nutritional adequacy of food supplies. *Food Security, 7*(3), 693–707.

Arthur, P., Kirkwood, B., Ross, D., Morris, S., Gyapong, J., Tomkins, A., & Addy, H. (1992). Impact of vitamin A supplementation on childhood morbidity in northern Ghana. *The Lancet, 339*, 361–362.

Ashworth, A. (1998). Effects of intrauterine growth retardation on mortality and morbidity in infants and young children. *European Journal of Clinical Nutrition, 52*, S34–S42.

Ashworth, A., & Feachem, R. G. (1985). Interventions for the control of diarrhoeal diseases among young children:

Prevention of low birthweight. *Bulletin of the World Health Organization, 63*, 165–184.

Awasthi, S., Peto, R., Read, S., Clark, S., Pande, V., Bundy. D., & DEVTA (Deworming and Enhanced Vitamin A) Team. (2013). Vitamin A supplementation every 6 months with retinol in 1 million pre-school children in north India: DEVTA, a cluster-randomised trial. *The Lancet*, *381*, 1469–1477.

Baker, P., & Friel, S. (2014). Processed foods and the nutrition transition: Evidence from Asia. *Obesity Reviews, 15*(7), 564–577.

Balarajan, Y., & Reich, M. R. (2016). Political economy challenges in nutrition. *Globalization and Health, 12*(1), 70.

Ballard, O., & Morrow, A. L. (2013). Human milk composition: Nutrients and bioactive factors. *Pediatric Clinics of North America, 60*(1), 49–74.

Banajeh, S. M. (2003). Is 12-monthly vitamin A supplementation of preschool children effective? An observational study of mortality rates for severe dehydrating diarrhea in Yemen. *South African Journal of Clinical Nutrition*, *16*, 137–142.

Baqui, A. H., Black, R. E., El Arifeen, S., Yunus, M., Chakraborty, J., Ahmed, S., & Vaughan, J. P. (2002). Effect of zinc supplementation started during diarrhea on morbidity and mortality in Bangladeshi children: Community randomized trial. *British Medical Journal, 325*, 1059–1065.

Baqui, A. H., Sack, R. B., Black, R. E., Chowdhury, H. R., Yunus, M., & Siddique, A. K. (1993). Cell-mediated immune deficiency and malnutrition are independent risk factors for persistent diarrhea in Bangladeshi children. *American Journal of Clinical Nutrition, 58*, 543–548.

Barber, T., Esteban-Pretel, G., Marin, M. P., & Timoneda, J. (2014) Vitamin A deficiency and alterations in the extracellular matrix. *Nutrients*, *6*, 4984–5017.

Barclay, A. J. G., Foster, A., & Sommer, A. (1987). Vitamin A supplements and mortality related to measles: A randomised clinical trial. *British Medical Journal*, *294,* 294–296.

Barker, D. J. P. (1998). *Mothers, babies and health in later life* (Vol. 2). Edinburgh, UK: Churchill Livingstone.

Bauserman, M., Lokangaka, A., Gado, J., Close, K., Wallace, D., Kodondi, K. K., & Bose, C. (2015). A cluster-randomized trial determining the efficacy of caterpillar cereal as a locally available and sustainable complementary food to prevent stunting and anaemia. *Public Health Nutrition, 18*(10), 1785–1792.

Beach, R. S., Gershwin, M. E., & Hurley, L. S. (1982). Gestational zinc deprivation in mice: Persistence of immunodeficiency for three generations. *Science, 218*, 469–471.

Beard, J. L. (2001). Iron biology in immune function, muscle metabolism and neuronal functioning. *Journal of Nutrition, 131*(2S-2), 568S–579S.

Beard, J. L. (2008). Why iron deficiency is important in infant development. *Journal of Nutrition, 138*(12), 2534–2536.

Beaton, G. H., Martorell, R., Aronson, K. J., Edmonston, B., McCabe, G., Ross, A. C., & Harvey, B. (1993). *Effectiveness of vitamin A supplementation in the control of young child morbidity and mortality in developing countries* (ACC/SCN State of the Art Series, Nutrition Policy Discussion Paper No. 13). Geneva, Switzerland: Administrative Committee on Coordination/Subcommittee on Nutrition. Retrieved from https://www.unscn.org/web/archives_resources/files/Policy_paper_No_13.pdf

Behrens, R. H., Lunn, P. G., Northrop, C. A., Hanlon, P. W., & Neale, G. (1987). Factors affecting the integrity of the intestinal mucosa of Gambian children. *American Journal of Clinical Nutrition, 45*, 1433–1441.

Beijer, M. R., Kraal, G., & den Haan, J. M. M. (2013). Vitamin A and dendritic cell differentiation. *Immunology*, *142*, 39–45.

Benn, C. S., Diness, B. R., Roth, A., Nante, E., Fisker, A. B., Lisse, I. M., . . . Aaby, P. (2008). Effect of 50,000 IU vitamin A given with BCG vaccine on mortality in infants in Guinea-Bissau: Randomized placebo controlled trial. *British Medical Journal, 336*, 1416–1420.

Benn, C. S., Martins, C. L., Fisker, A. B., Diness, B. R., Garly, M. L., Balde, I., . . . Aaby, P. (2014). Interaction between neonatal vitamin A supplementation and timing of measles vaccination: A retrospective analysis of three randomized trials from Guinea-Bissau. *Vaccine, 32*, 5468–5474.

Benn, C., Fisker, A. B., Napirna, B. M., Roth, A., Diness, B. R., Lausch, K. R., . . . Aaby, P. (2010). Vitamin A supplementation and BCG vaccination at birth in low birth weight neonates: Two by two randomised factorial trial. *British Medical Journal, 340.* doi:10.1136/bmj.c1101

Bhandari, N., Bahl, R., Mazumdar, S., Martines, J., Black, R. E., Bhan, M. K., & Infant Feeding Study Group. (2003). Effect of community-based promotion of exclusive breastfeeding on diarrhoeal illness and growth: A cluster randomised controlled trial. *The Lancet, 361*(9367), 1418–1423.

Bhandari, N., Bhan, M. K., & Sazawal, S. (1992). Mortality associated with acute watery diarrhea, dysentery, and persistent diarrhea in rural north India. *Acta Paediatrica Supplement, 381*, 3–6.

Bhandari, N., Mazumder, S., Taneja, S., Dube, B., Agarwal, R. C., Mahalanabis, D., . . . Bhan, M. K. (2008). Effectiveness of zinc supplementation plus oral rehydration salts compared with oral rehydration salts alone as a treatment for acute diarrhea in a primary care setting: A cluster randomized trial. *Pediatrics, 121*, e1279–e1285.

Bhandari, N., Mohan, S. B., Bose, A., Iyengar, S. D., Taneja, S., Mazumder, S., . . . Bahl, R. (2016). Efficacy of three feeding regimens for home-based management of children with uncomplicated severe acute malnutrition: A randomised trial in India. *BMJ Global Health, 1*(4), e000144.

Bhutta, Z. A., Das, J. K., Rizvi, A., Gaffey, M. F., Walker, N., Horton, S., & Black, R. E. (2013). Evidence-based interventions for improvement of maternal and child nutrition: What can be done and at what cost? *The Lancet, 382*(9890), 452–477.

Biesalski, H. K. (2016). Nutrition meets the microbiome: Micronutrients and the microbiota. *Annals of the New York Academy of Sciences*, *1372*, 53–64.

Black, M. M., & Aboud, F. E. (2011). Responsive feeding is embedded in a theoretical framework of responsive parenting. *Journal of Nutrition, 141*(3), 490–494.

Black, M. M., Gove, A., & Merseth, K. A. (2017). Platforms to reach children in early childhood. In D. A. P. Bundy, N. de Silva, S. Horton, D. T. Jamison, & G. C. Patton (Eds.), *Disease control priorities: Child and adolescent health and development* (3rd ed., Vol. 8, chapter 19). Washington, DC: World Bank.

Black, M. M., & Hurley, K. M. (2017). Responsive feeding: Strategies to promote healthy mealtime interactions. In R. E. Black, M. Makrides, & K. K. Ong, *Complementary feeding: Building the foundations for a healthy life* (Vol. 87, pp. 153–165). Singapore: Nestlé Nutrition Institute Workshop Series.

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., & *The Lancet* Early Childhood Development Series Steering Committee. (2017). Early childhood development coming of age: Science through the life course. *The Lancet, 389*(10064), 77–90.

Black, R. E. (1993). Persistent diarrhea in children of developing countries. *Pediatric Infectious Diseases Journal, 12*, 751–761.

Black, R. E. (1997). *Zinc for child health: Child Health Research Project special report.* Baltimore, MD: Johns Hopkins University. Retrieved from https://childhealthresearch.org/doc/spec1.pdf

Black, R. E., Allen, L. H., Bhutta, Z. A., Caulfield, L. E., de Onis, M., Ezzati, M., & Child Undernutrition Study Group. (2008). Maternal and child undernutrition: Global and regional exposures and health consequences. *The Lancet, 371*(9608), 243–260.

Black, R. E., Brown, K. H., & Becker, S. (1983). Influence of acute diarrhea on the growth parameters of children. In J. A. Bellanti (Ed.), *Acute diarrhea: Its nutritional consequences* (pp. 75–84). New York, NY: Vevey-Raven Press.

Black, R. E., Brown, K. H., & Becker, S. (1984a). Malnutrition is a determining factor in diarrheal duration, but not incidence, among young children in a longitudinal study in rural Bangladesh. *American Journal of Clinical Nutrition, 39*, 87–94.

Black, R. E., Brown, K. H., & Becker, S. (1984b). Effects of diarrhea associated with specific enteropathogens on the growth of children in rural Bangladesh. *Pediatrics, 73*, 799–805.

Black, R. E., Victora, C. G., Walker, S. P., Bhutta, Z. A., Christian, P., de Onis, M. . . . Uauy, R. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet, 382*(9890), 427–451.

Blencowe, H., Cousens, S., Modell, B., & Lawn, J. (2010). Folic acid to reduce neonatal mortality from neural tube disorders. *International Journal of Epidemiology, 39*(suppl 1), i110–i121.

Blomhoff, R., & Blomhoff, H. K. (2006). Overview of retinoid metabolism and function. *Journal of Neurobiology, 66*, 606–630.

Bobat, R., Coovadia, H., Stephen, C., Naidoo, K. L., McKerrow, N., Black, R. E., & Moss, W. J. (2005). Safety and efficacy of zinc supplementation for children with HIV-1 infection in South Africa: A randomized double-blind placebo-controlled trial. *The Lancet, 366*, 1862–1867.

Bode, L. (2015). The functional biology of human milk oligosaccharides. *Early Human Development, 91*(11), 619–622.

Boschi-Pinto, C., Dillip, T. R., & Costello, A. (2016). Association between community management of pneumonia and diarrhoea in high-burden countries and the decline in under-five mortality rates: An ecological analysis. *BMJ Open, 7*(2), e012639.

Bose, A., Coles, C. L., Gunavathi, J. H., Moses, P., Raghupathy, P., Kirubakaran, C., . . . Santosham, M. (2006). Efficacy of zinc in the treatment of severe pneumonia in hospitalized children < 2 y old. *American Journal of Clinical Nutrition, 83*, 1089–1096.

Bougma, K., Aboud, F. E., Harding, K. B., & Marquis, G. S. (2013). Iodine and mental development of children 5 years old and under: a systematic review and meta-analysis. *Nutrients, 5*(4), 1384–1416.

Bouis, H. E., & Saltzman, A. (2017). Improving nutrition through biofortification: A review of evidence from HarvestPlus, 2003 through 2016. *Global Food Security, 12*, 49–58.

Boyages, S. C. (1994). The damaged brain of iodine deficiency: Evidence for a continuum of effect on the population at risk. In J. B. Stanbury (Ed.), *The damaged brain of iodine deficiency* (pp. 251–257). New York, NY: Cognizant Communication.

Brooks, W. A., Yunus, M., Santosham, M., Wahed, M. A., Nahar, K., Yeasmin, S., & Black, R. E. (2004). Zinc for severe pneumonia in very young children: Double-blind placebo-controlled trial. *The Lancet, 363*, 1683–1688.

Brossaud, J., Pallet, V., & Corcuff, J. B. (2017). Vitamin A, endocrine tissues and hormones: Interplay and interactions. *Endocrine Connections*. pii: EC-17-0101. doi:10.1530/EC-17-0101

Brown, K. H., Black, R. E., Robertson, A. D., & Becker, S. (1985). Effects of season and illness on the dietary intake of weanlings during longitudinal studies in rural Bangladesh. *American Journal of Clinical Nutrition, 41*, 343–355.

Brown, K. H., Dewey, K. G., & Allen, L. H. (1998). *Complementary feeding of young children in developing countries: A review of current scientific knowledge*. Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/nutrition/publications/infantfeeding/WHO_NUT_98.1/en/

Brown, K. H., Engle-Stone, R., Kagin, J., Rettig, E., & Vosti, S. A. (2015). Use of optimization modeling for selecting national micronutrient intervention strategies: An example based on potential programs for control of vitamin A deficiency in Cameroon. *Food and Nutrition Bulletin, 36*(3 suppl), S141–S148.

Brown, K. H., Gastanaduy, A. S., Saavedra, J. M., Lembcke, J., Rivas, D., Robertson, A. D., . . . Sack, R. B. (1988). Effect of continued oral feeding on clinical and nutritional outcomes of acute diarrhea in children. *Journal of Pediatrics, 112*, 191–200.

Brown, K. H., Gilman, R. H., Gaffar, A., Alamgir, S. M., Strife, J. L., Kapikian, A. Z., & Sack, R. B. (1981). Infections associated with severe protein-calorie malnutrition in hospitalized infants and children. *Nutrition Research, 1*(1), 33–46.

Brown, K. H., Peerson, J. M., Rivera, J., & Allen, L. H. (2002). Effect of supplemental zinc on the growth and serum zinc concentration of prepubertal children: A meta-analysis of randomized controlled trials. *American Journal of Clinical Nutrition, 75*, 1062–1071.

Bukenya, G. B., Barnes, T., & Nwokolo, N. (1991). Low birthweight and acute childhood diarrhoea: Evidence of their association in an urban settlement of Papua New Guinea. *Annals of Tropical Paediatrics, 11*, 357–362.

Bunce, G. E., Lytton, F., Gunesekera, B., Vessal, M., & Kim, C. (1994). Molecular basis for abnormal parturition in zinc deficiency in rats. In L. Allen, J. King, & B. Lonnerdal (Eds.), *Nutrient regulation during pregnancy, lactation, and infant growth* (pp. 209–214). New York, NY: Plenum Press.

Bundy, D., & Alderman, H. (2012). School feeding programs and development: Are we framing the question correctly? *World Bank Research Observer, 27*(2), 204–221.

Burchi, F. (2010). Child nutrition in Mozambique in 2003: The role of mother's schooling and nutrition knowledge. *Economics and Human Biology, 8*(3), 331–345.

Buse, K., & Tanaka, S. (2011). Global public–private health partnerships: Lessons learned from ten years of experience and evaluation. *International Dental Journal, 61*, 2–10.

Campbell, D. I., Elia, M., & Lunn, P. G. (2003). Growth faltering in rural Gambian infants is associated with impaired small intestinal barrier function, leading to endotoxemia and systemic inflammation. *Journal of Nutrition, 133,* 1332–1338.

Cao, C., & O'Brien, K. O. (2013). Pregnancy and iron homeostasis: An update. *Nutrition Reviews, 71*(1), 35–51.

Casanovas, M. D. C., Lutter, C. K., Mangasaryan, N., Mwadime, R., Hajeebhoy, N., Aguilar, A. M., & Onywango, A. W. (2013). Multi-sectoral interventions for healthy growth. *The Lancet, 9*(suppl 2), 46–57.

Caspi, C. E., Sorensen, G., Subramanian, S. V., & Kawachi, I. (2012). The local food environment and diet: A systematic review. *Health & Place, 18*(5), 1172–1187.

Caulfield, L. E., de Onis, M., Blossner, M., & Black, R. E. (2004). Undernutrition as an underlying cause of child deaths associated with diarrhea, pneumonia, malaria, and measles. *American Journal of Clinical Nutrition, 80*, 193–198.

Caulfield, L. E., Zavaleta, N., Shankar, A., & Merialdi, M. (1998). Potential contribution of maternal zinc supplementation during pregnancy for maternal and child survival. *American Journal of Clinical Nutrition, 68*, S499–S508.

Cerqueiro, M. C., Murtagh, P., Halac, A., Avila, M., & Weissenbacher, M. (1990). Epidemiologic risk factors for children with acute lower respiratory tract infection in Buenos Aires, Argentina: A matched case-control study. *Reviews of Infectious Diseases, 12*, S1021–S1028.

Chaker, L., Bianco, A. C., Jonklaas, J., & Peeters, R. P. (2017). Hypothyroidism. *The Lancet, 390*(10101), 1550–1562.

Chaparro, C. M., & Dewey, K. G. (2010). Use of lipid-based nutrient supplements (LNS) to improve the nutrient adequacy of general food distribution rations for vulnerable sub-groups in emergency settings. *Maternal & Child Nutrition, 6*(suppl 1), 1–69.

Checkley, W., Buckley, G., Gilman, R. H., Assis, A. M. O., Guerrant, R. L., Morris, S. S., . . . Black, R. E. (2008). Multi-country analysis of the effects of diarrhea on childhood stunting. *International Journal of Epidemiology, 37*, 816–830.

Checkley, W., Epstein, L. D., Gilman, R. H., Black, R. E., Cabrera, L., & Sterling, C. R. (1998). Effects of *Cryptosporidium parvum* infection in Peruvian children: Growth faltering and subsequent catch-up growth. *American Journal of Epidemiology, 148*, 497–506.

Checkley, W., Epstein, L. D., Gilman, R. H., Cabrera, L., & Black, R. E. (2003). Effects of acute diarrhea on linear growth in Peruvian children. *American Journal of Epidemiology, 157*, 166–175.

Checkley, W., Gilman, R. H., Black, R. E., Lescano, A. G., Cabrera, L., Taylor, D. N., & Moulton, L. H. (2002). Effects of nutritional status on diarrhea in Peruvian children. *Journal of Pediatrics, 140*, 210–218.

Checkley, W., West, K. P., Jr., Wise, R. A., Baldwin, M. R., Wu, L., LeClerq, S. C., . . . Sommer, A. (2010). Maternal vitamin A supplementation and lung function in offspring. *New England Journal of Medicine, 362*, 1784–1794.

Chen, Y., Yu, S., & Li, W. (1988). Artificial feeding and hospitalization in the first 18 months of life. *Pediatrics, 81*, 58–62.

Christian, P. (2002). Recommendations for indicators: Night blindness during pregnancy: A simple tool to assess vitamin A deficiency in a population. *Journal of Nutrition*, *132*, 2884S–2888S.

Christian, P., Labrique, A. B., Ali, H., Richman, M., Wu, L., Rashid, M., & West, K. P., Jr. (2011). Maternal vitamin A and β-carotene supplementation and risk of bacterial vaginosis: A randomized controlled trial in rural Bangladesh. *American Journal of Clinical Nutrition*, *94*, 1643–1649.

Christian, P., Lee, S. E., Donahue Angel, M., Adair, L. S., Arifeen, S. E., Ashorn, P., & Black, R. E. (2013). Risk of childhood undernutrition related to small-for-gestational age and preterm birth in low- and middle-income countries. *International Journal of Epidemiology, 42*(5), 1340–1355.

Christian, P., Shaikh, S., Shamim, A. A., Mehra, S., Wu, L., Mitra, M., & West, K. P., Jr. (2015). Effect of fortified complementary food supplementation on child growth in rural Bangladesh: A cluster-randomized trial. *International Journal of Epidemiology, 44*(6), 1862–1876.

Christian, P., West, K. P., Jr., Khatry, S. K., Katz, J., Shrestha, S. R., Pradhan, E. K., . . . Pokhrel, R. P. (1998). Night blindness of pregnancy in rural Nepal: Nutritional and health risks. *International Journal of Epidemiology, 27*, 231–237.

Christian, P., West, K. P., Jr., Khatry, S. K., Kimbrough-Pradhan, E., LeClerq, S. C., Katz, J., . . . Sommer, A. (2000). Night blindness during pregnancy and subsequent mortality among women in Nepal: Effects of vitamin A and β-carotene supplementation. *American Journal of Epidemiology*, *152,* 542–547.

Chwang, L. C., Soemantri, A. G., & Pollitt, E. (1988). Iron supplementation and physical growth of rural Indonesian children. *American Journal of Clinical Nutrition, 47*(3), 496–501.

Clasen, T., Boisson, S., Routray, P., Torondel, B., Bell, M., Cumming, O., . . . Schmidt, W. P. (2014). Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: A cluster-randomised trial. *The Lancet Global Health, 2*(11), e645–e653.

Coffey, R., & Ganz, T. (2017). Iron homeostasis: An anthropocentric perspective. *Journal of Biological Chemistry, 292*(31), 12727–12734.

Cole, R. N., Ruczinski, I., Schulze, K., Christian, P., Herbrich, S., Wu, L., . . . West, K. P., Jr. (2013). The plasma proteome identifies expected and novel proteins correlated with micronutrient status in undernourished Nepalese children. *Journal of Nutrition*, *143*, 1540–1548.

Cole, T. J., & Parkin, J. M. (1977). Infection and its effect on the growth of young children: A comparison of the Gambia and Uganda. *Transactions of the Royal Society for Tropical Medicine and Hygiene, 71*, 196–198.

Coles, C. L., Rahmathullah, L., Kanungo, R., Thulasiraj, R. D., Katz, J., Santhosham, M., & Tielsch, J. M. (2001). Vitamin A supplementation at birth delays pneumococcal colonization in South Indian infants. *Journal of Nutrition*, *131*, 255–261.

Collins, S., Dent, N., Binns, P., Bahwere, P., Sadler, K., & Hallam, A. (2006). Management of severe acute malnutrition in children. *The Lancet, 368*(9551), 1992–2000.

Condon-Paoloni, D., Joaquin, C., Johnston, F. E., deLicardi, E. R., & Scholl, T. O. (1977). Morbidity and growth of infants and young children in a rural Mexican village. *American Journal of Public Health, 67*, 651–656.

Cordain, L., Eaton, S. B., Sebastian, A., Mann, N., Lindeberg, S., Watkins, B. A, & Brand-Miller, J. (2005). Origins and evolution of the Western diet: Health implications for the 21st century. *American Journal of Clinical Nutrition, 81*(2), 341. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15699220

Cousins, R. J., & Hempe, J. M. (1990). Zinc. In M. L. Brown (Ed.), *Present knowledge in nutrition* (6th ed., pp. 251–260). Washington, DC: ILSI Press.

Coutsoudis, A., Broughton, M., & Coovadia, H. M. (1991). Vitamin A supplementation reduces measles morbidity in young African children: A randomized, placebo-controlled, double-blind trial. *American Journal of Clinical Nutrition, 54*, 890–895.

Cruz, J. R., Pareja, G., de Fernandez, A., Peralta, F., Caceres, P., & Cano, F. (1990). Epidemiology of acute respiratory tract infections among Guatemalan ambulatory preschool children. *Reviews of Infectious Diseases, 12*, S1029–S1034.

Cunningham, K., Ruel, M., Ferguson, E., & Uauy, R. (2015). Women's empowerment and child nutritional status in South Asia: A synthesis of the literature. *Maternal & Child Nutrition, 11*(1), 1–19.

Dangour, A. D., Watson, L., Cumming, O., Boisson, S., Che, Y., Velleman, Y., & Uauy, R. (2013). Interventions to improve water quality and supply, sanitation and hygiene practices, and their

effects on the nutritional status of children (review). *Cochrane Database of Systematic Reviews, 8*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD009382.pub2/abstract;jsessionid=269B609B93B15ADEE13C61167EBB23DC.f01t04

Debes, A. K., Kohli, A., Walker, N., Edmond, K., & Mullany, L. C. (2013). Time to initiation of breastfeeding and neonatal mortality and morbidity: A systematic review. *BMC Public Health, 13*(suppl 3), S19.

Delange, F. (1994). The disorders induced by iodine deficiency. *Thyroid, 4*(1), 107–128.

DeLong, R., Tai, M., Xue-Yi, C., Xin-Min, J., Zhi-Hong, D., & Rakeman, M. A. (1994). The neuromotor deficit in endemic cretinism. In J. B. Stanbury (Ed.), *The damaged brain of iodine deficiency* (pp. 9–13). New York, NY: Cognizant Communication.

de Pee, S., Bloem, M. W., Satoto, T., Yip, R., Sukaton, A., Tjiong, R., . . . Kodyat, B. (1998). Impact of a social marketing campaign promoting dark-green leafy vegetables and eggs in Central Java, Indonesia. *International Journal of Vitamin and Nutrition Research, 68*, 389–398.

de Pee, S., Grais, R., Fenn, B., Brown, R., Briend, A., Frize, J., Shoham, J., & Kiess, L. (2015). Prevention of acute malnutrition: Distribution of special nutritious foods and cash, and addressing underlying causes: What to recommend, when, where, for whom, and how. *Food and Nutrition Bulletin, 36*(suppl 1), S24–S29.

De-Regil, L. M., Suchdev, P. S., Vist, G. E., Walleser, S., & Pena-Rosas, J. P. (2011). Home fortification of foods with multiple micronutrient powders for health and nutrition in children under two years of age. *Cochrane Database of Systematic Reviews, 9*, CD008959. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD008959.pub2/abstract

Dewey, K. G. (2013). The challenge of meeting nutrient needs of infants and young children during the period of complementary feeding: An evolutionary perspective. *Journal of Nutrition, 143*(12), 2050–2054.

Dewey, K. G., & Adu-Afarwuah, S. (2008). Systematic review of the efficacy and effectiveness of complementary feeding interventions in developing countries. *Maternal and Child Nutrition, 4*(suppl 1), 24–85.

Dewey, K. G., & Brown, K. H. (2003). Update on technical issues concerning complementary feeding of young children in developing countries and implications for intervention programs. *Food Nutrition Bulletin, 24*(1), 5–28.

Dewey, K. G., Mridha, M. K., Matias, S. L., Arnold, C. D., Cummins, J. R., Khan, M. S., & Vosti, S. A. (2017). Lipid-based nutrient supplementation in the first 1000 d improves child growth in Bangladesh: A cluster-randomized effectiveness trial. *American Journal of Clinical Nutrition, 105*(4), 944–957.

Dietz, W. H., Baur, L. A., Hall, K., Puhl, R. M., Taveras, E. M., Uauy, R., & Kopelman, P. (2015). Management of obesity: Improvement of health-care training and systems for prevention and care. *The Lancet*, *385*(9986), 2521–2533.

Di Giovanni, V., Bourdon, C., Wang, D. X., Seshadri, S., Senga, E., Versloot, C. J., & Bandsma, R. H. (2016). Metabolomic changes in serum of children with different clinical diagnoses of malnutrition. *Journal of Nutrition, 146*(12), 2436–2444.

Drakesmith, H., & Prentice, A. M. (2012). Hepcidin and the iron-infection axis. *Science, 338*(6108), 768–772.

Drewnowski, A., & Popkin, B. M. (1997). The nutrition transition: New trends in the global diet. *Nutrition Reviews, 55*(2), 31–43.

Eastman, C. J., & Zimmermann, M. B. (2017). The iodine deficiency disorders. In L. J. De Groot, G. Chrousos, K. Dungan, K. R. Feingold, A. Grossman, J. M. Hershman, . . . A. Vinik (Eds.), *Endotext*. South Dartmouth, MA: MDText.com, Inc.; 2000-. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK285556/.

Edmond, K. M., Newton, S., Shannon, C., O'Leary, M., Hurt, L., Thomas, G., . . . Kirkwood, B. R. (2014). Effect of early neonatal vitamin A supplementation on mortality during infancy in Ghana (Neovita): A randomized, double-blind, placebo-controlled trial. *The Lancet*, *385*, 1315–1323.

Efevbera, Y., Farmer, P. E., Bhabha, J., & Fink, G. (2017). Girl child marriage as a risk factor for early childhood development and stunting. *Social Science & Medicine, 185,* 91.

Elliot, V., Lutter, C., Lamstein, S., Koniz-Booher, P., & Caulfield, L. (2015). Systematic review of the dietary intakes of adolescent girls in low- and middle-income countries. *FASEB Journal, 29*(1), 898.4.

Ells, L. J., Hancock, C., Copley, V. R., Mead, E., Dinsdale, H., Kinra, S., & Rutter, H. (2015). Prevalence of severe childhood obesity in England: 2006–2013. *Archives of Disease in Childhood, 100*(7), 631–636.

El Samani, F. Z., Willett, W. C., & Ware, J. H. (1986). Predictors of simple diarrhoea disease surveillance in a rural Ghanaian pre-school child population. *Transactions of the Royal Society for Tropical Medicine and Hygiene, 80*, 208–213.

Engleberger, L., Aalbersberg, W., Dolodolotawake, U., Schierle, J., Humphries, J., Iuta, T., . . . Kaiririete, M. (2006). Carotenoid content of pandanus fruit cultivars and other foods of the Republic of Kiribati. *Public Health Nutrition, 9*, 631–643.

Engle-Stone, R., Aaron, G. J., Huang, J., Wirth, J. P., Namaste, S. M., Williams, A. M., & Suchdev, P. S. (2017). Predictors of anemia in preschool children: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *American Journal of Clinical Nutrition, 106*(suppl 1), 402S–415S.

Engle-Stone, R., Williams, T. N., Nankap, M., Ndjebayi, A., Gimou, M. M., Oyono, Y., & Green, R. (2017). Prevalence of inherited hemoglobin disorders and relationships with anemia and micronutrient status among children in Yaounde and Douala, Cameroon. *Nutrients, 9*(7). doi:10.3390/nu9070693

Ezzati, M., Obermeyer, Z., Tzoulaki, I., Mayosi, B. M., Elliott, P., & Leon, D. A. (2015). The contributions of risk factor trends and medical care to cardiovascular mortality trends. *Nature Reviews Cardiology*, *12*(9), 508.

Ezzati, M., & Riboli, E. (2013). Behavioral and dietary risk factors for non-communicable diseases. *New England Journal of Medicine, 369*(10), 954–964.

Fawzi, W. W., Chalmers, T. C., Herrera, G., & Mosteller, F. (1993). Vitamin A supplementation and child mortality: A meta-analysis. *Journal of the American Medical Association, 269*, 898–903.

Fenech, M. F. (2010). Dietary reference values of individual micronutrients and nutriomes for genome damage prevention: Current status and a road map to the future. *American Journal of Clinical Nutrition, 91*, 1438S–1454S.

Fenton, R., Nyamukapa, C., Gregson, S., Robertson, L., Mushati, P., Thomas, R., & Eaton, J. W. (2016). Wealth differentials in the impact of conditional and unconditional cash transfers on education: Findings from a community-randomised controlled trial in Zimbabwe. *Psychology, Health & Medicine, 21*(8), 909–917.

Ferguson, E., Chege, P., Kimiywe, J., Wiesmann, D., & Hotz, C. (2015). Zinc, iron and calcium are major limiting nutrients in the complementary diets of rural Kenyan children. *Maternal & Child Nutrition, 11*(suppl 3), 6–20.

Finkelstein, J. L., Mehta, S., Udipi, S. A., Ghugre, P. S., Luna, S. V., Wenger, M. J., & Haas, J. D. (2015). A randomized trial of iron-biofortified pearl millet in school children in India. *Journal of Nutrition, 145*(7), 1576–1581.

Fishman, S., Caulfield, L. E., de Onis, M., Blossner, M., Hyder, A. A., Mullany, L., & Black, R. E. (2004). Childhood and maternal underweight. In M. Ezzati, A. D. Lopez, A. Rodgers, & C. J. L. Murray (Eds.), *Comparative quantification of health risks: Global and regional burden of disease attributable to selected major risk factors* (pp. 39–162). Geneva, Switzerland: World Health Organization.

Fishman, S. M., Christian, P., & West, K. P. (2000). The role of vitamins in the prevention and control of anaemia. *Public Health Nutrition, 3*(2), 125–150.

Fonseca, W., Kirkwood, B. R., Victora, C. G., Fuchs, S. R., Flores, J. A., & Misago, C. (1996). Risk factors for childhood pneumonia among the urban poor in Fortaleza, Brazil: A case-control study. *Bulletin of the World Health Organization, 74*, 199–208.

Food and Agriculture Organization (FAO). (2012). *Food security and school nutrition in Cabo Verde*. Rome, Italy: Food and Agriculture Organization of the United Nations. Retrieved from http://www.fao.org/3/a-i3934e.pdf

Food and Agriculture Organization (FAO). (2014). *Nutrition-sensitive agriculture: Second International Conference on Nutrition 19–21 November 2014.* Rome, Italy: Food and Agriculture Organization of the United Nations. Retrieved from http://www.fao.org/about/meetings/icn2/documents/en/

Food and Agriculture Organization (FAO). (2015). *Nutrition and social protection.* Rome, Italy: Food and Agriculture Organization of the United Nations. Retrieved from http://www.fao.org/3/a-i4819e.pdf

Food and Agriculture Organization (FAO). (2017a). *The state of food security and nutrition in the world 2017: Building resilience for peace and food security.* Rome, Italy: Food and Agriculture Organization of the United Nations. Retrieved from http://www.fao.org/3/a-I7695e.pdf

Food and Agriculture Organization (FAO). (2017b). *Nutrition-sensitive agriculture and food systems in practice: Options for intervention.* Rome, Italy: Food and Agriculture Organization of the United Nations. Retrieved from http://www.fao.org/3/a-i7848e.pdf

Forouzanfar, M. H., Alexander, L., Anderson, H. R., Bachman, V. F., Biryukov, S., Brauer, M., & Cohen, A. (2015). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: A systematic analysis for the global burden of disease study 2013. *The Lancet, 386*(10010), 2287–2323.

Foster, B. A., Farragher, J., Parker, P., & Sosa, E. T. (2015). Treatment interventions for early childhood obesity: A systematic review. *Academic Pediatrics*, *15*(4), 353–361.

Fraser, A. M., Brockert, J. E., & Ward, R. H. (1995). Association of young maternal age with adverse reproductive outcomes. *New England Journal of Medicine, 332*(17), 1113–1118.

Frassetto, L. A., Schloetter, M., Mietus-Synder, M., Morris, J. R. C., & Sebastian, A. (2009). Metabolic and physiologic improvements from consuming a paleolithic, hunter–gatherer type diet. *European Journal of Clinical Nutrition, 63*(8), 947–955.

Frongillo, E. A., Nguyen, P. H., Saha, K. K., Sanghvi, T., Afsana, K., Haque, R., & Menon, P. (2017). Large-scale behavior-change initiative for infant and young child feeding advanced language and motor development in a cluster-randomized program evaluation in Bangladesh. *Journal of Nutrition, 147*(2), 256–263.

Gaiha, R., & Young, T. (1989). On the relationship between share of starchy staples, calories consumed and income in selected developing countries. *Journal of International Development, 1*(3), 373–386.

Ganatra, B., & Faundes, A. (2016). Role of birth spacing, family planning services, safe abortion services and post-abortion care in reducing maternal mortality. *Best Practice & Research Clinical Obstetrics & Gynaecology, 36*, 145–155.

Gannon, B., Kaliwile, C., Arscott, S. A., Schmaelzle, S., Chileshe, J., Kalungwana, N., . . . Tanumihardjo, S. A. (2014). Biofortified orange maize is as efficacious as a vitamin A supplement in Zambian children even in the presence of high liver reserves of vitamin A: A community-based, randomized placebo-controlled trial. *American Journal of Clinical Nutrition, 100,* 1541–1550.

Garnham, P. C. C. (1954). Malaria in the African child. *East African Medical Journal, 31*, 155–159.

Gera, T., Sachdev, H. P., Nestel, P., & Sachdev, S. S. (2007). Effect of iron supplementation on haemoglobin response in children: Systematic review of randomised controlled trials. *Journal of Pediatric Gastroenterology and Nutrition, 44*(4), 468–486.

Gernand, A. D., Schulze, K. J., Stewart, C. P., West, K .P., Jr., & Christian, P. (2016). Micronutrient deficiencies in pregnancy worldwide: Health effects and prevention. *National Reviews of Endocrinology*, *12*, 274–289.

Ghuman, S. J., Lee, H. J., & Smith, H. L. (2006). Measurement of women's autonomy according to women and their husbands: Results from five Asian countries. *Social Science Research, 35*(1), 1–28.

Gibson, R. S. (2005). *Principles of nutritional assessment* (2nd ed.). New York, NY: Oxford University Press.

Gibson, R. S., & Ferguson, E. L. (1998a). Assessment of dietary zinc in a population. *American Journal of Clinical Nutrition, 68*, 430S–434S.

Gibson, R. S., & Ferguson, E. L. (1998b). Nutrition intervention strategies to combat zinc deficiency in developing countries. *Nutrition Research Reviews, 11*, 115–131.

Gibson, R. S., Heywood, A., Yaman, C., Sohlstrom, A., Thompson, L. U., & Heywood, P. (1991). Growth in children from the Wosera subdistrict, Papua New Guinea, in relation to energy and protein intakes and zinc status. *American Journal of Clinical Nutrition, 53*, 782–789.

Gibson, R. S., & Huddle, J. M. (1998). Suboptimal zinc status in pregnant Malawian women: Its association with low intakes of poorly available zinc, frequent reproductive cycling, and malaria. *American Journal of Clinical Nutrition, 67*, 702–709.

Gill, R., & Stewart, D. E. (2011). Relevance of gender-sensitive policies and general health indicators to compare the status of South Asian women's health. *Women's Health Issues, 21*(1), 12–18.

Gillespie, S., Haddad, L., Mannar, V., Menon, P., & Nisbett, N. (2013). The politics of reducing malnutrition: Building commitment and accelerating progress. *The Lancet, 382*(9891), 552–569.

Gillespie, S., Menon, P., & Kennedy, A. L. (2015). Scaling up impact on nutrition: What will it take? *Advances in Nutrition, 6*(4), 440–451.

Girard, A. W., Self, J. L., McAuliffe, C., & Olude, O. (2012). The effects of household food production strategies on the health and nutrition outcomes of women and young children: A systematic review. *Paediatric and Perinatal Epidemiology, 26*, 205–222.

Gittelsohn, J., & Vastine, A. E. (2003). Sociocultural and household factors impacting on the selection, allocation and consumption of animal source foods: Current knowledge and application. *Journal of Nutrition, 133*(11 suppl 2), 4036S–4041S.

Global Panel on Agriculture and Food Systems for Nutrition (GLOPAN). (2016). *Foresight report: Food systems and diets: Facing the challenges of the 21st century.* London, UK: Author. Retrieved from http://glopan.org/sites/default/files/ForesightReport.pdf

Global Panel on Agriculture and Food Systems for Nutrition (GLOPAN). (2017). *Improving nutrition through enhanced food environments*. Policy Brief No. 7. London, UK: Author. Retrieved from https://glopan.org/sites/default/files/Downloads/FoodEnvironmentsBrief.pdf

Golden, M. H. (2002). The development of concepts of malnutrition. *Journal of Nutrition, 132*(7), 2117S–2122S.

Goldstein, R. F., Abell, S. K., Ranasinha, S., Misso, M., Boyle, J. A., Black, M. H., & Teede, H. J. (2017). Association of gestational weight gain with maternal and infant outcomes: A systematic review and meta-analysis. *Journal of the American Medical Association, 317*(21), 2207–2225.

Goli, S., Rammohan, A., & Singh, D. (2015). The effect of early marriages and early childbearing on women's nutritional status in India. *Maternal and Child Health Journal, 19*(8), 1864–1880.

Graham, N. M. H. (1990). The epidemiology of acute respiratory infections in children and adults: A global perspective. *Epidemiologic Reviews, 12*, 149–178.

Greene, M. D., Kabaghe, G., Musonda, M., & Palmer, A. C. (2017). Retail sugar from one Zambian community does not meet statutory requirements for vitamin A fortification. *Food and Nutrition Bulletin, 38,* 594–598.

Guerrant, R. I., Schorling, J. B., McAuliffe, J. F., & de Souza, M. A. (1992). Diarrhea as a cause and an effect of malnutrition: Diarrhea prevents catch-up growth and malnutrition increases diarrhea frequency and duration. *American Journal of Tropical Medicine and Hygiene, 47*, 28–35.

Guldan, G. S., Fan, H. C., Ma, X., Ni, Z. Z., Xiang, X., & Tang, M. Z. (2000). Culturally appropriate nutrition education improves infant feeding and growth in rural Sichuan, China. *Journal of Nutrition, 130*(5), 1204–1211.

Haas, J. D., Luna, S. V., Lung'aho, M. G., Wenger, M. J., Murray-Kolb, L. E., Beebe, S., & Egli, I. M. (2016). Consuming iron biofortified beans increases iron status in Rwandan women after 128 days in a randomized controlled feeding trial. *Journal of Nutrition, 146*(8), 1586–1592.

Haddad, L. (2013). How can we build an enabling political environment to fight undernutrition? *European Journal of Development Research, 25*(1), 13–20.

Halpern, J. P. (1994). The neuromotor deficit in endemic cretinism and its implications for the pathogenesis of the disorder. In J. B. Stanbury (Ed.), *The damaged brain of iodine deficiency* (pp. 15–24). New York, NY: Cognizant Communication.

Hambidge, K. M., & Krebs, N. F. (1999). Zinc, diarrhea, and pneumonia. *Journal of Pediatrics, 135*, 661–664.

Harikrishnan, S., Leeder, S., Huffman, M., Jeemon, P., & Prabhakaran, D. (2014). *A race against time: The challenge of cardiovascular disease in developing economies.* New Delhi, India: Centre for Chronic Disease Control. Retrieved from http://earth.columbia.edu/news/2004/images/raceagainsttime_FINAL_051104.pdf

Hasan, M. T., Soares Magalhaes, R. J., Williams, G. M., & Mamun, A. A. (2016). The role of maternal education in the 15-year trajectory of malnutrition in children under 5 years of age in Bangladesh. *Maternal & Child Nutrition, 12*(4), 929–939.

Haskell, M. J., & Brown, K. H. (1999). Maternal vitamin A nutriture and the vitamin A content of human milk. *Journal of Mammary Gland Biology Neoplasia*, *4*, 243–257.

Hawkes, C., & Ruel, M. T. (2012). Value chains for nutrition. In S. Fan & R. Pandya-Lorch (Eds.), *Reshaping agriculture for nutrition and health* (pp. 73–82). Washington, DC: International Food Policy Research Institute.

Hawkes, C., Smith, T. G., Jewell, J., Wardle, J., Hammond, R. A., Friel, S., & Kain, J. (2015). Smart food policies for obesity prevention. *The Lancet, 385*(9985), 2410–2421.

Headey, D. D. (2013). Developmental drivers of nutritional change: A cross-country analysis. *World Development, 42*(1), 76–88.

Hendrickse, R. G., Hasan, A. H., Olumide, L. O., & Akinkunmi, A. (1971). Malaria in early childhood: An investigation of five hundred seriously ill children in whom a "clinical" diagnosis of malaria was made on admission to the children's emergency room at University College Hospital, Ibadan. *Annals of Tropical Medicine and Parasitology, 65*, 1–20.

Hess, S. Y., Abbeddou, S., Jimenez, E. Y., Some, J. W., Vosti, S. A., Ouedraogo, Z. P., & Brown, K. H. (2015). Small-quantity lipid-based nutrient supplements, regardless of their zinc content, increase growth and reduce the prevalence of stunting and wasting in young Burkinabe children: A cluster-randomized trial. *PLoS One, 10*(3), e0122242.

Hetzel, B. S. (1989). *The story of iodine deficiency: An international challenge in nutrition*. Oxford, UK/Delhi, India: Oxford University Press. Retrieved from http://www.nejm.org/doi/full/10.1056/NEJM199011013231821#t=article

Hetzel, B. S. (1994). Historical development of the concepts of the brain–thyroid relationships. In J. B. Stanbury (Ed.), *The damaged brain of iodine deficiency* (pp. 1–7). New York, NY: Cognizant Communication.

High Level Panel of Experts (HLPE). (2012). Social protection for food security. A report by High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security: Rome, Italy. Retrieved from http://www.fao.org/fileadmin/user_upload/hlpe/hlpe_documents/HLPE_Reports/HLPE-Report-4-Social_protection_for_food_security-June_2012.pdf

High Level Panel of Experts (HLPE). (2017). Nutrition and food systems. A report by High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security: Rome, Italy. Retrieved from http://www.gainhealth.org/wp-content/uploads/2017/09/HLPE-12-Nutrition-and-food-systems.pdf

Hlaing, L. M., Fahmida, U., Htet, M. K., Utomo, B., Firmansyah, A., & Ferguson, E. L. (2016). Local food-based complementary feeding recommendations developed by the linear programming approach to improve the intake of problem nutrients among 12–23-month-old Myanmar children. *British Journal of Nutrition, 116*(suppl 1), S16–S26.

Hodge, J. (2016). Hidden hunger: Approaches to tackling micronutrient deficiencies. In S. Gillespie, J. Hodge, S. Yosef, & R. Pandya-Lorch (Eds.), *Nourishing millions: Stories of change in nutrition* (pp. 35–43). Washington, DC: International Food Policy Research Institute.

Hodge, J., Herforth, A., Gillespie, S., Beyero, M., Wagah, M., & Semakula, R. (2015). Is there an enabling environment for nutrition-sensitive agriculture in East Africa? *Food and Nutrition Bulletin, 36*(4), 503–519.

Hofmeyr, G. J., Lawrie, T. A., Atallah, A. N., Duley, L., & Torloni, M. R. (2014). Calcium supplementation during pregnancy for preventing hypertensive disorders and related problems. *Cochrane Database of Systematic Reviews, 6*, CD001059. Retrieved from http://www.cochrane.org/CD001059/PREG_calcium-supplementation-during-pregnancy-for-preventing-hypertensive-disorders-and-related-problems

Holdsworth, M., El Ati, J., Bour, A., Kameli, Y., Derouiche, A., Millstone, E., & Delpeuch, F. (2013). Developing national obesity policy in middle-income countries: A case study from North Africa. *Health Policy and Planning*, *28*(8), 858–870.

Hoogendoorn, A., Luthringer, C., Parvanta, I., & Garrett, G. S. (2016). *Food fortification global mapping study.* European Commission. Retrieved from https://ec.europa.eu/europeaid/sites/devco/files/2fas_global_mapping_study_2016_july_revision.pdf

Horton, S., & Ross, J. (2007). The economics of iron deficiency. *Food Policy, 32*(1), 141–143.

Hotz, C., & Brown, K. H. (Eds.). (2004). International Zinc Nutrition Consultative Group (IZiNCG) Technical Document No. 1: Assessment of the risk of zinc deficiency in populations and options for its control [Special issue]. *Food and Nutrition Bulletin, 25*(suppl 2), S94–S203.

Hoyle, B., Yunus, M., & Hen, L. C. (1980). Breastfeeding and food intake among children with acute diarrheal disease. *American Journal of Clinical Nutrition, 33*, 2365–2371.

Humphrey, J. H. (2009). Child undernutrition, tropical enteropathy, toilets, and handwashing. *The Lancet, 374*, 1032–1035.

Humphrey, J. H., Agoestina, T., Wu, L., Usman, A., Nurachim, M., Subardja, D., . . . Sommer, A. (1996). Impact of neonatal vitamin A supplementation on infant morbidity and mortality. *Journal of Pediatrics, 128*, 489–496.

Humphrey, J. H., Prendergast, A. J., Ntozini, R., & Gladstone, M. (2017, November 7). *The Sanitation Hygiene Infant Nutrition Efficacy (SHINE) Trial.* Symposium at the American Society of Tropical Medicine and Hygiene Annual Meeting, Baltimore, MD.

Hussey, G. D., & Klein, M. (1990). A randomized, controlled trial of vitamin A in children with severe measles. *New England Journal of Medicine, 323*, 160–164.

Iannotti, L. L., Dulience, S. J., Green, J., Joseph, S., Francois, J., Antenor, M. L., & Nickerson, N. M. (2014). Linear growth increased in young children in an urban slum of Haiti: A randomized controlled trial of a lipid-based nutrient supplement. *American Journal of Clinical Nutrition, 99*(1), 198–208.

Iannotti, L. L., Lutter, C. K., Stewart, C. P., Gallegos Riofrio, C. A., Malo, C., Reinhart, G., & Waters, W. F. (2017). Eggs in early complementary feeding and child growth: A randomized controlled trial. *Pediatrics, 140*(1), e20163459.

Institute of Medicine (IOM). (2003). *Improving birth outcomes: Meeting the challenge of the developing world*. Washington, DC: National Academies Press.

Institute of Medicine (IOM). (2005). *Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein, and amino acids*. Washington, DC: National Academies Press. Retrieved from https://www.nal.usda.gov/sites/default/files/fnic_uploads/energy_full_report.pdf

Institute of Medicine (IOM). (2009). *Weight gain during pregnancy: Reexamining the guidelines*. Washington, DC: National Academies Press. Retrieved from https://www.nap.edu/resource/12584/Report-Brief . . . Weight-Gain-During-Pregnancy.pdf

Intergovernmental Panel on Climate Change (IPCC). (2014). *Climate change 2014: Impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. New York, NY: Cambridge University Press. Retrieved from http://www.ipcc.ch/pdf/assessment-report/ar5/wg2/WGIIAR5-FrontMatterA_FINAL.pdf

International Food Policy Research Institute (IFPRI). (2016). *Global nutrition report: From promise to impact, ending malnutrition by 2030*. Washington, DC: Author. Retrieved from https://www.compassion.com/multimedia/global-nutrition-report-2016.pdf

International Food Policy Research Institute (IFPRI). (2017). *Global food policy report*. Washington, DC: Author. Retrieved from http://www.fao.org/3/a-bs201e.pdf

International Labor Organization. (2014). *Maternity and paternity at work: Law and practice across the world*. Geneva, Switzerland: Author. Retrieved from http://www.ilo.org/global/publications/ilo-bookstore/order-online/books/WCMS_242615/lang--en/index.htm

International Panel of Experts on Sustainable Food Systems. (2017). What makes urban food policy happen? Insights from five case studies. Retrieved from http://www.ipes-food.org/images/Reports/Cities_full.pdf.

Islam, S. M. S., Purnat, T. D., Phuong, N. T. A., Mwingira, U., Schacht, K., & Fröschl, G. (2014). Non-communicable diseases (NCDs) in developing countries: A symposium report. *Globalization and Health, 10*(1), 81.

Ittiravivongs, A., Songchitratna, K. S., Ratthapalo, S., & Pattara-Arechachai, J. (1991). Effect of low birthweight on severe childhood diarrhea. *Southeast Asian Journal of Tropical Medicine and Public Health, 22*, 557–562.

Jackson, B. D., & Black, R. E. (2017). A literature review of the effect of malaria on stunting. *Journal of Nutrition, 147*(suppl), 2163S–2168S.

James, J. W. (1972). Longitudinal study of the morbidity of diarrheal and respiratory infections in malnourished children. *American Journal of Clinical Nutrition, 25*, 690–694.

Jaramillo, E. G., Mupere, E., Opoka, R. O., Hodges, J. S., Lund, T. C., Georgieff, M. K., & Cusick, S. E. (2017). Delaying the start of iron until 28 days after antimalarial treatment is associated with lower incidence of subsequent illness in children with malaria and iron deficiency. *PLoS One, 12*(8), e0183977.

Johnson, W. B., Aderele, W. I., & Gbadero, D. A. (1992). Host factors and acute lower respiratory infections in pre-school children. *Journal of Tropical Pediatrics, 38*, 132–136.

Jonker, F. A. M., Te Poel, E., Bates, I., & Boele van Hensbroek, M. (2017). Anaemia, iron deficiency and susceptibility to infection in children in sub-Saharan Africa, guideline dilemmas. *British Journal of Haematology, 177*(6), 878–883.

Kahn, J. G., Yang, J. S., & Kahn, J. S. (2010). "Mobile" health needs and opportunities in developing countries. *Health Affairs, 29*(2), 252–258.

Kalamar, A. M., Lee-Rife, S., & Hindin, M. J. (2016). Interventions to prevent child marriage among young people in low- and middle-income countries: A systematic review of the published and gray literature. *Journal of Adolescent Health*, *59*(3), S16–S21.

Kalter, H. D., Burnham, G., Kolstad, P. R., Hossain, M., Schillinger, J. A., Khan, N. Z., . . . Black, R. E. (1997). Evaluation of clinical signs to diagnose anaemia in Uganda and Bangladesh, in areas with and without malaria. *Bulletin of the World Health Organization, 75*(suppl), 103–111.

Kamiya, Y. (2012). Socioeconomic determinants of nutritional status of children in Lao PDR: Effects of household and community factors. *Journal of Health, Population and Nutrition, 29*(4), 339–348.

Katz, J., Lee, A. C., Kozuki, N., Lawn, J. E., Cousens, S., Blencowe, H., & Black, R.E. (2013). Mortality risk in preterm and small-for-gestational-age infants in low- income and middle-income countries: A pooled country analysis. *The Lancet, 382*(9890), 417–425.

Katz, J., West, K. P., Jr., Khatry, S. K., Pradhan, E. K., LeClerq, S. C., & Christian, P. (2000). Maternal low-dose vitamin A or beta-carotene supplementation has no effect on fetal loss and early infant mortality: A randomized cluster trial in Nepal. *American Journal of Clinical Nutrition, 71*, 1570–1576.

Kavle, J. A., & Landry, M. (2017). Community-based distribution of iron–folic acid supplementation in low- and middle-income countries: A review of evidence and programme implications. *Public Health Nutrition*, 1–9. doi:10.1017/S1368980017002828

Keen, C. L., & Gershwin, M. E. (1990). Zinc deficiency and immune function. *Annual Review of Nutrition, 10*, 415–431.

Kelly, A. S., Barlow, S. E., Rao, G., Inge, T. H., Hayman, L. L., Steinberger, J., & Daniels, S. R. (2013). Severe obesity in children and adolescents: Identification, associated health risks, and treatment approaches. *Circulation, 128*(15), 1689–1712.

Kerac, M., Bunn, J., Chagaluka, G., Bahwere, P., Tomkins, A., Collins, S., & Seal, A. (2014). Follow-up of post-discharge growth and mortality after treatment for severe acute malnutrition (FuSAM study): A prospective cohort study. *PloS One, 9*(6), e96030.

Keusch, G. T., Denno, D. M., Black, R.E., Duggan, C., Guerrant, R. L., Lavery, J. V., & Brewer, T. (2014). Environmental enteric dysfunction: Pathogenesis, diagnosis, and clinical consequences. *Clinical Infectious Diseases, 59*(suppl 4), S207–S212.

Khin-Maung-U, Nyunt-Nyunt-Wai, Myo-Khin, Mu-Mu-Khin, Tin-U, & Thane-Toe. (1985). Effect on clinical outcome of breast feeding during acute diarrhoea. *British Medical Journal, 290*, 587–589.

Khor, G. L., Tan, S. Y., Tan, K. L., Chan, P. S., & Amarra, M. S. (2016). Compliance with WHO IYCF indicators and dietary intake adequacy in a sample of Malaysian infants aged 6–23 months. *Nutrients, 8*(12). doi:10.3390/nu8120778

Kim, S. S., Rawat, R., Mwangi, E. M., Tesfaye, R., Abebe, Y., Baker, J., & Menon, P. (2016). Exposure to large-scale social and behavior change communication interventions is associated with improvements in infant and young child feeding practices in Ethiopia. *PLoS One, 11*(10), e0164800.

Kirkwood, B. R., Hurt, L., Amenga-Etego, S., Tawiah, C., Zandoh, C., Danso, S., . . . Arthur, P. (2010). Effect of vitamin A supplementation in women of reproductive age on maternal survival in Ghana (ObaapaVitA): A cluster-randomised, placebo-controlled trial. *The Lancet, 375*, 1640–1649.

Kiserud, T., Piaggio, G., Carroli, G., Widmer, M., Carvalho, J., Neerup Jensen, L., & Platt, L. D. (2017). The World Health Organization fetal growth charts: A multinational longitudinal study of ultrasound biometric measurements and estimated fetal weight. *PLoS Medicine, 14*(1), e1002220.

Klemm, R. D. W., Labrique, A. B., Christian, P., Rashid, M., Shamim, A. A., Katz, J., & Sommer, A. (2008). Newborn vitamin A supplementation reduced infant mortality in rural Bangladesh. Pediatrics, 122, e242–e250.

Klemm, R. D., Palmer, A. C., Greig, A., Engle-Stone, R., & Dalmiya, N. (2016). A changing landscape for Vitamin A programs: Implications for optimal intervention packages, program monitoring, and safety. *Food and Nutrition Bulletin, 37*(2 suppl), S75–S86.

Knowles, J. M., Garrett, G. S., Gorstein, J., Kupka, R., Situma, R., Yadav, K., & Universal Salt Iodization Coverage Survey Team. (2017). Household coverage with adequately iodized salt varies greatly between countries and by residence ypye and socioeconomic status within countries: Results from 10 national coverage surveys. *Journal of Nutrition, 147*(5), 1004S–1014S.

Knutson, M. D. (2017). Iron transport proteins: Gateways of cellular and systemic iron homeostasis. *Journal of Biological Chemistry, 292*(31), 12735–12743.

Korat, A. A., Willett, W., & Hu, F. (2014). Diet, lifestyle, and genetic risk factors for type 2 diabetes: A review from the Nurses' Health Study, Nurses' Health Study 2, and health professionals' follow-up study. *Current Nutrition Reports, 3*(4), 345–354.

Kothari, M. T., Abderrahim, N., Colie, A., & Cheng, Y. (2014). *Nutritional status of women and children.* Rockville, MD: ICF International.

Kozuki, N., Katz, J., Lee, A. C., Vogel, J. P., Silveira, M. F., Sania, A., & Child Health Epidemiology Reference Group Small-for-Gestational-Age/Preterm Birth Working Group. (2015). Short maternal stature increases risk of small-for-gestational-age and preterm births in low- and middle-income countries: Individual participant data meta-analysis and population attributable fraction. *Journal of Nutrition, 145*(11), 2542–2550.

Kramer, M. S. (1987). Determinants of low birth weight: Methodological assessment and meta-analysis. *Bulletin of the World Health Organization, 65*(5), 663–737.

Krebs, N. F., Mazariegos, M., Chomba, E., Sami, N., Pasha, O., Tshefu, A., & Hambidge, K. M. (2012). Randomized controlled trial of meat compared with multimicronutrient-fortified cereal in infants and toddlers with high stunting rates in diverse settings. *American Journal of Clinical Nutrition, 96*(4), 840–847.

Kutlesic, V., Brewinski Isaacs, M., Freund, L. S., Hazra, R., & Raiten, D. J. (2017). Executive summary: Research gaps at the intersection of pediatric neurodevelopment, nutrition, and inflammation in low-resource settings. *Pediatrics, 139*(S1), S1–S11.

Lachat, C., Otchere, S., Roberfroid, D., Abdulai, A., Seret, F. M. A., Milesevic, J., & Kolsteren, P. (2013). Diet and physical activity for the prevention of noncommunicable diseases in low-and middle-income countries: A systematic policy review. *PLoS Medicine*, *10*(6), e1001465.

Lagarde, M., Haines, A., & Palmer, N. (2007). Conditional cash transfers for improving uptake of health interventions in low- and middle-income countries: A systematic review. *Journal of the American Medical Association, 298*(16), 1900–1910.

Lamberti, L. M., Fischer Walker, C. L., Taneja, S., Mazumder, S., & Black, R. E. (2015). The association between provider practice and knowledge of ORS and zinc supplementation for the treatment of childhood diarrhea in Bihar, Gujarat and Uttar

Pradesh, India: A multi-site cross-sectional study. *PLoS One, 10*(6), e0130845.

Lamberti, L. M., Taneja, S., Mazumder, S., LeFevre, A., Black, R. E., & Walker, C. L. (2015). An external evaluation of the Diarrhea Alleviation through Zinc and ORS Treatment (DAZT) program in Gujarat and Uttar Pradesh, India. *Journal of Global Health, 5*(2), 020409.

Lamberti, L. M., Walker, C. L., Chan, K. Y., Jian, W. Y., & Black, R. E. (2013). Oral zinc supplementation for the treatment of acute diarrhea in children: A systematic review and meta-analysis. *Nutrients, 5*(11), 4715–4740.

Launer, L. J., Habicht, J. P., & Kardjati, S. (1990). Breast feeding protects infants in Indonesia against illness and weight loss due to illness. *American Journal of Epidemiology, 131*, 322–331.

Lazzerini, M., Rubert, L., & Pani, P. (2013). Specially formulated foods for treating children with moderate acute malnutrition in low- and middle-income countries. *Cochrane Database of Systematic Reviews, 6*, CD009584. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD009584.pub2/abstract

Lazzerini, M., & Wanzira, H. (2016). Oral zinc for treating diarrhoea in children. *Cochrane Database of Systematic Reviews, 12*, CD005436. Retrieved from https://dx.doi.org/10.1002/14651858.CD005436.pub3

Lee, G., Paredes Olortegui, M., Penataro Yori, P., Black, R. E., Caulfield, L., Banda Chavez, C., . . . Kosek, M. (2014). Effects of *Shigella*-, *Campylobacter*- and ETEC-associated diarrhea on childhood growth. *Pediatric Infectious Disease Journal, 33*(10), 1004–1009.

Lee, S. E., Talegawkar, S. A., Merialdi, M., & Caulfield, L. E. (2012). Dietary intakes of women during pregnancy in low- and middle-income countries. *Public Health Nutrition, 16,* 1340–1353.

Lenters, L. M., Wazny, K., Webb, P., Ahmed, T., & Bhutta, Z. A. (2013). Treatment of severe and moderate acute malnutrition in low- and middle-income settings: A systematic review, meta-analysis and Delphi process. *BMC Public Health, 13*(suppl 3), S23.

Leroy, J. L., Ruel, M., & Verhofstadt, E. (2009). The impact of conditional cash transfer programmes on child nutrition: A review of evidence using a programme theory framework. *Journal of Development Effectiveness, 1*(2), 103–129.

Li, M., & Eastman, C. J. (2012). The changing epidemiology of iodine deficiency. *Nature Reviews Endocrinology, 8*(7), 434–440.

Liu, L., Oza, S., Hogan, D., Chu, Y., Perin, J., Zhu, J., . . . Black, R. E. (2016). Global, regional, and national causes of under-5 mortality in 2000–15: An updated systematic analysis with implications for the Sustainable Development Goals. *The Lancet, 388*(10063), 3027–3035.

Liu, X., Cui, T., Li, Y., Wang, Y., Wang, Q., Li, X., . . . Chen, J. (2014). Vitamin A supplementation in early life enhances the intestinal immune response of rats with gestational vitamin A deficiency by increasing the number of immune cells. *PLoS One, 9*(12), e114934.

Lobstein, T., Jackson-Leach, R., Moodie, M. L., Hall, K. D., Gortmaker, S. L., Swinburn, B. A., & McPherson, K. (2015). Child and adolescent obesity: Part of a bigger picture. *The Lancet*, *385*(9986), 2510–2520.

Lönnerdal, B., & Hernell, O. (2013). Iron: Physiology, dietary sources, and requirements. In B. Caballero (Ed.), *Encyclopedia of human nutrition* (Vol. 3, pp. 39–46). Amsterdam, Netherlands: Elsevier.

Low, J. W., Arimond, M., Osman, N., Cunguara, B., Zano, F., & Tschirley, D. (2007). A food-based approach introducing orange-fleshed sweet potatoes increased vitamin A intake and serum retinol concentrations in young children in rural Mozambique. *Journal of Nutrition, 137*(5), 1320–1327.

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., & Ahn, S. Y. (2013). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease study 2010. *The Lancet, 380*(9859), 2095–2128.

Lozoff, B., Beard, J., Connor, J., Barbara, F., Georgieff, M., & Schallert, T. (2006). Long-lasting neural and behavioral effects of iron deficiency in infancy. *Nutrition Reviews, 64*(5 Pt 2), S34–S43.

Lu, Z., O'Dell, D., Srinivasan, B., Rey, E., Wang, R., Vemulapati, S., . . . Erickson, D. (2017). Rapid diagnostic testing platform for iron and vitamin A deficiency. *Proceedings of the National Academy of Sciences*, *114*(51), 13513–13518.

Luby, S. P., Kazembe, P. N., Redd, S. C., Ziba, C., Nwanyanwu, O. C., Hightower, A. W., . . . Olivar, M. A. (1995). Using clinical signs to diagnose anaemia in African children. *Bulletin of the World Health Organization, 73*, 477–482.

Luby, S. P., Rahman, M., Arnold, B. F., Unicomb, L., Ashraf, S., Winch, P. J., & Colford, J. M., Jr. (2018). Effects of water quality, sanitation, handwashing and nutritional interventions on diarrhoea and child growth in rural Bangladesh: A cluster randomized trial. *The Lancet Global Health, 6(3), e302-e315*.

Lumbiganon, P., Laopaiboon, M., Intarut, N., Vogel, J. P., Souza, J. P., Gulmezoglu, A. M., & Newborn Health Research Network. (2014). Indirect causes of severe adverse maternal outcomes: A secondary analysis of the WHO Multicountry Survey on Maternal and Newborn Health. *British Journal of Obstetrics and Gynecology, 121*(suppl 1), 32–39.

Lunn, P. G., Northrop-Clewes, C. A., & Downes, R. M. (1991). Intestinal permeability, mucosal injury, and growth faltering in Gambian infants. *The Lancet, 338*, 907–910.

Luthringer, C. L., Rowe, L. A., Vossenaar, M., & Garrett, G. S. (2015). Regulatory monitoring of fortified foods: Identifying barriers and good practices. *Global Health: Science and Practice, 3,* 446–461.

Lutter, C. K., Mora, J. O., Habicht, J. P., Rasmussen, K. M., Robson, D. S., Sellers, S. G., . . . Herrera, M. G. (1989). Nutritional supplementation: Effects on child stunting because of diarrhea. *American Journal of Clinical Nutrition, 50*, 1–8.

Mahalanabis, D., Lahiri, M., Paul, D., Gupta, S., Gupta, A., Wahed, M. A., & Khaled, M. (2004). Randomized, double-blind, placebo-controlled clinical trial of the efficacy of treatment with zinc or vitamin A in infants and young children with severe acute lower respiratory infection. *American Journal of Nutrition, 79*, 430–436.

Makoka, D., & Masibo, P. K. (2015). Is there a threshold level of maternal education sufficient to reduce child undernutrition? Evidence from Malawi, Tanzania and Zimbabwe. *BMC Pediatrics, 15*(1), 96. doi:10.1186/s12887-015-0406-8

Malaba, L. C., Iliff, P. J., Nathoo, K. J., Marinda, E., Moulton, L. H., Zijenah, L.S., . . . Humphrey, J. H. (2005). Effect of postpartum maternal or neonatal vitamin A supplementation on infant mortality among infants born to HIV-negative mothers in Zimbabwe. *American Journal of Clinical Nutrition, 81*, 454–460.

Malapit, H. J. L., & Quisumbing, A. R. (2015). What dimensions of women's empowerment in agriculture matter for nutrition in Ghana? *Food Policy*, *52*, 54–63.

Maleta, K. M., Phuka, J., Alho, L., Cheung, Y. B., Dewey, K. G., Ashorn, U., & Ashorn, P. (2015). Provision of 10–40 g/d lipid-based nutrient supplements from 6 to 18 months of age does not prevent linear growth faltering in Malawi. *Journal of Nutrition, 145*(8), 1909–1915.

Malik, V. S., Willett, W. C., & Hu, F. B. (2013). Global obesity: Trends, risk factors and policy implications. *Nature Reviews Endocrinology, 9*(1), 13–27.

Man, W. D., Weber, M., Palmer, A., Schneider, G., Wadda, R., Jaffar, S., . . . Greenwood, B. M. (1998). Nutritional status of children admitted to hospital with different diseases and its relationship to outcome in the Gambia, West Africa. *Tropical Medicine and International Health, 3*, 678–686.

Mannar, M. G. V., & Dunn, J. T. (1995). *Salt iodization for the elimination of iodine deficiency.* Amsterdam, Netherlands: International Council for Control of Iodine Deficiency Disorders. Retrieved from http://apps.who.int/iris/handle/10665/36917.

Mariath, J. G., Lima, M. C., & Santos, L. M. (1989). Vitamin A activity of buriti (*Mauritia vinifera* Mart) and its effectiveness in the treatment and prevention of xerophthalmia. *American Journal of Clinical Nutrition, 49*, 849–853.

Martin, S. L., Omotayo, M. O., Pelto, G. H., Chapleau, G. M., Stoltzfus, R. J., & Dickin, K. L. (2017). Adherence-specific social support enhances adherence to calcium supplementation regimens among pregnant women. *Journal of Nutrition, 147*(4), 688–696.

Martorell, R., Yarbrough, C., Lechtig, A., Habicht, J. P., & Klein, R. E. (1975). Diarrheal diseases and growth retardation in preschool Guatemalan children. *American Journal of Physical Anthropology, 43*, 341–346.

Martorell, R., & Zongrone, A. (2012). Intergenerational influences on child growth and undernutrition. *Paediatric and Perinatal Epidemiology, 26*(suppl 1), 302–314.

Masanja, H., Smith, E. R., Muhihi, A., Briegleb, C., Mshamu, S., Ruben, J., . . . Fawzi, W. W. (2015). Effect of neonatal vitamin A supplementation on mortality in infants in Tanzania (Neovita): A randomized, double-blind, placebo-controlled trial. *The Lancet, 385*, 1324–1332.

Masset, E., Haddad, L., Cornelius, A., & Isaza Castro, J. (2012). Effectiveness of agricultural interventions that aim to improve nutritional status of children: Systematic review. *British Medical Journal, 344,* d8222. doi:https://doi.org/10.1136/bmj.d8222

Mata, L. J., Cromal, R. A., Urrutia, J. J., & Garcia, B. (1977). Effect of infection of food intake and the nutritional state: Perspectives as viewed from the village. *American Journal of Clinical Nutrition, 30*, 1215–1227.

Matopoulos, A., Vlachopoulou, M., Manthou, V., & Manos, B. (2007). A conceptual framework for supply chain collaboration: Empirical evidence from the agri-food industry. *Supply Chain Management, 12*(3), 177–186.

Mayo-Wilson, E., Junior, J. A., Imdad, A., Dean, S., Chan, X. H., Chan, E. S., & Bhutta, Z. A. (2014). Zinc supplementation for preventing mortality, morbidity, and growth failure in children aged 6 months to 12 years of age. *Cochrane Database of Systematic Reviews, 5*, CD009384. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD009384.pub2/abstract

Mazumder, S., Taneja, S., Bhatio, K., Yoshida, S., Kaur, J., Dube, B., . . . Bhandari, N. (2015). Efficacy of early neonatal supplementation with vitamin A to reduce mortality in infancy in Haryana, India (Neovita): A randomized, double-blind, placebo-controlled trial. *The Lancet, 385*, 1333–1342.

McCauley, M. E., Van Den Broek, N., Dou, L., & Othman, M. (2015). Vitamin A supplementation during pregnancy for maternal and newborn outcomes. *Cochrane Database of Systematic Reviews, 10*, CD008666. Retrieved from http://www.cochrane.org/CD008666/PREG_vitamin-supplementation-during-pregnancy-maternal-and-newborn-health-outcomes

McGovern, M. E., Krishna, A., Aguayo, V. M., & Subramanian, S. V. (2017). A review of the evidence linking child stunting to economic outcomes. *International Journal of Epidemiology*, *46*(4), 1171–1191.

McMahon, R. J., & Cousins, R. J. (1998). Mammalian zinc transporters. *Journal of Nutrition, 128*, 667–670.

Mead, E., Brown, T., Rees, K., Azevedo, L. B., Whittaker, V., Jones, D., . . . Ells, L. J. (2017). Diet, physical activity and behavioural interventions for the treatment of overweight or obese children from the age of 6 to 11 years. *Cochrane Database of Systematic Reviews, 6,* CD012651. doi:10.1002/14651858.CD012651.pub1

Menon, P., Nguyen, P. H., Saha, K. K., Khaled, A., Kennedy, A., Tran, L. M., & Rawat, R. (2016). Impacts on breastfeeding practices of at-scale strategies that combine intensive interpersonal counseling, mass media, and community mobilization: Results of cluster-randomized program evaluations in Bangladesh and Viet Nam. *PLoS Medicine, 13*(10), e1002159.

Menon, P., Nguyen, P. H., Saha, K. K., Khaled, A., Sanghvi, T., Baker, J., & Rawat, R. (2016). Combining intensive counseling by frontline workers with a nationwide mass media campaign has large differential impacts on complementary feeding practices but not on child growth: Results of a cluster-randomized program evaluation in Bangladesh. *Journal of Nutrition, 146*(10), 2075–2084.

Messer, E. (1997). Intra-household allocation of food and health care: Current findings and understandings: Introduction. *Social Science & Medicine, 44*(11), 1675–1684.

Micha, R., Wallace, S. K., & Mozaffarian, D. (2010). Red and processed meat consumption and risk of incident coronary heart disease, stroke, and diabetes mellitus. *Circulation*, *121*(21), 2271–2283.

Monteiro, C. A., & Cannon, G. (2012). The impact of transnational "big food" companies on the south: A view from Brazil. *PLoS Medicine, 9*(7), e1001252.

Monteiro, C. A., Cannon, G., Moubarac, J. C., Levy, R. B., Louzada, M., & Jaime, P. C. (2017). The UN decade of nutrition, the NOVA food classification and the trouble with ultra-processing. *Public Health Nutrition, 21*, 1–13.

Moodie, R., Stuckler, D., Monteiro, C., Sheron, N., Neal, B., Thamarangsi, T., & Casswell, S. (2013). Profits and pandemics: Prevention of harmful effects of tobacco, alcohol, and ultra-processed food and drink industries. *The Lancet, 381*(9867), 670–679.

Mora, J. R., & Iwata, M. (2015). Retinoids and the immune system. In P. Dolle & K. Niederreither (Eds.), *The retinoids: Biology, biochemistry, and disease* (pp. 465–483). Hoboken, NJ: Wiley & Sons.

Moreira, P., Baraldi, L. G., Moubarac, J., Monteiro, C. A., Newton, A., Capewell, S., & O'Flaherty, M. (2015). Comparing different policy scenarios to reduce the consumption of ultra-processed foods in UK: Impact on cardiovascular disease mortality using a modelling approach. *PloS One, 10*(2), e0118353.

Morley, D. (1964). The severe measles of West Africa. *Proceedings of the Royal Society of Medicine, 57*, 846–849.

Morris, S. S., Cogill, B., & Uauy, R. (2008). Effective international action against undernutrition: Why has it proven so difficult and what can be done to accelerate progress? *The Lancet, 371*(9612), 608–621.

Moubarac, J., Batal, M., Louzada, M. L., Martinez Steele, E., & Monteiro, C. A. (2017). Consumption of ultra-processed foods predicts diet quality in Canada. *Appetite, 108*, 512–520.

Moursi, M. M., Arimond, M., Dewey, K. G., Treche, S., Ruel, M. T., & Delpeuch, F. (2008). Dietary diversity is a good predictor of the micronutrient density of the diet of 6- to 23-month-old children in Madagascar. *Journal of Nutrition, 138*(12), 2448–2453.

Moy, R. J. D., Marshall, T. F. de C., Choto, R. G. A. B., McNeish, A. S., & Booth, I. W. (1994). Diarrhoea and growth faltering in rural Zimbabwe. *European Journal of Clinical Nutrition, 48*, 810–821.

Mozaffarian, D. (2016). Dietary and policy priorities for cardiovascular disease, diabetes, and obesity: A comprehensive review. *Circulation*, *133*(2), 187–225.

Muller, M. (1974). *The baby killer: A War on Want investigation into the promotion and sale of powdered baby milks in the Third World*. London, UK: War on Want. Retrieved from http://www.waronwant.org/resources/baby-killer

Murray, M. J., Murray, N. J., Murray, A. B., & Murray, M. B. (1975). Refeeding-malaria and hyperferraemia. *The Lancet, 1*, 653–654.

Myers, S. S., Smith, M. R., Guth, S., Golden, C. D., Vaitla, B., Mueller, N. D., . . . Huybers, P. (2017). Climate change and global food systems: Potential impacts on food security and undernutrition. *Annual Review of Public Health, 38*, 259–277.

Myers, S. S., Wessells, K. R., Kloog, I., Zanobetti, A., & Schwartza, J. (2015). Rising atmospheric CO_2 increases global threat of zinc deficiency. *The Lancet Global Health, 3*(10), e639.

Myers, S. S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A. D. B., Bloom, A. J., . . . Usui, Y. (2014). Increasing CO_2 threatens human nutrition. *Nature, 510*(7503), 139–142.

Mwangi, M. N., Prentice, A. M., & Verhoef, H. (2017). Safety and benefits of antenatal oral iron supplementation in low-income countries: A review. *British Journal of Hematology, 177*, 884–895.

Nacul, L. C., Arthur, P., Kirkwood, B. R., Morris, S. S., Cameiro, A. C., & Benjamin, A. F. (1998). The impact of vitamin A supplementation given during a pneumonia episode on the subsequent morbidity of children. *Tropical Medicine and International Health, 3,* 661–666.

Namaste, S. M., Rohner, F., Huang, J., Bhushan, N. L., Flores-Ayala, R., Kupka, R., & Suchdev, P. S. (2017). Adjusting ferritin concentrations for inflammation: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *American Journal of Clinical Nutrition, 106*(suppl 1), 359S–371S.

Nazki, F. H., Sameer, A. G., & Ganaie, B. A. (2014) Folate: Metabolism, genes, polymorphisms and the associated diseases. *Gene, 533,* 11–20.

NCD Risk Factor Collaboration. (2016). Worldwide trends in diabetes since 1980: A pooled analysis of 751 population-based studies with 4.4 million participants. *The Lancet, 387*(10027), 1513–1530.

Nelson, G. C., Rosegrant, M. W., Palazzo, A., Gray, I., Ingersoll, C., Robertson, R. D., & You, L. (2010). *Food security, farming and climage change to 2050: Scenarios, results, and policy options*. Washington, DC: International Food Policy Research Institute. Retrieved from http://www.ifpri.org/publication/food-security-farming-and-climate-change-2050

Nishida, C., Uauy, R., Kumanyika, S., & Shetty, P. (2004). The Joint WHO/FAO Expert Consultation on diet, nutrition and the prevention of chronic diseases: process, product and policy implications. *Public Health Nutrition*, *7*(1a), 245-250.

Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., & Abera, S. F. (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: A systematic analysis for the global burden of disease study 2013. *The Lancet, 384*(9945), 766–781.

Null, C., Stewart, C. P., Pickering, A. J., Dentz, H. N., Arnold, B. F., Arnold, C. D., & Colford, J. M., Jr. (2018). Effects of water quality, sanitation, handwashing and nutritional interventions on diarrhea and child growth in rural Kenya: A cluster-randomised trial. *The Lancet Global Health, 6*(3), e316-e329.

Ochola, S., & Masibo, P. K. (2014). Dietary intake of schoolchildren and adolescents in developing countries. *Annals of Nutrition and Metabolism, 64*(s2), 24–40.

O'Dempsey, T. J., McArdle, R. F., Lloyd-Evans, N., Baldeh, I., Lawrence, B. E., Secka, O., & Greenwood, B. (1996). Pneumococcal disease among children in a rural area of West Africa. *Pediatric Infectious Disease Journal, 15*, 431–437.

Olivares, M., Hertrampf, E., & Uauy, R. (2007). Copper and zinc interactions in anemia: A public health perspective. In K. Kraemer & M. B. Zimmermann (Eds.), *Nutritional anemia* (pp. 99–109). Basel, Switzerland: Sight and Life Press.

Olney, D. K., Pedehombga, A., Ruel, M. T., & Dillon, A. (2015). A 2-year integrated agriculture and nutrition and health behavior change communication program targeted to women in Burkina Faso reduces anemia, wasting, and diarrhea in children 3–12.9 months of age at baseline: A cluster-randomized controlled trial. *Journal of Nutrition, 145*(6), 1317–1324.

Olofin, I., McDonald, C. M., Ezzati, M., Flaxman, S., Black, R. E., Fawzi, W. W., . . . Danaei, G. (2013). Associations of suboptimal growth with all-cause and cause-specific mortality in children under five years: A pooled analysis of ten prospective studies. *PLoS One, 8*(5), e64636. doi:10.1371/journal.pone.0064636

Osendarp, S. J. M., van Raaij, J. M. A., Darmstadt, G. L., Baqui, A. H., Hautvast, J. G. A., & Fuchs, G. J. (2001). Zinc supplementation during pregnancy and effects on growth and morbidity in low birthweight infants: A randomized placebo controlled trial. *The Lancet, 357*, 1080–1085.

Ota, E., Mori, R., Middleton, P., Tobe-Gai, R., Mahomed, K., Miyazaki, C., & Bhutta, Z. A. (2015). Zinc supplementation for improving pregnancy and infant outcome. *Cochrane Database of Systematic Reviews, 2*, CD000230. Retrieved from http://www.cochrane.org/CD000230/PREG_zinc-supplementation-improving-pregnancy-and-infant-outcome

Otten, J. J., Hellwig, J. P., & L. D. Meyers (Eds.). (2006). *Dietary reference intakes: Essential guide to nutrient requirements.* Institute of Medicine. Washington, DC: National Academies Press. Retrieved from https://www.nal.usda.gov/sites/default/files/fnic_uploads/DRIEssentialGuideNutReq.pdf

Pachon, H., Spohrer, R., Mei, Z., & Serdula, M. K. (2015). Evidence of the effectiveness of flour fortification programs on iron status and anemia: a systematic review. *Nutrition Reviews, 73*(11), 780–795.

Palmer, A. C., Healy, K., Barffour, M. A., Siamusantu, W., Chileshe, J., Schulze, K. J., . . . Labrique, A. B. (2016). Provitamin A carotenoid-biofortified maize consumption increases pupillary responsiveness among Zambian children in a randomized controlled trial. *Journal of Nutrition, 146*, 2551–2558.

Palmer, A. C., Schulze, K. J., Khatry, S. K., De Luca, L. M., & West, K. P., Jr. (2015). Maternal vitamin A supplementation increases natural antibody concentrations of pre-adolescent offspring in rural Nepal. *Nutrition*, *31*, 813–819.

Palmer, A. C., West, K. P., Jr., Dalmiya, N., & Schultink, W. (2012). The use and interpretation of serum retinol distributions in evaluating the public health impact of vitamin A programmes. *Public Health Nutrition, 15*(7), 1201–1215.

Pan American Health Organization (PAHO). (2003). *Guiding principles for complementary feeding of the breastfed child*. Washington, DC: Author. Retrieved from http://www.who.int/nutrition/publications/guiding_principles_compfeeding_breastfed.pdf

Pan American Health Organization (PAHO). (2015). *Ultra-processed food and drink products in Latin America: Trends, impact on obesity, policy implications.* Washington, DC: Author. Retrieved from http://iris.paho.org/xmlui/bitstream/handle/123456789/7699/9789275118641_eng.pdf

Pasricha, S. R., Hayes, E., Kalumba, K., & Biggs, B. A. (2013). Effect of daily iron supplementation on health in children aged 4–23 months: A systematic review and meta-analysis of randomised controlled trials. *The Lancet Global Health, 1*(2), E77–E86.

Patil, S. R., Arnold, B. F., Salvatore, A. L, Briceno, B., Ganguly, S., Colford, J. M., & Gertler, P. J. (2014). The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: A cluster randomized controlled trial. *PLoS Medicine, 11*(8), e1001709.

Pelletier, D. L., Frongillo, E. D., & Habicht, J. P. (1993). Epidemiologic evidence for a potentiating effect of malnutrition on child mortality. *American Journal of Public Health, 83*, 1130–1133.

Pena-Rosas, J. P., De-Regil, L. M., Garcia-Casal, M. N., & Dowswell, T. (2015). Daily oral iron supplementation during pregnancy. *Cochrane Database Systematic Reviews, 7*, CD004736. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD004736.pub4/abstract;jsessionid=49A9B0CB1F8B5928B50F8572FD9BD1F7.f01t03

Penny, M. E., Creed-Kanashiro, H. M., Robert, R. C., Narro, M. R., Caulfield, L. E., & Black, R. E. (2005). Effectiveness of an educational intervention delivered through the health services to improve nutrition in young children: A cluster-randomised controlled trial. *The Lancet, 365*(9474), 1863–1872.

Perkins, A. V., & Vanderlelie, J. J. (2016). Multiple micronutrient supplementation and birth outcomes: The potential importance of selenium. *Placenta, 36,* S61–S65.

Pickering, A. J., Djebbari, H., Lopez, C., Coulibaly, M., & Alzua, M. L. (2015). Effect of a community-led total sanitation intervention on child diarrhoea and child growth in rural Mali: A cluster-randomised controlled trial. *The Lancet Global Health, 3*(11), e701–e711.

Pinstrup-Anderson, P. (1993). *The political economy of food and nutrition policies.* Washington, DC: International Food Policy Research Institute. Retrieved from http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/126459

Pinstrup-Andersen, P. (2007). Agricultural research and policy for better health and nutrition in developing countries: A food systems approach. *Agricultural Economics, 37*(suppl 1), 187–198.

Popkin, B. M. (2006). Technology, transport, globalization and the nutrition transition food policy. *Food Policy, 31*(6), 554–569.

Popkin, B. M. (2013). Bellagio declaration 2013. *Obesity Reviews, 14*(suppl 2), 9–10.

Popkin, B. M., Adair, L. S., & Ng, S. W. (2012). Global nutrition transition and the pandemic of obesity in developing countries. *Nutrition Reviews, 70*(1), 3–21.

Popkin, B. M., & Hawkes, C. (2016). Sweetening of the global diet, particularly beverages: Patterns, trends, and policy responses. *The Lancet Diabetes Endocrinology, 4*, 174–186.

Prado, E. L., & Dewey, K. G. (2014). Nutrition and brain development in early life. *Nutrition Reviews, 72*(4), 267–284.

Prado, E. L., Ullman, M. T., Muadz, H., Alcock, K. J., Shankar, A. H., & SUMMIT Study Group. (2012). The effect of maternal multiple micronutrient supplementation on cognition and mood during pregnancy and postpartum in Indonesia: A randomized trial. *PLoS One, 7,* e32519.

Prasad, A. S. (1985). Clinical manifestations of zinc deficiency. *Annual Review of Nutrition, 5*, 341–363.

Prentice, A. (1995). Regional variations in the composition of human milk. In R. G. Jensen (Ed.), *Handbook of milk composition* (pp. 115–221). San Diego, CA: Academic Press.

Prentice, A. M. (2005). The emerging epidemic of obesity in developing countries. *International Journal of Epidemiology, 35*(1), 93–99.

Pullar, J., Allen, L., Townsend, N., Williams, J., Foster, C., Roberts, N., & Wickramasinghe, K. (2016). Poverty and development interventions that address behavioural risk factors of non-communicable diseases in low and lower middle income countries: A systematic review. *The Lancet*, *388*, S97.

Qian, M., Wang, D., Watkins, W. E., Gebski, V., Yan, Y. Q., Li, M., & Chen, Z. P. (2005). The effects of iodine on intelligence in children: A meta-analysis of studies conducted in China. *Asia Pacific Journal of Clinical Nutrition, 14*(1), 32–42.

Quisumbing, A. R., & Maluccio, J. A. (2003). Resources at marriage and intrahousehold allocation: Evidence from Bangladesh, Ethiopia, Indonesia, and South Africa. *Oxford Bulletin of Economics and Statistics, 65*(3), 283–327.

Rahman, M. M., Huq, F., Sack, D. A., Butler, T., Azad, A. K., Alam, A., . . . Islam, M. (1990). Acute lower respiratory infections in hospitalized patients with diarrhea in Dhaka, Bangladesh. *Reviews of Infectious Diseases, 12*, S899–S906.

Rahmathullah, L., Tielsch, J. M., Thulasirai, R. D., Katz, J., Coles, C., Devi, S., . . . Kamaraj, C. (2003). Impact of supplementing newborn infants with vitamin A on early infant mortality: Community based randomized trial in southern India. *British Medical Journal, 327*, 254–259.

Rahmathullah, L., Underwood, B. A., Thulasiraj, R. D., Milton, R. C., Ramaswamy, K., Rahmathullah, R., & Babu, G. (1990). Reducing mortality among children in southern India receiving a small weekly dose of vitamin A. *New England Journal of Medicine, 323*, 929–935.

Ramakrishnan, U., Aburto, N., McCabe, G., & Martorell, R. (2004). Multimicronutrient interventions but not vitamin A or iron interventions alone improve child growth: Results of 3 meta-analyses. *Journal of Nutrition, 134*, 2592–2602.

Ranganathan, J., Vennard, D., Waite, R., Lipinski, B., Searchinger, T., Dumas, P., & Mungkung, R. (2016). *Shifting diets for a sustainable food future: Installment 11 of "Creating a sustainable food future."* Working paper. Washington, DC: World Resources Institute. Retrieved from http://www.wri.org/sites/default/files/Shifting_Diets_for_a_Sustainable_Food_Future_0.pdf

Rawat, R., Nguyen, P. H., Tran, L. M., Hajeebhoy, N., Nguyen, H. V., Baker, J., & Menon, P. (2017). Social franchising and a nationwide mass media campaign increased the prevalence of adequate complementary feeding in Vietnam: A

cluster-randomized program evaluation. *Journal of Nutrition, 147*(4), 670–679.

Reddy, V. (1991). Protein-energy malnutrition. In P. Stanfield, M. Brueton, M. Chan, M. Parkin, & T. Waterston (Eds.), *Diseases of children in the subtropics and tropics* (4th ed., pp. 335–357). London, UK: Hodder & Stoughton.

Reifen, R., Nur, T., Ghebermeskel, K., Zaiger, G., Urizky, R., & Pines, M. (2002). Vitamin A deficiency exacerbates inflammation in a rat model of colitis through activation of nuclear factor–kappa B and collagen formation. *Journal of Nutrition*, *132*, 2743–2747.

Reinhardt, K., & Fanzo, J. (2014). Addressing chronic malnutrition through multi-sectoral, sustainable approaches: A review of the causes and consequences. *Frontiers in Nutrition, 1*, 13. doi:10.3389/fnut.2014.00013

Richard, S. A., McCormick, B. J., Miller, M. A., Caulfield, L. E., & Checkley, W. (2014). Modeling environmental influences on child growth in the MAL-ED cohort study: Opportunities and challenges. *Clinical Infectious Diseases*, *59*(suppl 4), S255–S260.

Roberto, C. A., Swinburn, B., Hawkes, C., Huang, T. T., Costa, S. A., Ashe, M., & Brownell, K. D. (2015). Patchy progress on obesity prevention: Emerging examples, entrenched barriers, and new thinking. *The Lancet, 385*(9985), 2400–2409.

Rohner, F., Leyvraz, M., Konan, A. G., Esso, L. J. C. E., Wirth, P., Norte, A., . . . Aaron, G. J. (2016). The potential of food fortification to add micronutrients in young children and women of reproductive age: Findings from a cross-sectional survey in Abidjan, Côte d'Ivoire. *PLoS One.* doi:10.1371 / journal.pone.0158552

Rohner, F., Namaste, S. M., Larson, L. M., Addo, O. Y., Mei, Z., Suchdev, P. S., & Northrop-Clewes, C. A. (2017). Adjusting soluble transferrin receptor concentrations for inflammation: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *American Journal of Clinical Nutrition, 106*(suppl 1), 372S–382S.

Rohner, F., Zimmermann, M., Jooste, P., Pandav, C., Caldwell, K., Raghavan, R., & Raiten, D. J. (2014). Biomarkers of nutrition for development: Iodine review. *Journal of Nutrition, 144*(8), 1322S–1342S.

Rollins, N. C., Bhandari, N., Hajeebhoy, N., Horton, S., Lutter, C. K., Martines, J. C., & *The Lancet* Breastfeeding Series Group. (2016). Why invest, and what it will take to improve breastfeeding practices? *The Lancet, 387*(10017), 491–504.

Roth, D. E., Leung, M., Mesfin, E., Qamar, H., Watterworth, J., & Papp, E. (2017) Vitamin D supplementation during pregnancy: State of the evidence from a systematic review of randomised trials. *British Medical Journal, 359,* j5237. doi:10.1136/bmj.j5237

Roth, D. E., Richard, S. A., & Black, R. E. (2010). Zinc supplementation for the prevention of acute lower respiratory infection in children in developing countries: Meta-analysis and meta-regression of randomized trials. *International Journal of Epidemiology, 39*, 795–808.

Rotondi, M. A., & Khobzi, N. (2010). Vitamin A supplementation and neonatal mortality in the developing world: A meta-regression of cluster-randomized trials. *Bulletin of the World Health Organization,* 88, 697–702.

Rousham, E. K., & Mascie-Taylor, C. G. (1994). An 18-month study of the effect of periodic anthelminthic treatment on the growth and nutritional status of pre-school children in Bangladesh. *Annals of Human Biology, 21*, 315–324.

Rowland, M. G. M., Cole, T. J., & Whitehead, R. G. (1977). A quantitative study into the role of infection in determining nutritional status in Gambian village children. *British Journal of Nutrition, 37*, 441–450.

Rowland, M. G. M., Rowland, S. G. J. G., & Cole, T. J. (1988). Impact of infection on the growth of children from 0 to 2 years in an urban West African community. *American Journal of Clinical Nutrition, 47*, 134–138.

Roy, S. K., Chowdhury, A. K., & Rahaman, M. M. (1983). Excess mortality among children discharged from hospital after treatment for diarrhea in rural Bangladesh. *British Medical Journal, 287*, 1097–1099.

Ruel, M. T., Alderman, H., & Maternal and Child Nutrition Study Group. (2013). Nutrition-sensitive interventions and programmes: How can they help to accelerate progress in improving maternal and child nutrition? *The Lancet, 382*(9891), 536–551.

Ruel-Bergeron, J., Stevens, G. A., Sugimoto, J. D., Roos, F. F., Ezzati, M., Black, R. E., & Kraemer, K. (2015). Global update and trends of hidden hunger, 1995–2011: The hidden hunger index. *PLoS One, 10*(12), e0143497.

Saari, J. C. (2016), Vitamin A and vision. *Subcellular Biochemistry, 81*, 231–259.

Sablah, M., Klopps, J., Steinberg, D., Touaoro, Z., Laillou, A., & Baker, S. (2012). Thriving public–private partnership to fortify cooking oil in the West African Economic and Monetary Union (UEMOA) to control vitamin A deficiency: Faire Tache 'Huille en Afrique de l'Ouest. *Food and Nutrition Bulletin, 33*, S310–S320.

Salam, R. A., & Bhutta, Z. A. (2015). Adolescent nutrition. In B. Koletzko, J. Bhatia, Z. A. Bhutta, P. Cooper, M. Makrides, R. Uauy, & W. Wang (Eds.), *Pediatric nutrition in practice* (pp. 113, 122–126). Basel, Switzerland: Karger.

Salam, R. A., MacPhail, C., Das, J. K., & Bhutta, Z. A. (2013). Effectiveness of micronutrient powders (MNP) in women and children. *BMC Public Health, 13*(suppl 3), S22. doi:10.1186/1471-2458-13-S3-S22

Saleem, A. F., Mahmud, S., Baig-Ansari, N., & Zaidi, A. (2014). Impact of maternal education about complementary feeding on their infants' nutritional outcomes in low- and middle-income households: A community-based randomized interventional study in Karachi, Pakistan. *Journal of Health, Population and Nutrition, 32*(4), 623–633.

Sankar, M. J., Sinha, B., Chowdhury, R., Bhandari, N., Taneja, S., Martines, J., & Bahl, R. (2015). Optimal breastfeeding practices and infant and child mortality: A systematic review and meta-analysis. *Acta Paediatrica, 104*(467), 3–13.

Santhya, K. G. (2011). Early marriage and sexual and reproductive health vulnerabilities of young women: A synthesis of recent evidence from developing countries. *Current Opinion in Obstetrics & Gynecology, 23*(5), 334–339.

Santika, O., Fahmida, U., & Ferguson, E. L. (2009). Development of food-based complementary feeding recommendations for 9- to 11-month-old peri-urban Indonesian infants using linear programming. *Journal of Nutrition, 139*(1), 135–141.

Sazawal, S., Black, R. E., Bhan, M. K., Bhandari, N., Sinha, A., & Jalla, S. (1995). Zinc supplementation in young children with acute diarrhea in India. *New England Journal of Medicine, 333*, 839–844.

Sazawal, S., Black, R. E., Menon, V. P., Dhingra, P., Caulfield, L. E., Dhingra, U., & Bagati, A. (2001). Zinc supplementation in infants born small for gestational age reduces mortality: A prospective randomized controlled trial. *Pediatrics, 108*, 1280–1286.

Sazawal, S., Black, R. E., Ramsan, M., Chwaya, H. M., Dutta, A., Dhingra, U., . . . Kabole, F. M. (2007). Effect of zinc supplementation on mortality in children aged 1–48 months: A community-based randomized placebo-controlled trial. *The Lancet, 369*, 927–934.

Sazawal, S., Black, R. E., Ramsan, M., Chwaya, H. M., Stoltzfus, R. J., Dutta, A., & Kabole, F. M. (2006). Effects of routine prophylactic supplementation with iron and folic acid on admission to hospital and mortality in preschool children in a high malaria transmission setting: Community-based, randomised, placebo-controlled trial. *The Lancet, 367*(9505), 133–143.

Schmitz, J., West, K. P., Jr., Khatry, S. K., Wu, L., LeClerq, S. C., Karna, S. L., . . . Pillion, J. (2012). Vitamin A supplementation in preschool children and risk of hearing loss as adolescents and young adults in rural Nepal: Randomized trial cohort follow-up study. *British Medical Journal, 344*, d7962. doi:10.1136/bmj.d7962

Scott, S. P., Chen-Edinboro, L. P., Caulfield, L. E., & Murray-Kolb, L. E. (2014). The impact of anemia on child mortality: An updated review. *Nutrients, 6*(12), 5915–5932.

Scrimshaw, N. S., Taylor, C. E., & Gordon, J. E. (1968). *Interactions of nutrition and infection* (WHO Monograph Series). Geneva, Switzerland: World Health Organization. Retrieved from http://apps.who.int/iris/handle/10665/41782

Selwyn, B. J. (1990). The epidemiology of acute respiratory tract infection in young children: Comparison of findings from several developing countries. *Reviews of Infectious Diseases, 12*, S870–S888.

Semba, R. D. (2007). Selenium. In K. Kraemer & M. B. Zimmermann (Eds.), *Nutritional anemia* (pp. 189–197). Basel, Switzerland: Sight and Life Press.

Semba, R. D., & Bloem, M. J. (2002). The anemia of vitamin A deficiency: Epidemiology and pathogenesis. *European Journal of Clinical Nutrition*, *56*, 271–281.

Semba, R. D., de Pee, S., Sun, K., Sari, M., Akhter, N., & Bloem, M. W. (2008). Effect of parental formal education on risk of child stunting in Indonesia and Bangladesh: A cross-sectional study. *The Lancet, 371*(9609), 322–328.

Seto, K. C., & Ramankutty, N. (2016). Hidden linkages between urbanization and food systems. *Science, 352*(6288), 943–945.

Shamim, A. A ., Schulze, K., Merrill, R. D., Kabir, A., Christian, P., & Shaikh, S. (2015). First-trimester plasma tocopherols are associated with risk of miscarriage in rural Bangladesh. *American Journal of Clinical Nutrition, 101,* 294–301.

Shankar, A., Genton, B., Semba, R. D., Maisor, M., Paino, J., Tamja, S., . . . West, K. P., Jr. (1999). Effect of vitamin A supplementation on morbidity due to *Plasmodium falciparum* in young children in Papua New Guinea: A randomised trial. *The Lancet, 354*, 203–209.

Shankar, A. H., & Prasad, A. S. (1998). Zinc and immune function: The biological basis of altered resistance to infection. *American Journal of Clinical Nutrition, 68*, 447S–463S.

Shann, F., Barker, J., & Poore, P. (1989). Clinical signs that predict death in children with severe pneumonia. *Pediatric Infectious Disease Journal, 8*, 852–855.

Shete, V., & Quadro, L. (2013). Mammalian metabolism of β-carotene: Gaps in knowledge. *Nutrients, 5*, 4849–4868.

Siberry, G. K., Ruff, A. J., & Black, R. (2002). Zinc and human immunodeficiency virus infection. *Nutrition Research, 22*, 527–538.

Siega-Riz, A. M., Estrada Del Campo, Y., Kinlaw, A., Reinhart, G. A., Allen, L. H., Shahab-Ferdows, S., & Bentley, M. E. (2014). Effect of supplementation with a lipid-based nutrient supplement on the micronutrient status of children aged 6–18 months living in the rural region of Intibuca, Honduras. *Paediatric and Perinatal Epidemiology, 28*(3), 245–254.

Singh, V., & West, K. P., Jr. (2004). Vitamin A deficiency and xerophthalmia among school-aged children in southeastern Asia. *European Journal of Clinical Nutrition*, *58*, 1342–1349.

Smit, E. N., Koopmann, M., Boersma, E. R., & Muskiet, F. A. (2000). Effect of supplementation of arachidonic acid (AA) or a combination of AA plus docosahexaenoic acid on breastmilk fatty acid composition. *Prostaglandins, Leukotrienes, and Essential Fatty Acids, 62*(6), 335–340.

Smitasiri, S., Attig, G. A., & Dhanamitta, S. (1992). Participatory action for nutrition education: Social marketing vitamin A–rich foods in Thailand. *Ecology of Food and Nutrition, 28*, 199–210.

Smith, E. R., Shankar, A. H., Wu, L. S., Aboud, S., Adu-Afarwuah, S., Ali, H., & Sudfeld, C. R. (2017). Modifiers of the effect of maternal multiple micronutrient supplementation on stillbirth, birth outcomes, and infant mortality: A meta-analysis of individual patient data from 17 randomised trials in low-income and middle-income countries. *The Lancet Global Health, 5*(11), e1090–e1100.

Smith, T. A., Lehman, D., Coakley, C., Spooner, V., & Alpers, M. P. (1991). Relationships between growth and acute lower-respiratory infections in children < 5 years in a highland population of Papua New Guinea. *American Journal of Clinical Nutrition, 53*, 963–970.

Solomons, N. W. (1999). Malnutrition: Secondary malnutrition. In M. J. Sadler, J. J. Strain, & B. Caballero (Eds.), *Encyclopedia of human nutrition* (pp. 1254–1259). San Diego, CA: Academic Press.

Solon, F. S. (2006). Good governance for nutrition in the Philippines: Elements, experiences, and lessons learned. *Food and Nutrition Bulletin, 27*(4), 343–352.

Sommer, A. (1982). *Nutritional blindness: Xerophthalmia and keratomalacia.* New York, NY: Oxford University Press.

Sommer, A. (1995). *Vitamin A deficiency and its consequences: A field guide to detection and control* (3rd ed.). Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/nutrition/publications/vad_consequences.pdf

Sommer, A., Hussaini, G., Tarwotjo, I., & Susanto, D. (1983). Increased mortality in children with mild vitamin A deficiency. *The Lancet, 2*, 585–588.

Sommer, A., & West, K. P., Jr. (1996). *Vitamin A deficiency: Health, survival, and vision.* New York, NY: Oxford University Press. Retrieved from https://global.oup.com/academic/product/vitamin-a-deficiency-9780195088243?cc=us&lang=en

Soofi, S., Ariff, S., Sadiq, K., Habib, A., Bhatti, Z., Ahmad, I., . . . Bhutta, Z. A. (2017). Evaluation of the uptake and impact of neonatal vitamin A supplementation delivered through the Lady Health Worker programme on neonatal and infant morbidity and mortality in rural Pakistan: An effectiveness trial. *Archives of Disease in Childhood, 102*, 216–223.

Spears, D., Ghosh, A., & Cumming, O. (2013). Open defecation and childhood stunting in India: An ecological analysis of new data from 112 districts. *PLoS One, 8*(9), e73784.

Stein, S. A. (1994). Molecular and neuroanatomical substrates of motor and cerebral cortex abnormalities in fetal thyroid hormone disorders. In J. B. Stanbury (Ed.), *The damaged brain of iodine deficiency* (pp. 67–102). New York, NY: Cognizant Communication.

Stevens, G. A., Bennett, J. E., Hennocq, Q., Lu, Y., De-Regil, L. M., Danaei, G., . . . Ezzati, M. (2015). Trends and mortality effects of vitamin A deficiency in children in 138 low-income and middle-income countries between 1991 and 2013: A pooled analysis of population-based surveys. *The Lancet Global Health*, *3*, e528–e536.

Stevens, G. A., Finucane, M. M., De-Regil, L. M., Paciorek, C. J., Flaxman, S. R., Branca, F., & Nutrition Impact Model Study Group. (2013). Global, regional, and national trends in haemoglobin concentration and prevalence of total and severe anaemia in children and pregnant and non-pregnant women for 1995–2011: A systematic analysis of population-representative data. *The Lancet Global Health, 1*(1), e16–e25.

Stipanuk, M. H. (2000). *Biochemical and physiological aspects of human nutrition.* Philadelphia, PA: W. B. Saunders.

Stoltzfus, R. J., Mullany, L., & Black, R. E. (2004). Iron deficiency anaemia. In M. Ezzati, A. D. Lopez, A. Rodgers, & C. J. L. Murray (Eds.), *Comparative quantification of health risks: Global and regional burden of disease attributable to selected major risk factors* (Vol. 1, pp. 163–209). Geneva, Switzerland: World Health Organization.

Subramanyam, M. A., Kawachi, I., Berkman, L. F., & Subramanian, S. V. (2011). Is economic growth associated with reduction in child undernutrition in India? *PLoS Medicine, 8*(3), e1000424.

Sudfeld, C. R., McCoy, D. C., Danaei, G., Fink, G., Ezzati, M., Andrews, K. G., & Fawzi, W. W. (2015). Linear growth and child development in low- and middle-income countries: A meta-analysis. *Pediatrics, 135*(5), e1266–e1275.

SUMMIT Study Group (Supplementation with Multiple Micronutrients Intervention Trial Study Group). (2008). Effect of maternal multiple micronutrient supplementation on fetal loss and infant death in Indonesia: A double-blind cluster-randomised trial. *The Lancet, 371,* 215–227.

Sununtnasuk, C., D'Agostino, A., & Fiedler, J. L. (2016). Iron + folic acid distribution and consumption through antenatal care: Identifying barriers across countries. *Public Health Nutrition, 19*(4), 732–742.

Swinburn, B., Dominick, C., & Vandevijvere, S. (2014). *Benchmarking food environments: Experts' assessments of policy gaps and priorities for the New Zealand government.* Auckland, NZ: University of Auckland, Faculty of Medical and Health Sciences, School of Population Health.

Swinburn, B., Kraak, V., Rutter, H., Vandevijvere, S., Lobstein, T., Sacks, G., & Magnusson, R. (2015). Strengthening of accountability systems to create healthy food environments and reduce global obesity. Retrieved from http://researchonline.lshtm.ac.uk/2121411/.

Swinburn, B., Sacks, G., Vandevijvere, S., Kumanyika, S., Lobstein, T., Neal, B., & Walker, C. (2013). INFORMAS (International Network for Food And Obesity/Non-communicable Diseases Research, Monitoring and Action Support): Overview and key principles. *Obesity Reviews, 14*(S1), 1–12.

Talukder, A., Islam, N., Klemm, R., & Bloem, M. (1993). *Home gardening in South Asia: The complete handbook.* Dhaka, Bangladesh: Helen Keller International. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=XF2015009484

Tanumihardjo, S. A., Russell, R. M., Stephensen, C. B., Gannon, B. M., Craft, N. E., Haskell, M. J., . . . Raiten, D. J. (2016). Biomarkers of nutrition for development (BOND): Vitamin A review. *Journal of Nutrition, 146*(suppl), 1816S–1848S.

Taylor-Robinson, D. C., Maayan, N., Soares-Weiser, K., Donegan, S., & Garner, P. (2015). Deworming drugs for soil-transmitted intestinal worms in children: Effects on nutritional indicators, haemoglobin, and school performance. *Cochrane Database of Systematic Reviews, 7*, CD000371. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD000371.pub6/abstract;jsessionid=EC7D26DF40BF16D3685A4A87BD4521EC.f01t04

te Lintelo, D. J. H., & Lakshman, R. W. D. (2015). *Evidence Report No. 150: Reducing hunger and undernutrition: The hunger and nutrition commitment index (HANCI).* Sussex, UK: Institute of Development Studies.

Thapa, S., Choe, M. K., & Retherford, R. D. (2005). Effects of vitamin A supplementation on child mortality: Evidence from Nepal's 2001 Demographic and Health Survey. *Tropical Medicine and International Health, 10*, 782–789.

Tharrey, M., Olaya, G. A., Fewtrell, M., & Ferguson, E. (2017). Adaptation of new Colombian food-based complementary feeding recommendations using linear programming. *Journal of Pediatric Gastroenterology and Nutrition*. [Epub ahead of print]. doi:10.1097/MPG.0000000000001662

Thompson, J., Biggs, B. A., & Pasricha, S. R. (2013). Effects of daily iron supplementation in 2- to 5-year-old children: Systematic review and meta-analysis. *Pediatrics, 131*(4), 739–753.

Thornton, K. A., Mora-Plazas, M., Marin, C., & Villamor, E. (2014). Vitamin A deficiency is associated with gastrointestinal and respiratory morbidity in school-age children. *Journal of Nutrition, 144*, 496–503.

Thurnham, D. I., Mburu, A. S., Mwaniki, D. L., & De Wagt, A. (2005). Micronutrients in childhood and the influence of subclinical inflammation. *Proceedings of the Nutrition Society, 64*, 502–509.

Thurnham, D. I., McCabe, L. D., Haldar, S., Wieringa, F. T., Northrop-Clewes, C. A., & McCabe, G. P. (2010). Adjusting plasma ferritin concentrations to remove the effects of subclinical inflammation in the assessment of iron deficiency: A meta-analysis. *American Journal of Clinical Nutrition, 92*(3), 546–555.

Thurnham, D. I., & Northrop-Clewes, C. A. (2007). Infection and the etiology of anemia. In K. Kraemer & M. B. Zimmerman (Eds.), *Nutritional anemia* (pp. 231–256). Basel, Switzerland: Sight and Life Press.

Tielsch, J. M., Khatry, S. K., Stoltzfus, R. J., Katz, J., LeClerq, S. C., Adhikari, R., . . . Shresta, S. (2007). Effect of daily zinc supplementation on child mortality in southern Nepal: A community-based, cluster randomized, placebo-controlled trial. *The Lancet, 370*, 1230–1239.

Tielsch, J. M., Rahmathullah, L., Thulasiraj, R. D., Katz, J., Coles, C., Sheeladevi, S., . . . Prakash, K. (2007). Newborn vitamin A dosing reduces the case fatality but not incidence of common childhood morbidities in South India. *Journal of Nutrition, 137*, 2470–2474.

Tilman, C., & Clark, M. (2014). Global diets link environmental sustainability and human health. *Nature, 515*(7528), 518–522.

Timmer, C. B., Falcon, W. P., & Pearson, S. R. (1983). *Food policy analysis.* Baltimore, MD: Johns Hopkins University Press.

Tirivayi, N., Knowles, M., & Davis, B. (2016). The interaction between social protection and agriculture: A review of evidence. *Global Food Security, 10*, 52–62.

Tome, P., Reyes, H., Rodriguez, L., Guiscafre, H., & Gutierrez, G. (1996). Death caused by acute diarrhea in children: A study of prognostic factors. *Salud Publica de Mexico, 38*, 227–235.

Torheim, L. E., Ferguson, E. L., Penrose, K., & Arimond, M. (2010). Women in resource-poor settings are at risk of inadequate intakes of multiple micronutrients. *Journal of Nutrition, 140,* 2051S–2058S.

Trehan, I., Goldbach, H. S., LaGrone, L. N., Meuli, G. J., Wang, R. J., Maleta, K. M., & Manary, M. J. (2013). Antibiotics as part of the management of severe acute malnutrition. *New England Journal of Medicine, 368*(5), 425–435.

Trehan, I., & Manary, M. J. (2015). Management of severe acute malnutrition in low-income and middle-income countries. *Archives of Disease in Childhood, 100*(3), 283–287.

Trowbridge, F. L., Harris, S. S., Cook, J., Dunn, J. T., Florentino, R. F., Kodyat, B. A., . . . Yip, R. (1993). Coordinated strategies for controlling micronutrient malnutrition: A technical workshop. *Journal of Nutrition, 123*, 775–787.

Tupasi, T. E., Mangubat, N. V., Sunico, M. E., Magdangal, D. M., Navarro, E. E., Leonor, Z. A., . . . Lucero, M. G. (1990). Malnutrition and acute respiratory tract infections in Filipino children. *Review of Infectious Diseases, 12*, S1047–S1054.

Uauy, R., Kain, J., & Corvalan, C. (2011). How can the Developmental Origins of Health and Disease (DOHaD) hypothesis contribute to improving health in developing countries? *American Journal of Clinical Nutrition, 94*(6 suppl), 1759S–1764S.

United Nations (UN). (2015b). *Trends in contraceptive use worldwide—2015. Economic and social affairs.* New York, NY: Author. Retrieved from http://www.un.org/en/development/desa/population/publications/pdf/family/trendsContraceptiveUse2015Report.pdf

United Nations Children's Fund (UNICEF). (1998). *The state of the world's children 1998.* New York, NY: Oxford University Press.

United Nations Children's Fund (UNICEF). (2015). *Levels and trends in child mortality—2015.* New York, NY: Author. Retrieved from http://www.who.int/maternal_child_adolescent/documents/levels_trends_child_mortality_2015/en/

United Nations Children's Fund (UNICEF) & Centers for Disease Control and Prevention (CDC). (2013). *Global assessment of home fortification interventions, 2011*. Geneva, Switzerland: Home Fortification Technical Advisory Group. Retrieved from http://www.incap.int/index.php/es/publicaciones/publicaciones-externas/doc_view/388-global-assessment-of-home-fortification-interventions-2011

United Nations Children's Fund (UNICEF), World Health Organization (WHO), & World Bank. (2017). *Levels and trends in child malnutrition. UNICEF/WHO/World Bank group joint child malnutrition estimates: Key findings of the 2017 edition.* New York, NY: UNICEF. Retrieved from http://www.who.int/nutgrowthdb/jme_brochoure2017.pdf

Urwin, H. J., Zhang, J., Gao, Y., Wang, C., Li, L., Song, P., & Yaqoob, P. (2013). Immune factors and fatty acid composition in human milk from river/lake, coastal and inland regions of China. *British Journal of Nutrition, 109*(11), 1949–1961.

van den Bold, M., Kohli, N., Gillespie, S., Zuberi, S., Rajeesh, S., & Chakraborty, B. (2015). Is there an enabling environment for nutrition-sensitive agriculture in South Asia? Stakeholder perspectives from India, Bangladesh, and Pakistan. *Food and Nutrition Bulletin, 36*(2), 231–247.

van den Bold, M., Quisumbing, A. R., & Gillespie, S. (2013). *Women's empowerment and nutrition: An evidence review.* Washington, DC: International Food Policy Research Institute. Retrieved from http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/127840

Verstraeten, R., Roberfroid, D., Lachat, C., Leroy, J. L., Holdsworth, M., Maes, L., & Kolsteren, P. W. (2012). Effectiveness of preventive school-based obesity interventions in low- and middle-income countries: A systematic review. *American Journal of Clinical Nutrition*, *96*(2), 415–438.

Victora, C. G., Bahl, R., Barros, A. J., Franca, G. V., Horton, S., Krasevec, J., & *The Lancet* Breastfeeding Series Group. (2016). Breastfeeding in the 21st century: Epidemiology, mechanisms, and lifelong effect. *The Lancet, 387*(10017), 475–490.

Victora, C. G., Barros, F. C., Kirkwood, B. R., & Vaughan, J. P. (1990). Pneumonia, diarrhea, and growth in the first 4 years of life: A longitudinal study of 5914 urban Brazilian children. *American Journal of Clinical Nutrition, 52*, 391–396.

Victora, C. G., de Onis, M., Hallal, P. C., Blossner, M., & Shrimpton, R. (2010). Worldwide timing of growth faltering: revisiting implications for interventions. *Pediatrics, 125*(3), e473–e480.

Victora, C. G., Kirkwood, B. R., Ashworth, A., Black, R. E., Rogers, S., Sazawal, S., . . . Gove, S. (1999). Potential interventions for the prevention of childhood pneumonia in developing countries: Improving nutrition. *American Journal of Clinical Nutrition, 70*, 309–320.

Victora, C. G., Smith, P. G., Barros, F. C., Vaughan, J. P., & Fuchs, S. C. (1989). Risk factors for deaths due to respiratory infections among Brazilian infants. *International Journal of Epidemiology, 18*, 918–925.

Villar, J., Cheikh Ismail, L., Victora, C. G., Ohuma, E. O., Bertino, E., Altman, D. G., & Newborn Growth Consortium for the 21st Century. (2014). International standards for newborn weight, length, and head circumference by gestational age and sex: The Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. *The Lancet, 384*(9946), 857–868.

Viner, R. M., Coffey, C., Mathers, C., Bloem, P., Costello, A., Santelli, J., & Patton, G. C. (2011). 50-year mortality trends in children and young people: A study of 50 low-income, middle-income, and high-income countries. *The Lancet, 377*(9772), 1162–1174.

Vitamin A and Pneumonia Working Group. (1995). Potential interventions for the prevention of childhood pneumonia in developing countries: A meta-analysis of data from field trials to assess the impact of vitamin A supplementation on pneumonia morbidity and mortality. *Bulletin of the World Health Organization, 73*, 609–619.

Vosti, S. A., Kagin, J., Engle-Stone, R., & Brown, K. H. (2015). An economic optimization model for improving the efficiency of vitamin A interventions: An application to young children in Cameroon. *Food and Nutrition Bulletin, 36*(3 suppl), S193–S207.

Walker, S. P., Wachs, T. D., Gardner, J. M., Lozoff, B., Wasserman, G. A., Pollitt, E., & Carter, J. A. (2007). Child development: risk factors for adverse outcomes in developing countries. *The Lancet, 369*(9556), 145–157.

Wamani, H., Tylleskär, T., Åstrøm, A. N., Tumwine, J. K., & Peterson, S. (2004). Mothers' education but not fathers' education, household assets or land ownership is the best predictor of child health inequalities in rural Uganda. *International Journal for Equity in Health, 3*(1), 9. doi:10.1186/1475-9276-3-9

Waters, E., de Silva-Sanigorski, A., Burford, B. J., Brown, T., Campbell, K. J., Gao, Y., & Summerbell, C. D. (2011). Interventions for preventing obesity in children. *Cochrane Database of Systematic Reviews, 7*(12), CD001871.

Weissenbacher, M., Carballal, G., Avila, M., Salomon, H., Harisiadi, J., Catalano, M., . . . Murtagh, P. (1990). Etiologic and clinical evaluation of acute lower respiratory tract infections in young Argentinian children: An overview. *Review of Infectious Diseases, 12*, S889–S898.

Welch, V. A., Ghogumu, E., Hossain, A., Awasthi, S., Bhutta, Z. A., Cumberbatch, C., & Wells, G. A. (2017). Mass deworming to improve developmental health and wellbeing of children in low-income and middle-income countries: A systematic review and network meta-analysis. *The Lancet Global Health, 5*(1), e40–e50.

Wessells, K. R., & Brown, K. H. (2012). Estimating the global prevalence of zinc deficiency: Results based on zinc availability in national food supplies and the prevalence of stunting. *PLoS One, 7*(11), e50568. doi:10.1371/journal.pone.0050568

West, K. P., Jr., Christian, P., Labrique, A. B., Rashid, M., Shamim, A. A., Klemm, R. D.W., . . . Sommer, A. (2011). Effects of vitamin A or beta-carotene supplementation on pregnancy-related mortality and infant mortality in rural Bangladesh: A cluster-randomized trial. *Journal of the American Medical Association*, *305*, 1986–1995.

West, K .P., Jr., Katz, J., Khatry, S. K., LeClerq, S. C., Pradhan, E. K., Shrestha, S.R ., . . . Sommer, A. (1999). Double blind, cluster randomised trial of low dose supplementation with vitamin A or β-carotene on mortality related to pregnancy in Nepal. *British Medical Journal, 318*, 570–575.

West, K. P., Jr., Pokhrel, R. P., Katz, J., LeClerq, S .C., Khatry, S. K., Shrestha, S. R., . . . Sommer, A. (1991). Efficacy of vitamin A in reducing preschool child mortality in Nepal. *The Lancet*, *338*, 67–71.

West, K.P., Jr., Shamim, A. A., Mehra, S., Labrique, A. B., Ali, H., Shaikh, S., . . . Christian, P. (2014). Effect of maternal multiple micronutrient versus iron–folic acid supplementation on infant mortality and adverse birth outcomes in rural Bangladesh: The JiVitA-3 randomized trial. *Journal of the American Medical Association, 312*, 2649–2658.

West, K. P., Jr., Sommer, A., Palmer, A. C., Schultink, W., & Habicht, J. P. (2015). Commentary: Vitamin A policies need rethinking. *International Journal of Epidemiology*, *44*(1), 292–294; discussion 294–296.

Williams, C. D. (1933). A nutritional disease of childhood associated with a maize diet. *Archives Diseases in Childhood, 8*(48), 423–433.

Wirth, J. P., Petry, N., Tanumihardjo, S. A., Rogers, L. M., McLean, E., Greig, A., . . . Rohner, F. (2017). Vitamin A supplementation programs and country-level evidence of vitamin A deficiency. *Nutrients, 9*(3). pii: E190. doi:10.3390/nu9030190

Wirth, J. P., Woodruff, B. A., Engle-Stone, R., Namaste, S. M., Temple, V. J., Petry, N., & Aaron, G. J. (2017). Predictors of anemia in women of reproductive age: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *American Journal of Clinical Nutrition, 106*(suppl 1), 416S–427S.

Wondafrash, M., Huybregts, L., Lachat, C., Bouckaert, K. P., & Kolsteren, P. (2016). Dietary diversity predicts dietary quality regardless of season in 6–12-month-old infants in south-west Ethiopia. *Public Health Nutrition, 19*(14), 2485–2494.

World Food Programme (WFP). (2017). *Fill the Nutrient Gap Tool: Analysis, consensus, decision-making.* Rome, Italy: Author. Retrieved from http://documents.wfp.org/stellent/groups/public/documents/communications/wfp288102.pdf

World Health Organization (WHO). (1993). *Global prevalence of iodine deficiency disorders*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/nutrition/publications/micronutrients/iodine_deficiency/54015_mdis_workingpaper1/en/

World Health Organization (WHO). (2006). *Implementing the new recommendations on the clinical management of diarrhoea: Guidelines for policy makers and program managers*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/maternal_child_adolescent/documents/9241594217/en/

World Health Organization (WHO). (2007). *Assessing the iron status of populations*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/nutrition/publications/micronutrients/anaemia_iron_deficiency/9789241596107/en/

World Health Organization (WHO). (2008a). *Indicators for assessing infant and young child feeding practices: Part 1: Definitions*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/maternal_child_adolescent/documents/9789241596664/en/

World Health Organization (WHO). (2008b). *Indicators for assessing infant and young child feeding practices: Part 3: Country profiles*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/nutrition/publications/infantfeeding/9789241599757/en/

World Health Organization (WHO). (2009). *Global prevalence of vitamin A deficiency in populations at risk 1995–2005. WHO Global Database on Vitamin A Deficiency.* Geneva, Switzerland: Author. Retrieved from http://apps.who.int/iris/bitstream/10665/44110/1/9789241598019_eng.pdf

World Health Organization (WHO). (2010a). *Increasing access to health workers in remote and rural areas through improved retention: Policy recommendations.* Geneva, Switzerland: Author. Retrieved from http://www.searo.who.int/nepal/mediacentre/2010_increasing_access_to_health_workers_in_remote_and_rural_areas.pdf

World Health Organization (WHO). (2010b). *Nutrition Landscape Information System (NLIS) country profile indicators: Interpretation guide.* Geneva, Switzerland: Author. Retrieved from http://www.who.int/nutrition/nlis_interpretationguide_isbn9789241599955/en/

World Health Organization (WHO). (2011a). Serum retinol concentrations for determining the prevalence of vitamin A deficiency in populations. Vitamin and Mineral Information System. Geneva, Switzerland: Author. Retrieved from http://www.who.int/vmnis/indicators/retinol.pdf

World Health Organization (WHO). (2011b). *Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/vmnis/indicators/haemoglobin/en/

World Health Organization (WHO). (2011c). *Unsafe abortion: Global and regional estimates of the incidence of unsafe abortion and associated mortality in 2008.* Geneva, Switzerland: Author. Retrieved from http://www.who.int/reproductivehealth/publications/unsafe_abortion/9789241501118/en/

World Health Organization (WHO). (2011d). *Vitamin A supplementation for infants and children 6-59 months of age.* Geneva, Switzerland: Author. Retrieved from http://www .who .int/nutrition/publications/micronutrients/guidelines/vas_6to59_months/en/

World Health Organization (WHO). (2013a). *WHO recommendations on postnatal care of the mother and newborn.*

Geneva, Switzerland: Author. Retrieved from http://www.who.int/maternal_child_adolescent/documents/postnatal-care-recommendations/en/

World Health Organization (WHO). (2013b). *Guideline: Updates on the management of severe acute malnutrition in infants and children*. Geneva, Switzerland: Author. Retrieved from http://apps.who.int/iris/bitstream/10665/95584/1/9789241506328_eng.pdf

World Health Organization (WHO). (2014a). *Global nutrition targets 2025: Low birth weight policy brief*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/nutrition/publications/globaltargets2025_policybrief_lbw/en/

World Health Organization (WHO). (2014b). *Global nutrition targets 2025: Policy brief series*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/nutrition/publications/globaltargets2025_policybrief_overview/en/

World Health Organization (WHO). (2014c). *Global status report on noncommunicable diseases*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/nmh/publications/ncd-status-report-2014/en/

World Health Organization (WHO). (2015a). *The global prevalence of anaemia in 2011*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/nutrition/publications/micronutrients/global_prevalence_anaemia_2011/en/

World Health Organization (WHO). (2015b). *Healthy diet*. Fact Sheet No. 394. Geneva, Switzerland: WHO Media Centre. Retrieved from http://www.who.int/mediacentre/factsheets/fs394/en/

World Health Organization (WHO). (2016a). *Guideline: Daily iron supplementation in infants and children*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/nutrition/publications/micronutrients/guidelines/daily_iron_supp_childrens/en/

World Health Organization (WHO). (2016b). *Guideline: Use of multiple micronutrient powders for point-of-use fortification of foods consumed by infants and young children aged 6–23 months and children aged 2–12 years*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/nutrition/publications/micronutrients/guidelines/mmpowders-infant6to23mons-children2to12yrs/en/

World Health Organization (WHO). (2016c). *Obesity and overweight*. Fact Sheet. Geneva, Switzerland: Author. Retrieved from http://www.who.int/mediacentre/factsheets/fs311/en/

World Health Organization (WHO). (2017). *Malnutrition*. Fact Sheet. Geneva, Switzerland: Author. Retrieved from http://www.who.int/mediacentre/factsheets/malnutrition/en/

World Health Organization (WHO) & United Nations Children's Fund (UNICEF). (2009). *Baby-Friendly Hospital Initiative: Revised, updated and expanded for integrated care*. Geneva, Switzerland: WHO. Retrieved from http://www.who.int/nutrition/publications/infantfeeding/bfhi_trainingcourse/en/

World Health Organization (WHO), United Nations Children's Fund (UNICEF), & International Baby Food Action Network (IBFAN). (2016). *Marketing of breast-milk substitutes: National implementation of the International Code, Status Report 2016*. Geneva, Switzerland: WHO. Retrieved from http://www.who.int/nutrition/publications/infantfeeding/code_report2016/en/

World Health Organization (WHO), United Nations Children's Fund (UNICEF), & International Council for Control of Iodine Deficiency Disorders (ICCIDD). (2007). *Assessment of iodine deficiency disorders and monitoring their elimination: A guide for program managers* (3rd ed.). Geneva, Switzerland: WHO. Retrieved from http://apps.who.int/iris/bitstream/10665/43781/1/9789241595827_eng.pdf

World Health Organization (WHO), World Food Programme (WFP), United Nations System Standing Committee on Nutrition, & United Nations Children's Fund (UNICEF). (2007). *Community-based management of severe acute malnutrition*. Geneva, Switzerland: WHO. Retrieved from http://www.who.int/nutrition/topics/Statement_community_based_man_sev_acute_mal_eng.pdf

World Health Organization Multicentre Growth Reference Study Group. (2006a). WHO child growth standards based on length/height, weight and age. *Acta Paediatrica, 450*(suppl), 76–85.

World Health Organization Multicentre Growth Reference Study Group. (2006b). Assessment of differences in linear growth among populations in the WHO Multicentre Growth Reference Study. *Acta Paediatrica, 450*(suppl), 56–65.

World Health Organization Secretariat. (2007). Conclusions and recommendations of the WHO consultation on prevention and control of iron deficiency in infants and young children in malaria-endemic areas. *Food and Nutrition Bulletin, 28*(4), S621–S631.

Xu, J., Luntamo, M., Kulmala, T., Ashorn, P., & Cheung, Y. B. (2014). A longitudinal study of weight gain in pregnancy in Malawi: Unconditional and conditional standards. *American Journal of Clinical Nutrition, 99*(2), 296–301.

Yoong, J., Rabinovich, L., & Diepeveen, S. (2012). *The impact of economic resource transfers to women versus men: A systematic review*. Technical Report. London, UK: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Yount, K., VanderEnde, K., Dodell, S., & Cheong, Y. (2016). Measurement of women's agency in Egypt: A national validation study. *Social Indicators Research, 128*(3), 1171–1192.

Zakiyah, N., van Asselt, D. I., Roijmans, F., & Postma, M. J. (2016). Economic evaluation of family planning interventions in low and middle income countries; A systematic review. *PLoS One, 11*(12), e0168447.

Zimmermann, M. B. (2004). Assessing iodine status and monitoring progress of iodized salt programs. *Journal of Nutrition, 134*(7), 1673–1677.

Zimmermann, M. B. (2012). The effects of iodine deficiency in pregnancy and infancy. *Paediatric and Perinatal Epidemiology, 26*(suppl 1), 108–117.

Zimmermann, M. B., & Andersson, M. (2012). Assessment of iodine nutrition in populations: Past, present, and future. *Nutrition Reviews, 70*(10), 553–570.

Zinc Investigators' Collaborative Group. (1999). Prevention of diarrhea and pneumonia by zinc supplementation in children in developing countries: Pooled analysis of randomized controlled trials. *Journal of Pediatrics, 135*, 689–697.

Zoghbi, W. A., Duncan, T., Antman, E., Barbosa, M., Champagne, B., Chen, D., & Komajda, M. (2014). Sustainable Development Goals and the future of cardiovascular health: A statement from the Global Cardiovascular Disease Taskforce. *Journal of the American Heart Association, 3*(5), e000504.

Zumrawi, F. Y., Dimond, H., & Waterlow, J. C. (1987). Effects of infection on growth in Sudanese children. *Human Nutrition: Clinical Nutrition, 41*, 453–461.

CHAPTER 8

Chronic Diseases and Risks

Gillian P. Christie, Sandeep P. Kishore, David J. Heller, and Derek Yach

DISCLOSURE STATEMENT

Statements made and opinions expressed in this chapter are those of the authors and should not be construed as representing an official position of the organizations where the authors are employed.

Introduction

In the 1920s, Yale University's Charles-Edward Amory Winslow taught that the goal of each generation was to redefine the unacceptable. With chronic diseases now arguably "the social justice issue of our generation," this chapter intends to redefine the unacceptable by building a renewed case on the imperative to invest in and advocate for chronic disease prevention (Horton, 2015). It calls for young professionals and continuing students to engage in sustained and coordinated actions to combat the growing burden of chronic disease on global populations.

This chapter covers three objectives: (1) define and outline the global prevalence of chronic diseases, their health and economic impact, and their four key risk factors; (2) describe stakeholder efforts to date in chronic disease prevention; and (3) propose stakeholder actions—both policies and systems of care—that can further control the rising prevalence of chronic diseases.

The Burden of Chronic Disease: Causes and Impacts

Defining Chronic Diseases

At its founding in 1946, the World Health Organization (WHO) defined health as "a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity." At that time, most human death and disease was due to acute, infectious diseases from which people either recovered or died. Today, chronic conditions—those that cause illness gradually over the course of many months or years—have become the leading cause of human disability and death. Chronic diseases are among the greatest threats to the state of global human health that WHO was founded to protect.

Chronic diseases often emerge in middle age, after a long exposure to adverse social, environmental, genetic, behavioral, and lifestyle factors. Such diseases can affect the young and healthy—including children—and can emerge quickly and without warning. For example, an adolescent newly diagnosed

with type 1 diabetes and a seemingly healthy young woman who suffers an acute stroke due to undiagnosed high blood pressure both suffer from chronic disease. In affluent countries as well as in low- and middle-income countries (LMICs), chronic diseases often disproportionately affect lower-income persons and communities. Although most chronic conditions are not "infectious" in the typical sense, emerging evidence suggests that chronic conditions such as diabetes and atherosclerotic cardiovascular disease tend to cluster in families and communities with common lifestyle and environmental risk factors. These risk factors can *transmit* disease from one person to another in a social unit through shifts in behavior.

Despite their many forms, most chronic diseases have common underlying characteristics. These include a few common risk factors that act independently and synergistically; a long latency between cumulative exposure to risk and disease outcomes; a high degree of preventability; a low cure rate, necessitating decades of treatment and care coordination; considerable comorbidity; and strong linkages to poverty and socioeconomic development. Chronic diseases are predominantly caused by noninfectious risk factors, but can be caused or exacerbated by infection as well. The leading chronic diseases include cardiovascular disease, stroke, chronic respiratory diseases, cancer, diabetes, human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS), mental illness, and traumatic injury. Fortunately, all of these conditions—and the morbidity and mortality they cause—are preventable and treatable through policies and programs that treat a few common risk factors, such as unhealthy diet, tobacco use, physical inactivity, and alcohol and other substance misuse.

Epidemiology of Chronic Diseases: Rising Risk

In the past 100 years, chronic and largely noncommunicable diseases have silently overtaken infectious diseases as the leading cause of death and disability worldwide (Stuckler, 2011). Initially proposed by Abdel Omran in 1971, this "epidemiologic transition" to chronic, noncommunicable diseases began in the early twentieth century in high-income countries such as the United States. Pneumonia, tuberculosis, and gastrointestinal infections were the leading killers in 1900, and by the early twenty-first century had spread to lower-income countries. This transition, which is largely due to improvements in sanitation, vaccination, and antibiotics that significantly curbed disease and death from infection, increased global life expectancy from 61.7 years in 1980 to 71.8 years in 2015. Nevertheless, achieving further advances in longevity will require a different approach, as human life expectancy now chiefly depends on the growing toll of chronic diseases (Global Burden of Disease [GBD] Collaborators, 2016). Each year, chronic diseases kill an estimated 39 million people, with cardiovascular diseases the leading cause of death (17.9 million people), followed by cancers (8.8 million people) and respiratory diseases (3.8 million people) (GBD Collaborators, 2016). Most of these deaths (approximately 28 million) occur in LMICs (WHO, 2017b). **TABLE 8-1** identifies the 10 leading causes of death globally based on income, while **TABLE 8-2** indicates changes in deaths from chronic diseases over time.

Individual Risk Factors for Chronic Disease. A few common, modifiable behavioral risk factors accumulate over the lifespan and drive the majority of all

TABLE 8-1 The Ten Leading Causes of Deaths in the World, and in Low-, Middle-, and High-Income Countries, 2015

Disease or Injury	Deaths (per 100,000)	Percentage Change (1990–2015)	Disease or Injury	Deaths (per 100,000)	Percentage Change (1990–2015)
World			*Low-Income Countries*		
1. Cardiovascular diseases	243.1	2.4	1. Lower respiratory infections	80.4	–52.5
2. Neoplasms	118.9	11.0	2. Malaria	64.5	–64.1
3. Diarrhea/lower respiratory infections/other	67.3	–55.4	3. Diarrheal diseases	59.0	–66.1
4. Chronic respiratory	51.5	–22.4	4. Ischemic heart disease	47.9	–4.6

Disease or Injury	Deaths (per 100,000)	Percentage Change (1990–2015)	Disease or Injury	Deaths (per 100,000)	Percentage Change (1990–2015)
5. Diabetes/urogenital/ blood/endocrine diseases	46.3	44.1	5. HIV/AIDS	44.5	76.4
6. HIV/AIDS and tuberculosis	31.3	–7.5	6. Tuberculosis	33.6	–43.9
7. Neurologic disorders	30.7	46.8	7. Hemorrhagic stroke	29.2	–18.0
8. Neonatal disorders	29.4	–55.7	8. Neonatal preterm birth complications	23.4	–51.2
9. Unintentional injuries	25.0	–31.5	9. Neonatal encephalopathy due to birth asphyxia and trauma	21.9	–43.2
10. Transport injuries	19.9	–12.7	10. Protein-energy malnutrition	21.6	–66.5
Middle-Income Countries			*High-Income Countries*		
1. Cardiovascular diseases	237.0	25.3	1. Ischemic heart disease	193.1	–16.0
2. Neoplasms	100.0	24.6	2. Alzheimer's disease	71.2	66.3
3. Chronic respiratory	61.3	–26.0	3. Ischemic stroke	65.5	–15.8
4. Diabetes/urogenital/ blood/endocrine diseases	46.4	59.0	4. Lung cancer	51.4	12.2
5. Diarrhea/lower respiratory infections/other	42.2	–64.9	5. Lower respiratory infections	39.7	30.2
6. HIV/AIDS and tuberculosis	24.4	–29.1	6. Chronic obstructive pulmonary disease	37.4	12.1
7. Transport injuries	23.8	–2.0	7. Hemorrhagic stroke	35.4	–15.8
8. Unintentional injuries	22.5	–42.8	8. Colorectal cancer	28.8	15.6
9. Neonatal disorders	19.2	–69.2	9. Diabetes	18.9	17.6
10. Cirrhosis	18.7	3.1	10. Self-harm	18.2	1.1

Data from Institute of Health Metrics and Evaluation (IHME) at the University of Washington, 2015.

TABLE 8-2 Changes in Deaths from Selected Chronic Diseases, 2005–2015

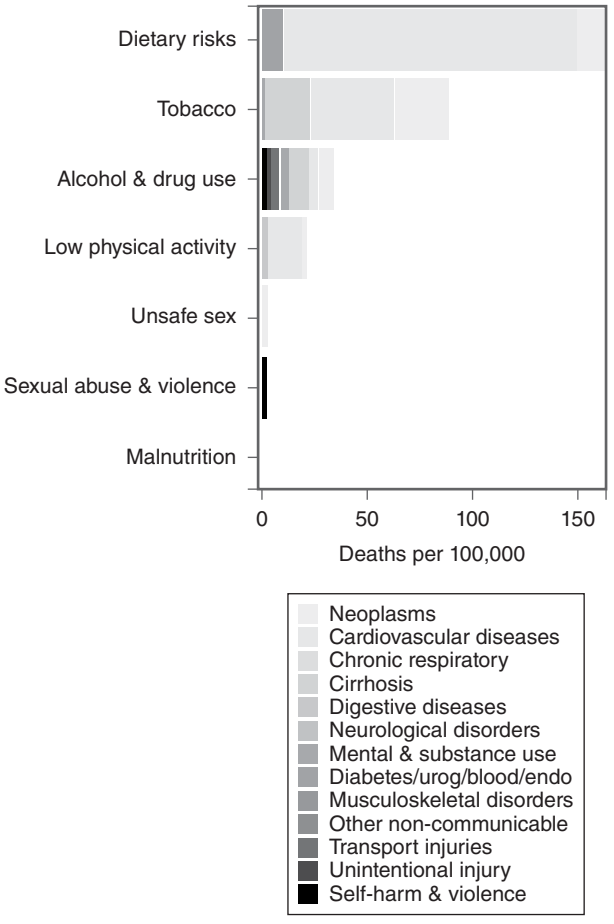
Chronic Disease	Deaths, 2005 (thousands)	Deaths, 2015 (thousands)	Percent Change (%)
Overall	34,835.6	39,804.2	14.3
Neoplasms	7,492.8	8,764.6	17.0
Cardiovascular diseases	15,933.7	17,921.0	12.5
Chronic respiratory diseases	3,709.1	3,795.5	2.3
Neurologic disorders	1,671.0	2,258.9	35.2
Mental and substance use disorders	305.9	324.9	6.2
Drug use disorders	128.8	169.9	31.8
Diabetes, urogenital, blood, and endocrine diseases	2,635.3	3,409.3	29.4
Musculoskeletal disorders	76.2	90.1	18.2

Note: Between 2005 and 2015, global deaths from chronic diseases increased by 14.3%. The greatest changes were in neurologic disorders; drug use disorders; and diabetes, urogenital, blood, and endocrine diseases.

chronic disease, from diabetes to cancer to cardiovascular conditions (**FIGURE 8-1** and **FIGURE 8-2**). These include unhealthy diet, tobacco use, abuse of alcohol and other substances, and low physical activity (**TABLE 8-3**). Chronic disease is also preventable through adherence to prescription medications and promotion of mental well-being, among other factors. Collectively, these risk factors drive a substantial portion (up to 60%) of the global burden of disease, both directly and by causing other risk factors for chronic disease such as elevated cholesterol, glucose, and blood pressure levels.

To explain the shift in chronic diseases from affluent countries, and predict its onset in LMICs, Omran conceptualized three stages of disease epidemiology: pestilence and famine, receding epidemics, and degenerative and human-made chronic diseases. Three major sets of determinants drive the transition toward the third "chronic disease" stage in each country: ecobiologic changes; socioeconomic, political, and cultural changes; and medical and public health interventions. Omran correctly predicted that chronic diseases would eventually impose a greater burden in all countries in the coming decades. In practice, however, chronic diseases have not displaced acute infectious diseases in LMICs. Rather, these countries suffer from a polarized and protracted double, triple, or quadruple burden of disease (Frenk, Bobadilla, Sepúlveda & Cervantes, 1989; Mayosi et al., 2009).

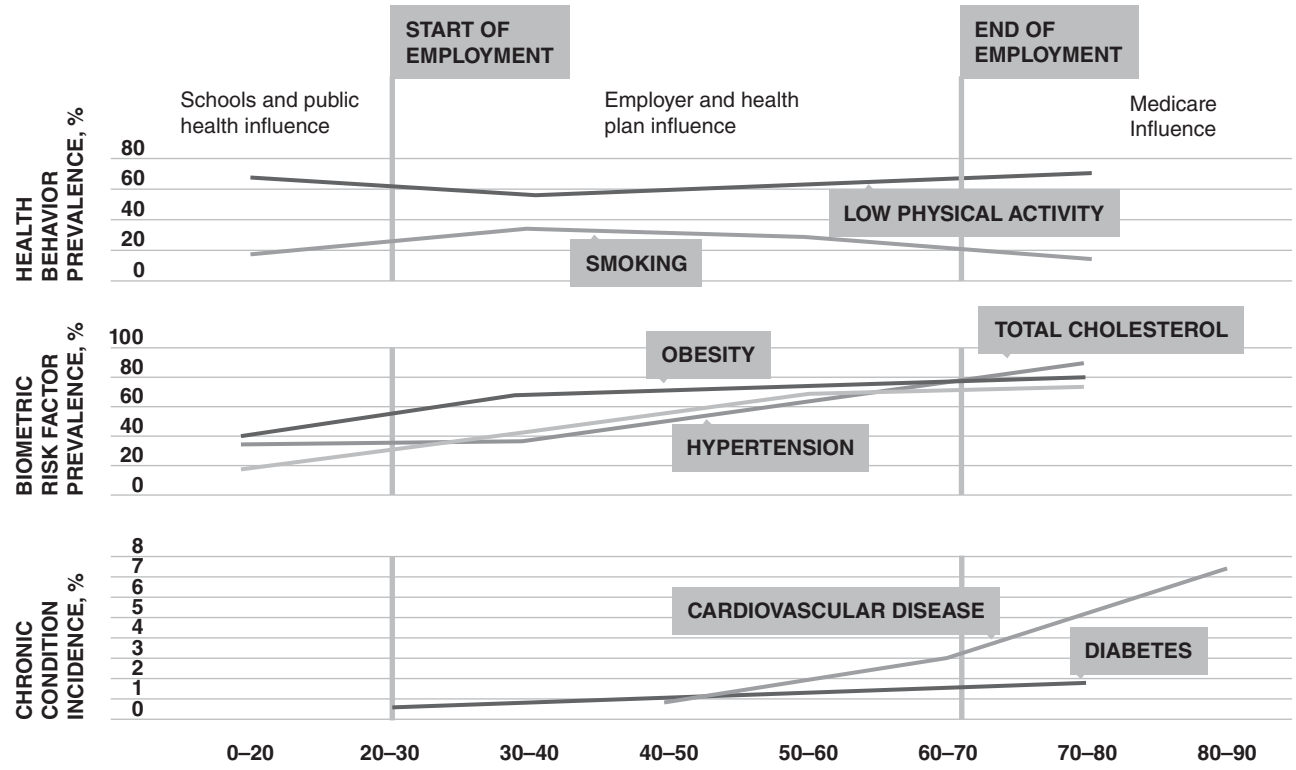
The four major chronic disease risk factors presented in Figure 8-1 are leading causes of the changes Omran described, especially with respect to



Note: *Four modifiable risk factors – dietary risks, tobacco, alcohol and drug use, and low physical activity – contribute to a significant majority of chronic diseases globally. Figure represents both sexes, all ages, 2015.*

FIGURE 8-1 Risk factors contributing to chronic diseases.

Reproduced from Institute for Health Metrics and Evaluation (IHME). (2017). *GBD Compare.* Seattle, WA: IHME, University of Washington. Available from https://vizhub.healthdata.org/gbd-compare/



Note: As behavioral risk factors build up over the life course, they can lead to chronic diseases later in life. In many countries, governments bear the burden of costly treatments for debilitating diseases at the end of life.

FIGURE 8-2 Risk accumulation over the lifespan.

Reproduced from Vitality Institute Commissioners. (2014). *Investing in prevention: A national imperative.* Retrieved from: http://thevitalityinstitute.org/commission.

TABLE 8-3 Ranking of Risk Factors: Ten Leading Risk Factor Causes of Death by Income Group, 2015

Risk Factor	Deaths (per 100,000)	Percentage Change (1990–2015)	Risk Factor	Deaths (per 100,000)	Percentage Change (1990–2015)
World			*Low-Income Countries*		
1. High systolic blood pressure	145.2	7.1	1. Child wasting	77.9	–73.0
2. Smoking	86.9	–8.2	2. High systolic blood pressure	76.4	–9.2
3. High fasting plasma glucose	71.1	39.2	3. Household air pollution	75.6	–42.4
4. High total cholesterol	58.5	0.2	4. Unsafe sex	70.5	11.7
5. Ambient particulate matter pollution	57.5	–12.2	5. Unsafe water source	53.3	–65.7
6. Diet high in sodium	56.0	7.2	6. Unsafe sanitation	40.1	–66.5

(continues)

TABLE 8-3 Ranking of Risk Factors: Ten Leading Risk Factor Causes of Death by Income Group, 2015 *(continued)*

Risk Factor	Deaths (per 100,000)	Percentage Change (1990–2015)	Risk Factor	Deaths (per 100,000)	Percentage Change (1990–2015)
World			*Low-Income Countries*		
7. High body mass index	53.7	28.3	7. Ambient particulate matter	39.9	–42.7
8. Diet low in whole grains	42.6	4.3	8. No access to hand-washing facility	37.1	–62.0
9. Diet low in fruits	39.7	2.9	9. Child underweight	29.1	–77.6
10. Household air pollution from solid fuels	38.7	–39.2	10. High fasting plasma glucose	28.2	2.2
Middle-Income Countries			*High-Income Countries*		
1. High systolic blood pressure	147.4	42.0	1. High systolic blood pressure	199.3	–17.2
2. Smoking	86.2	12.8	2. Smoking	139.6	–12.0
3. Diet high in sodium	74.9	23.6	3. High total cholesterol	100.3	–23.1
4. High fasting plasma glucose	72.9	66.0	4. High fasting plasma glucose	87.5	3.7
5. Ambient particulate matter	62.2	–1.95	5. High body mass index	84.3	3.9
6. High total cholesterol	49.5	53.8	6. Diet low in whole grains	50.0	–16.5
7. High body mass index	48.5	86.8	7. Diet high in sodium	47.6	–14.4
8. Diet low in whole grains	46.9	28.3	8. Ambient particulate matter	46.0	–19.3
9. Diet low in fruits	44.1	21.1	9. Low glomerular filtration	44.8	4.8
10. Household air pollution	40.2	–43.6	10. Diet low in fruits	43.8	–14.6

Countries grouped by socio-demographic index as per IHME classifications.
Data from Institute of Health Metrics and Evaluation (IHME) at the University of Washington, 2015.

socioeconomic, political, and cultural determinants of health. For example, rates of cigarette smoking and other tobacco use have declined in many high-income countries due in part to increased awareness of the chronic disease risks associated with this factor. Nonetheless, tobacco use has risen steadily in many lower-income countries—countries where tobacco was previously either unaffordable or unavailable, or fell outside local cultural norms. As a result, WHO (2017c) estimates that 80% of the world's 1 billion smokers now live in LMICs. Similar transitions have been reported in consumption of energy-dense, low-nutrition foods, as well as use of alcohol and other recreational substances. Although the causes of these shifts are complex, increasing disposable income, increasing contact with high-income countries and

their cultural norms, and aggressive marketing from food, beverage, alcohol, and tobacco industries all play a role in their emergence.

Macro Determinants of Chronic Disease. Chronic diseases are caused by multiple individual-level risk factors, and are also driven by health determinants at the level of entire economies and societies. These factors often correlate with levels of economic development, but society does not need to wait for rapid economic growth to take action against chronic diseases (see Exhibit 8-6, "Eight Myths of Chronic Disease Burden and Control," later in this chapter).

Global Population Aging

Populations around the world are living longer due to two demographic changes: (1) increases in life expectancy among older adults and (2) declines in fertility rates among younger adults. In 2015, the average life expectancy at birth was 71.4 years of age. Babies born in the year 1900 would not normally live beyond age 50 (National Institute on Ageing, 2011). Since then, scientific and behavioral advances, such as reductions in maternal mortality in childbirth, childhood undernutrition, and poor access to safe drinking water and sanitation, have driven the increases in life expectancy and led to a better understanding of the epidemiology underlying modifiable prevention risk factors. Discoveries such as advances in antibiotics and vaccines, as well as screening tools, have yielded more recent health gains.

It is expected that the global population age 60 and older will increase from 900 million in 2015 to 2 billion in 2050, increasing this group's share from 12% to 22% of the global population. The size of this population will also surpass the number of children younger than age 5 by 2020 (WHO, 2015a). Japan is the first country where the proportion of older adults exceeds 30%, but most of North America, Europe, China, and Russia will follow suit by 2050 (WHO, 2015b).

Despite people living longer, they are living longer in poorer health. James Fries's Compression of Morbidity Hypothesis proposes that with public health intervention, "the age of onset of chronic illness may be postponed more than the age at death, squeezing most of the morbidity in life into a shorter period with less lifetime disability." In other words, a larger proportion of life can be healthy. Many countries, however, have failed to follow this trend (Fries, Bruce, & Chakravarty, 2011). In the United States, life expectancy increased from 75.2 years to 78.2 years (3 years) between 1990 and 2010, though healthy life expectancy increased by only 2.3 years, from 65.8 years to 68.1 years (Salomon et al., 2013). Thus, longer life does not necessarily mean more freedom from chronic disease.

Globalization and Trade

Following World War II, nations signed the General Agreement on Tariffs and Trade (GATT) to coordinate international trade agreements. Between 1948 and the early 1990s, tariffs on trading products declined from 40% to 4%. GATT was eventually replaced with the World Trade Organization (WTO) in 1995.

Global trade agreements impact the cost of goods sold in a given country. Based on price differentials (and cultural preferences), the type and quantity of products and services vary among countries, and their trade can impact chronic disease risks. Production of palm oil, for example, takes place largely in Southeast Asia (led by Indonesia and Malaysia) and requires large-scale tropical rainforest deforestation. Palm oil contributes to many chronic diseases, including heart diseases and type 2 diabetes. Despite the health and environmental concerns, palm oil is produced inexpensively and sold on the international market, where it appears in packaged foods as well as toothpaste and detergent. The World Wildlife Fund estimates that 50% of packaged goods sold in supermarkets contain palm oil. Unregulated global trade agreements permit continued production and consumption of palm oil, despite the product's clear contributions to poor health and environmental damage (Yach et al., 2010).

Agricultural Subsidies

Governments provide agricultural subsidies to farmers to their boost incomes and influence the cost and supply of selected commodities. In many countries, agricultural policies are not aligned with health goals. The United States, for example, offers the majority of its agricultural subsidies to producers of grains, livestock, and dairy products. This leads to a surplus of these commodities and shortages of other foods such as fruits and vegetables. Farmers are penalized for growing "specialty crops," including fruits and vegetables, if they have received subsidies for planting other crops.

Despite evidence that fruits and vegetables prevent chronic disease, there is an estimated 22% supply shortage in the amount of fruits and vegetables needed for the world's population to meet their daily recommended servings, due in part to policies that penalize their cultivation (Siegel, Ali, Srinivasiah, Nugent, & Narayan, 2014). Meanwhile, subsidized unhealthy commodities such as grains, meats, and dairy, which are often found in low-nutrition, high-calorie "junk" foods, are over-produced.

Categories of Chronic Disease

This section details leading types of chronic disease that emerge from preventable risk factors. Each major category of chronic disease is described in terms of the trends in its prevalence and severity. These categories include neoplasms; cardiovascular diseases; chronic respiratory diseases; diabetes, urogenital, blood, and endocrine diseases; neurologic, mental, and substance use disorders; and musculoskeletal disorders.

Quantifying Chronic Disease: Deaths and Years of Life Lost

Measurement of health in across high-, middle-, and low-income countries has centered on deaths, years of life lost (YLLs), and disability-adjusted life years (DALYs) (**EXHIBIT 8-1**).

Neoplasms

Neoplasms include all benign and malignant cancers, which are caused by uncontrolled growth of cells that invade and spread to distant sites of the body (WHO, 2016a). The global prevalence of neoplasms increased by 17% between 2005 and 2015 (**TABLE 8-4**). This trend is largely due to preventable changes in lifestyle (e.g., sedentary behavior, tobacco, and alcohol use) associated with cancer as well as exposure to environmental pollutants and toxins that cause cancer. In men, the leading neoplasms causing death are lung, prostate, colorectal, stomach, and liver cancers. For women, they include breast, colorectal, lung, uterine, cervix, and stomach cancers (WHO, 2016a).

Recent data estimate that 30% of cancer deaths could be prevented by modifying underlying risk factors. Beyond behavioral risks, regular preventive screenings can assist with early detection by

EXHIBIT 8-1 Mortality Versus YLLs Versus DALYs

- **Mortality:** Measure of the number of deaths within a given population.
- **Years of life lost (YLLs):** Measure that estimates the average number of years a person would have lived if he or she had not died prematurely. Incorporates the age at which death occurs by weighing more heavily deaths at a younger age.
- **Disability-adjusted life years (DALYs):** Measure of the time lived with disability combined with the time lost due to premature mortality.

TABLE 8-4 Changes in Global Deaths from Selected Neoplasms, 2005–2015

Neoplasm	All-Age Deaths, 2005 (thousands)	All-Age Deaths, 2015 (thousands)	Percentage Change, 2005–2015 (%)
Overall	7,492.8	8,764.6	17.0
Esophageal cancer	459.3	439.0	–4.4
Stomach cancer	824.5	818.9	–0.7
Colon and rectum cancer	675.5	832.0	23.2
Liver cancer	726.7	810.5	11.5
Tracheal, bronchus, and lung cancer	1,434.5	1,722.5	20.1
Nonmalignant skin cancer	36.3	51.9	42.9
Breast cancer	439.8	533.6	21.3
Ovarian cancer	133.8	161.1	20.4
Prostate cancer	277.4	365.9	31.9
Brain and nervous system cancer	190.4	228.8	20.1

Data from Global Burden of Disease (GBD) Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388* (10053), 1459–1544.

identifying abnormal cells that may develop into cancer, improving the effectiveness of treatment on survival (Vitality Institute, 2016).

Cardiovascular Diseases

Cardiovascular diseases are the leading cause of death globally. These include diseases of the heart vessels and valves, such as ischemic heart disease and rheumatic heart disease, as well as diseases involving other blood vessels, such as cerebrovascular disease, hypertensive heart disease, and peripheral arterial disease (WHO, 2016b). Between 2005 and 2015, global deaths from cardiovascular diseases increased by 12.5%, with the greatest percent increases due to hypertensive and ischemic heart diseases (**TABLE 8-5**).

Cardiovascular diseases are largely preventable through interventions that target behavioral risk factors. Many of these risk factors are similar to those for neoplasms and cancer, including sedentary behavior, diets poor in fruits and vegetables, and tobacco use. These behavioral risks, in addition to nonmodifiable risk factors such as male gender and advanced age, contribute to development of silent chronic conditions that culminate in overt cardiovascular disease—for example, high blood pressure, elevated cholesterol levels, and elevated fasting glucose levels. Approximately 75% of all cases of cardiovascular diseases occur in LMICs.

Chronic Respiratory Diseases

Chronic respiratory diseases are diseases in the airways and other structures of the lung. The most common are chronic obstructive pulmonary disease (COPD), pneumoconiosis, and asthma. Tobacco use is the leading risk factor underlying chronic respiratory diseases, although air pollution, occupational chemicals and dusts, and lower respiratory infections during childhood all increase the risk of developing chronic respiratory diseases (WHO, 2016c). Chronic respiratory diseases are largely incurable, but can be managed through treatments that improve shortness of breath and dilate air passages.

Deaths from chronic respiratory diseases increased by 2.3% between 2005 and 2015 (**TABLE 8-6**). The relatively small increase in mortality from chronic respiratory diseases is largely a result of the effective tobacco control programs that have been implemented in recent decades.

Diabetes, Urogenital, and Endocrine Diseases

Deaths associated with diabetes, urogenital, blood, and endocrine diseases increased 29.4% between 2005 and 2015 (**TABLE 8-7**). The majority of these deaths are attributable to type 1 and 2 diabetes mellitus (**EXHIBIT 8-2** explains the difference between the types). Persons with type 1 diabetes are entirely dependent on insulin injections to survive, but persons with type 2 diabetes can also take oral medication to help the body use its existing insulin. Type 2 diabetes is both preventable and treatable with behavioral modifications, including changes to diet, physical activity, medications, and regular blood glucose screenings. The Diabetes Prevention Program is one approach to addressing diabetes (**EXHIBIT 8-3**).

Neurologic Disorders

Neurologic disorders are diseases of the central and peripheral nervous systems, which include the

TABLE 8-5 Changes in Global Deaths from Selected Cardiovascular Diseases, 2005–2015

Cardiovascular Disease	All-Age Deaths, 2005 (thousands)	All-Age Deaths, 2015 (thousands)	Percentage Change, 2005–2015 (%)
Overall	15,933.7	17,921.0	12.5
Rheumatic heart disease	333.2	319.4	–4.1
Ischemic heart disease	7,648.4	8,917.0	16.6
Cerebrovascular disease	6,020.9	6,326.1	5.1
Hypertensive heart disease	760.5	962.4	26.5

Data from Global Burden of Disease (GBD) Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388* (10053), 1459–1544.

TABLE 8-6 Changes in Global Deaths from Selected Chronic Respiratory Diseases, 2005–2015

Chronic Respiratory Disease	All-Age Deaths, 2005 (thousands)	All-Age Deaths, 2015 (thousands)	Percentage Change, 2005–2015 (%)
Overall	3,709.1	3,795.5	2.3
Chronic obstructive pulmonary disease	3,100.5	3,188.3	2.8
Pneumoconiosis	31.9	36.1	13.2
Asthma	449.9	397.1	–11.7

Data from Global Burden of Disease (GBD) Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388* (10053), 1459–1544.

TABLE 8-7 Changes in Global Deaths from Diabetes, Urogenital, Blood, and Endocrine Diseases, 2005–2015

Diabetes, Urogenital, Blood, and Endocrine Diseases	All-Age Deaths, 2005 (thousands)	All-Age Deaths, 2015 (thousands)	Percentage Change, 2005–2015 (%)
Overall	2,635.3	3,409.3	29.4
Diabetes mellitus	1,150.2	1,519.0	32.1
Chronic kidney disease	937.7	1,234.9	31.7

Data from Global Burden of Disease (GBD) Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388* (10053), 1459–1544.

EXHIBIT 8-2 Diabetes: What's the Difference?

- **Diabetes mellitus (diabetes):** Disease caused by an increased concentration of glucose in the blood, due to dysregulation of insulin, the hormone that removes glucose from the bloodstream. This condition arises either from the pancreas not producing enough insulin (type 1 diabetes) or from the body not effectively being able to use the insulin it produces (type 2 diabetes). Type 2 diabetes is far more common (95%), and is caused by hormonal changes arising from excess body weight, excess intake of sugars and carbohydrates, and physical inactivity.
- **Urogenital diseases:** Diseases affecting the urinary tracts, including the kidneys, ureters, bladder, or urethra, or the reproductive organs (Nature, 2016).
- **Endocrine diseases:** Diseases affecting hormonal glands such as the thyroid, ovaries, testes, pancreas, and pituitary gland.

Data from World Health Organization (WHO). (2016d). Retrieved from: http://www.who.int/topics/diabetes_mellitus/en.

brain, spinal cord, and cranial and peripheral nerves. Deaths from neurologic disorders increased by 35.2% between 2005 and 2015, due largely to an increase in age-related conditions such as Alzheimer's disease and Parkinson's disease as global life expectancy and median age have increased (**TABLE 8-8**).

Alzheimer's disease, the most common form of dementia, is characterized by memory loss and deterioration of cognitive abilities that interfere with daily activities. Symptoms of Parkinson's disease include tremors, slowed movement, and rigid muscles. Both diseases are incurable, but treatments exist to alleviate their symptoms to some extent. It is estimated that the number of people affected by dementia will double every 20 years (WHO, 2006a). Both Alzheimer's and Parkinson's diseases,

EXHIBIT 8-3 Diabetes Prevention Program

The Diabetes Prevention Program (DPP) was an evidence-based program to demonstrate that lifestyle changes (modest weight loss through dietary changes and increases in physical activity) or pharmacologic therapy (through the oral drug metformin) can prevent or delay the onset of type 2 diabetes. The program was administered to 3,234 participants at 27 centers across the United States. Participants, who were 45% racial and ethnic minorities, were initially classified as having prediabetes, meaning they had abnormally high blood glucose levels but had not yet developed diabetes. The study found that the lifestyle intervention resulted in a 58% reduction in the risk of developing diabetes (U.S. Department of Health and Human Services [DHHS], National Institutes of Health, & National Institute of Diabetes and Digestive and Kidney Diseases, 2008). The pharmacologic therapy group taking metformin reduced their risk by 31%. With the effectiveness of the DPP, organizations such as Omada Health (2016) have emerged to translate the program to an online platform. Following rigorous testing, the U.S. Centers for Medicare and Medicaid Services announced that Omada Health's program would be reimbursed for at-risk seniors in 2016.

Data from US Department of Health and Human Services, National Institutes of Health, and National Institute of Diabetes and Digestive and Kidney Diseases. (2008). Diabetes prevention program (DPP). National Institutes of Health. Retrieved from: https://www.niddk.nih.gov/about-niddk/research-areas/diabetes/diabetes-prevention-program-dpp/Documents/DPP_508.pdf and Omada Health. (2016). Medicare will cover diabetes prevention program for at-risk seniors. Omada Health. Retrieved from: https://www.omadahealth.com/news/medicare-will-cover-diabetes-prevention-program-for-at-risk-seniors.

TABLE 8-8 Changes in Global Deaths from Selected Neurologic Disorders, 2005–2015

Neurologic Disorders	All-Age Deaths, 2005 (thousands)	All-Age Deaths, 2015 (thousands)	Percentage Change, 2005–2015 (%)
Overall	1,671.0	2,258.9	35.2
Alzheimer's disease and other dementias	1,380.8	1,908.2	38.2
Parkinson's disease	82.4	117.4	42.4
Epilepsy	119.0	124.9	5.0

Global Burden of Disease (GBD) Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388* (10053), 1459–1544.

like other neurologic disorders, are associated with social stigma in many countries, which may cause persons suffering from these conditions to delay evaluation and treatment.

Mental and Substance Use Disorders

WHO (2014) defines mental health as "a state of well-being in which every individual realizes his or her own potential, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to her or his community." Although the slogan "no health without mental health" aims to integrate physical and mental health, on a practical level the two often occupy separate worlds of treatment and prevention (Horton, 2016). Between 2005 and 2015, all-age deaths from mental and substance use disorders increased by 6.2% (**TABLE 8-9**) (GBD Collaborators, 2015).

The global productivity loss attributable to poor mental health is approximately $1 trillion each year, or 10 billion days of lost work (Mnookin, 2016). This global cost of mental disorders is predicted to reach $6 trillion by 2030, an increase of 240% from 2010. With that said, mental and substance disorders are treatable with cost-effective interventions such as cognitive-behavioral therapy and oral medication (see the *Global Mental Health* chapter for more details on mental health). Access to mental health care can increase participation in the workforce, improve functioning at work, and reduce rates of absenteeism.

As with neurologic disorders, the stigma associated with poor mental health often inhibits effective treatment. This stigma can contribute to social

TABLE 8-9 Changes in Global Deaths from Selected Mental and Substance Use Disorders, 2005–2015

Mental and Substance Use Disorders	All-Age Deaths, 2005 (thousands)	All-Age Deaths, 2015 (thousands)	Percentage Change, 2005–2015 (%)
Overall	305.9	324.9	6.2
Alcohol use disorders	157.4	137.5	–12.6

Data from Global Burden of Disease (GBD) Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388* (10053), 1459–1544.

TABLE 8-10 Changes in Global Deaths from Drug Use Disorders, 2005–2015

Drug Use Disorders	All-Age Deaths, 2005 (thousands)	All-Age Deaths, 2015 (thousands)	Percentage Change, 2005–2015 (%)
Overall	128.8	169.9	31.8
Opioid use disorders	94.2	122.1	29.6
Cocaine use disorders	7.4	11.1	49.7
Amphetamine use disorders	7.3	12.2	67.5

Data from Global Burden of Disease (GBD) Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388* (10053), 1459–1544.

isolation, low self-esteem, and fewer chances for succeeding in maintaining employment, undertaking education, and securing housing.

Substance Use Disorders

A substance in this context is "a psychoactive compound with the potential to cause health and social problems." The most severe manifestation of a substance or drug use disorder is addiction (DHHS, 2016). Deaths from drug use disorders rose by 31.8% between 2005 and 2015 (**TABLE 8-10**).

The rising global substance-related mortality is being driven by opioid, cocaine, and amphetamine use disorders. In the United States, recent studies suggest increases in substance-related mortality among middle-aged white men, but not in other ethnic groups such as blacks and Latinos (Case & Deaton, 2015). In 2016, the Office of the U.S. Surgeon General released its first report on addiction and substance abuse, which outlined several solutions with global applicability. These include new policies and programs, investment in expanding the scientific evidence base, and a cultural shift in how addiction is viewed and treated, as the stigma of substance dependence often impairs access to treatment.

Musculoskeletal Disorders

Musculoskeletal disorders relate to the body's movement system, including the joints, ligaments, muscles, nerves, and tendons. They are the most common cause of severe long-term pain and physical disability worldwide. Deaths from musculoskeletal disorders increased by 18.2% between 2005 and 2015, largely a result of aging global populations (**TABLE 8-11**). New technologies—particularly those focused on artificial intelligence and robotics—are emerging that have the potential to enable people to maintain their movement in older age.

Country-Level Analysis

With the burden of chronic diseases becoming more prevalent in LMICs through the epidemiologic transition, various countries are attempting to tackle the prevention and management of chronic diseases. **TABLE 8-12** identifies the levels of major chronic

TABLE 8-11 Changes in Global Deaths from Musculoskeletal Disorders, 2005–2015

Musculoskeletal Disorders	All-Age Deaths, 2005 (thousands)	All-Age Deaths, 2015 (thousands)	Percentage Change, 2005–2015 (%)
Overall	76.2	90.1	18.2
Rheumatoid arthritis	26.5	30.0	13.2

Data from Global Burden of Disease (GBD) Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388* (10053), 1459–1544.

TABLE 8-12 Country-Level Data on Chronic Diseases and Risk Factors for Selected Countries

Chronic Disease Indicator	High Income: United States	Upper Middle Income: China	Middle Income: India	Lower Middle Income: Nigeria	Low Income: Rwanda
Risk Factors (DALYs per 100,000)					
Tobacco	3,324.50	3,166.13	2,080.53	355.4	627.69
High body mass index (BMI)	3,295.25	1,576.46	1,261.52	510.75	492.64
Dietary risks	3,143.06	4,022.99	3,159.35	686.33	816.49
Alcohol and drug use	2,787.24	1,944.27	1,259.82	1,372.99	1,874.15
Low physical activity	425.12	334.74	336.10	71.5	62.45
Chronic Diseases (DALYs per 100,000)					
Cardiovascular diseases	4,591.46	5,713.55	4,987.41	1,382.93	1,904.42
Neoplasms	4,231.83	4,285.42	1,772.68	1,300.64	1,707.27
Chronic respiratory	1,449.04	1,303.01	2,251.40	475.59	750.63
Diabetes	2,227.00	1,485.19	1,987.82	1,516.62	1,097.35
Mental and substance use	3,724.87	2,200.73	1,983.87	1,705.66	1,839.89

Data from Global Burden of Disease (GBD) Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388* (10053), 1459–1544.

diseases and their associated risk factors for a selection of upper-, middle-, and low-income countries.

Despite the growing prevalence of chronic diseases in many LMICs, governments have been slow to introduce effective interventions. For example, 80% of the world's tobacco smokers reside in LMICs, yet few of the governments in these countries have introduced excise taxes on tobacco. If a 33% price increase on tobacco cigarettes were implemented globally, 22 million to 65 million smoking-related deaths would be averted over the next 50 years—almost 90% of those deaths would occur in LMICs (Gaziano & Pagidipati, 2013). **TABLE 8-13** highlights the most cost-effective interventions for chronic diseases in

TABLE 8-13 Cost-Effective Interventions for Chronic Diseases in Resource-Constrained Environments

Risk Factor/Disease	Interventions
Tobacco use	■ Tax Increases ■ Smoke-free indoor workplaces and public places ■ Health information and warnings ■ Bans on tobacco advertising, promotion, and sponsorship
Harmful alcohol use	■ Tax increases ■ Restricted access to retailed alcohol ■ Bans on alcohol advertising ■ Reduced salt intake in food
Unhealthy diet and physical inactivity	■ Replacement of trans fat with polyunsaturated fat ■ Public awareness through mass media on diet and physical Inactivity
Cardiovascular disease (CVD) and diabetes	■ Counseling and multidrug therapy for people with a high risk of developing heart attacks and strokes (including those with established CVD) ■ Treatment of heart attacks with aspirin
Cancer	■ Hepatitis B immunization to prevent liver cancer (already scaled up) ■ Screening and treatment of precancerous lesions to prevent cervical cancer

Note: NCD = noncommunicahle disease.

resource-constrained environments. Country examples of effective chronic disease control strategies are detailed later in this chapter.

Clinical Features of Common Chronic Diseases

After a chronic disease has developed, it can often be treated, controlled, or even cured with early detection, a process known as *secondary prevention*. Although some chronic conditions can be diagnosed only with laboratory tests and other investigations, many individuals display signs and symptoms that can be detected with minimal medical training. This allows physicians and community health workers alike to refer high-risk persons for further specialized testing. **TABLE 8-14** summarizes common clinical signs of key chronic conditions.

Comorbid Conditions

Comorbidity refers to the simultaneous occurrence of two or more disorders in one person. A common chronic disease comorbidity is diabetes and cardiovascular disease. Heart attacks, for example, are significantly more common in people with diabetes, as are deaths following a first heart attack (International Diabetes Federation, 2003). Similarly, high cholesterol and hypertension are both causative of and concomitant with coronary heart disease and cerebrovascular disease, and both are growing more common with urban development (see **EXHIBIT 8-4**). In addition, persons with chronic conditions have a greater risk of developing mental disorders such as depression (WHO, 2003a). The proportion of patients with depression who also have other common chronic diseases, such as cardiovascular disease, diabetes, and cancer, ranges from 22% to 33% (WHO, 2003a).

There are also interactive effects between certain infectious and noninfectious diseases. Several infectious agents cause cancer: Hepatitis B virus causes liver cancer; human papillomavirus (HPV) causes cervical cancer; *Helicobacter pylori* causes stomach cancer; HIV causes several cancers, including Kaposi's sarcoma and non-Hodgkin's lymphoma; and *Schistosoma haematobium* causes bladder cancer (Stewart & Kleihaus, 2003). All of these cancers are common in LMICs, especially where resources for treatment are inadequate. Vaccines to prevent these infections and effective drugs to treat them could greatly reduce the cancer burden in these countries.

TABLE 8-14 Clinical Features of Common Chronic Conditions

Conditions	Typical Clinical Features	Diagnostic Requirements
Hypertension	No obvious symptoms	Measure blood pressure using standardized procedures
Hyperlipidemia	No obvious symptoms	Measure total cholesterol, high-density lipoprotein cholesterol, and triglyceride levels in fasting blood samples
Obesity	Fatigue, daytime somnolence, excessive snoring, and osteoarthritis	Measure height (meters) and weight (kilograms); calculate body mass index (BMI) = weight/height^2; measure waist circumference
Diabetes	Malaise, excessive thirst, excessive urination, hunger, blurred vision, and tendency to develop infections	Measure fasting or random blood glucose, perform glucose tolerance test, or check blood hemoglobin A1c (HbA1c) level
Asthma	Wheezing, difficulty with breathing, and coughing (bronchospasm) relieved by asthma medication	Peak flow measurements and other lung function tests, chest x-ray, and relief of symptoms with bronchodilators
Chronic bronchitis	Productive cough for 3 months per year in 2 consecutive years, shortness of breath, and frequent chest infections	Lung function tests, chest x-ray
Myocardial infarction (heart attack)	Sudden onset of severe crushing chest pain that could radiate down left arm, to the neck or jaw, with associated sweating, faintness, shortness of breath, and nausea	Clinical examination, electrocardiograph, blood tests for cardiac enzymes
Cerebrovascular disease (stroke)	Sudden weakness; loss of motor or sensory function, usually unilateral; inability to speak; vision disturbances; or unconsciousness	Neurologic and full clinical examination, head computed tomography (CT) scan if available, identify underlying causes
Angina	Central chest pain precipitated by exertion and relieved by resting	Electrocardiograph, pain relieved by angina medication
Transient ischemic attack	Same presentation as stroke—weakness, loss of motor or sensory function, inability to speak, or vision disturbance—but resolving within 24 hours	Neurologic and full clinical examination, head CT scan if available, identify underlying causes
Neoplasms and cancers	Unexplained loss of weight, malaise, and tiredness	Full clinical examination; biopsy, imaging, and/or special blood tests as appropriate

Finally, there are interactions between chronic disease risk factors and nonchronic infectious diseases. Tobacco increases the death rate from tuberculosis (TB)—a common disease of poverty—in those persons already infected with the TB bacterium. In India, smokers are 4.5 times more likely to die of TB than are nonsmokers (Gajalakshmi, Peto, Kanaka, & Jha, 2003). An estimated 80% of TB-infected patients smoke. As a result, tobacco is probably the major cause of death in treated TB patients (Yach & Raviglione, 2004).

The current burden of chronic diseases reflects cumulative risks over people's lifetimes. The accumulation of chronic disease risk begins in fetal life, and

EXHIBIT 8-4 Hypercholesterolemia and Hypertension in Africa

After many decades of scientific debate, data from large community-based trials in the 1970s and 1980s proved that high total blood cholesterol levels are an independent major risk factor for atherosclerotic chronic diseases such as coronary heart disease and cerebrovascular disease (strokes). Total blood cholesterol levels vary considerably between populations with different dietary patterns. For example, people in Africa who follow traditional diets may have lower blood cholesterol levels than people in Europe or the United States, including those who migrated from LMICs and adopted typical Western lifestyles.

These differences are present from a young age. In Johannesburg and Soweto, South Africa, Steyn and colleagues (2000) found that the mean total cholesterol level in ethnically African and multiracial 5-year-old children was 3.9 mmol/L, compared with 4.1 mmol/L for Indian children and 4.4 mmol/L for white children. Several studies in Africa have shown that total cholesterol levels usually differ between urban and rural settings independent of ethnicity, reflecting the effects of urbanization on increasing total cholesterol levels (Knuiman, Hermus, & Hautvast, 1980; Seftel et al., 1993; Swai et al., 1993).

It is estimated that between 10 million and 20 million people in sub-Saharan Africa alone have hypertension and, further, that adequate hypertension treatment could prevent approximately 250,000 deaths in this region (Cappuccio, Plange-Rhule, Phillips, & Eastwood, 2000). Unfortunately, hypertension is universally underdiagnosed and inadequately treated, such that extensive end-organ damage and premature death are often seen. Furthermore, hypertension frequently coexists with other chronic diseases such as diabetes.

Early surveys in sub-Saharan African countries showed that the lowest prevalence of hypertension occurred in the lowest-income countries; as affluence increased, however, prevalence increased. Researchers also found that hypertension is more common in urban than rural settings (Nissinen, Bothig, Granroth, & Lopez, 1988). The Kenyan Luo migration study conducted by Poulter and colleagues (1990) showed that people migrating from traditional rural villages on the northern shores of Lake Victoria to the urban settings of Nairobi had higher body weights, blood pressure, pulse rates, and urinary sodium–potassium ratios than those who remained in the rural areas. This result suggests a marked change in the diet of the new arrivals in Nairobi, including higher salt and calorie intakes, along with a reduced potassium intake due to consumption of fewer fruits and vegetables.

marches forward starting early in infancy and childhood (**FIGURE 8-3**). Thus, emerging evidence supports starting health promotion efforts during pregnancy and early childhood and continuing prevention efforts throughout the life course (Institute of Medicine, 2010).

Early-life and prenatal risk factors—such as suboptimal diet, early termination of breastfeeding, exposure to tobacco and alcohol, exposure to indoor air pollution from biomass fuels, and repeated respiratory infections—are important for the development of chronic diseases in adulthood. Recent research from birth cohort studies has documented how and when these life course influences happen (Aboderin et al., 2001; Batty & Leon, 2002). Some life-course influences are disease-specific, whereas others are cohort-specific, but nearly all depend on a few risk factors. For example, 80% to 90% of patients who develop clinically significant congenital heart disease and more than 95% of patients who have experienced a fatal congenital heart disease event have at least one of the major cardiac risk factors—smoking, diabetes, hypertension, or hypercholesterolemia (Canto & Iskandrian, 2003). Studies also indicate that early postnatal nutrition permanently affects the major components of the metabolic syndrome that determine the individual's propensity to cardiovascular disease (Singhal & Lucas, 2004) and that influences in fetal life and early childhood are related to systolic blood pressure (Levitt et al., 1999).

Chronic obstructive pulmonary disease and lung dysfunction in adults are also the result of cumulative exposures that start early in life. South African studies have shown that what were assumed to be genetic differences in lung size are probably due to early childhood respiratory infections occurring in crowded homes where biomass fuel is used, combined with tobacco use and adverse occupational exposures (Goldin, Louw, & Joubert, 1996). A recent population survey of chronic bronchitis in 5,671 men and 8,155 women in South Africa estimated that 25% of these cases were due to tobacco use, 14% to occupational exposure, and 10% to past TB. In women, these fractions were 14% for use of smoky domestic fuel, 10% for past TB, and 11% for tobacco use (Ehrlich et al., 2004). This study demonstrates how multiple assaults on the lung—including infectious agents, fuel use, exposure at work, and tobacco use—are important in chronic lung disease. All of these factors have been shown to be related to poverty as well.

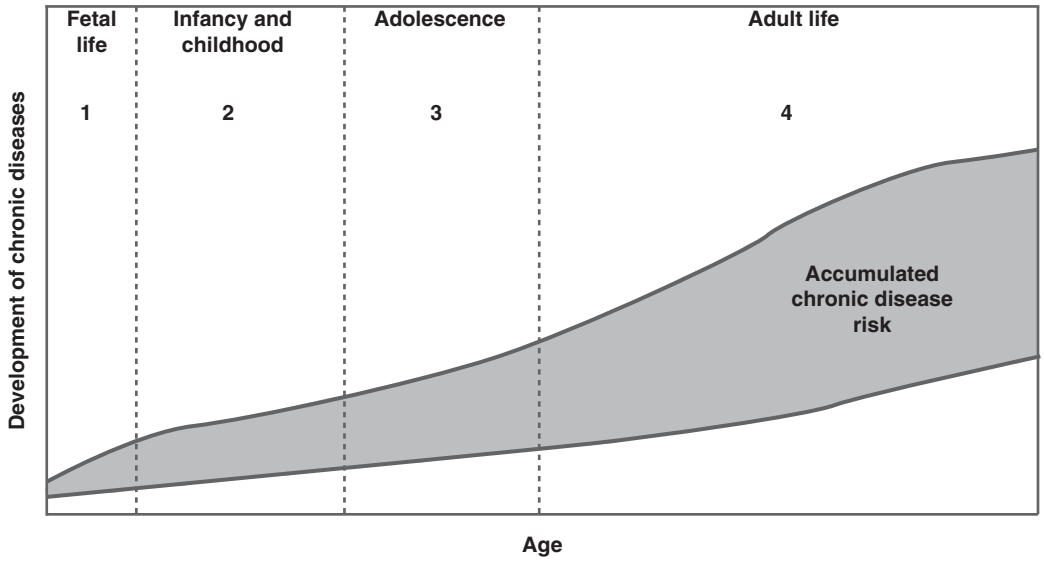


FIGURE 8-3 A life course approach to chronic disease prevention.

Reprinted from Aboderin, I., Kalache, A., Ben-Shlomo, Y., Lynch, J. W., Yajnik, C. D., Kuh, D., & Yach, D. (2001). *Life course perspectives in coronary heart disease, stroke and diabetes: Key issues and implications for policy and research*. Geneva, Switzerland: World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/10665/67173/1/WHO_NMH_NPH_01.4.pdf

Global Costs and Consequences of Chronic Diseases

Although chronic diseases will exact a grave financial toll if not addressed by global stakeholders, their potential impact on other aspects of human well-being, such as environmental welfare and socioeconomic inequity, is also severe. This section explores the global costs and consequences of chronic diseases.

Direct Economic Costs

By 2030, it is estimated that the cost of chronic diseases will amount to $30 trillion worldwide (Bloom et al., 2011). In the United States, for example, chronic diseases are the primary healthcare cost drivers, with diabetes ($101 billion), heart disease ($88 billion), and low back and neck pain ($86 billion) being the top three in terms of healthcare expenditures (Dieleman et al., 2016). Chronic diseases not only threaten the fiscal integrity of health budgets and shorten productive work careers, but also undermine the long-term sustainability of systems that deliver universal health coverage (Greenberg, Leeder, & Raymond, 2016). Unfortunately, chronic diseases remain greatly underfunded in LMICs (Clinton & Sridhar, 2017). None of the major health development funders, such as the Bill and Melinda Gates Foundation, WHO, World Bank, Global Fund to Fight HIV/AIDS, TB and Malaria, and Global Alliance of Vaccine Initiatives, is dedicated exclusively to chronic diseases.

Further Macroeconomic Impacts

Indirect productivity losses due to chronic diseases will affect the global economy because premature death and disability will dampen the engine of productivity and reduce economic growth in LMICs (Leeder, Raymond, Greenberg, Liu, & Esson, 2004). Some economists argue the potential returns for high-income-country investors from LMICs could be higher than from high-income countries over the long term (Clark & Hebb, 2002; Heller, 2003; Kimmis, Gottchalk, Armendariz, & Griffith-Jones, 2002). If chronic diseases do diminish productivity as predicted, the effects will, therefore, impact both wealthy countries and lower income countries alike. Both multinational corporations and pension funds face risks from this source, as demonstrated by the HIV/AIDS pandemic, which showed that chronic conditions can place heavy financial burdens on a company when treatment is expensive.

Impact on Health Inequalities

In 1952, Gunnar Myrdal asserted that people are sick because they are poor, and they become poorer because they are sick. This downward spiral certainly occurs with chronic diseases. Increased cumulative exposure to risk factors over the life course, combined with social and economic inequalities, leads to inequalities seen in later adult life.

Chronic diseases impose a significant burden on low-income populations. In high-income countries, the relationships among poverty and cardiovascular disease, cancer, diabetes, and their associated risk

factors are well described, and chronic diseases drive health inequalities by social class, ethnicity, and gender (Aboderin et al., 2001; Batty & Leon, 2002; Brands & Yach, 2001; Kogevinas, Pearce, Susser, & Boffetta, 1997; Mackenbach et al., 2000; Marmot, Adelstein, Robinson, & Rose, 1978; Wong, Shapiro, Boscardin, & Ettner, 2002).

In LMICs, the poorest populations already exhibit the highest risk of tobacco and alcohol use (Jha & Chaloupka, 2000), but the relationship with intermediate risk factors and diseases is complex and varies between countries:

- India: Currently, rates of hypertension, cholesterol, diabetes, and cardiovascular disease increase directly with socioeconomic status (Singh et al., 1999; Vikram et al., 2003). Yet, high rates of hypertension and elevated cholesterol levels are now being measured in urban slums (Misra et al., 2001), and tobacco consumption is higher among the most poorly educated. Death rates from TB—a disease associated with poverty— are four times higher among people who smoke relative to nonsmokers (Gajalakshmi et al., 2003). Poorer people also suffer from relatively higher rates of complications from diabetes, owing to their frequent exposure to multiple risk factors (Ramachandran, Snehalatha, Vijay, & King, 2002).
- Brazil: In urban Brazil, research in the 1980s demonstrated an inverse relationship between socioeconomic status and smoking and alcohol consumption in men and women (Duncan et al., 1993). Researchers also found inverse links between socioeconomic status and hypertension in men, and between sedentary lifestyle and obesity in women.
- Jamaica: Obesity increases with income level strongly in men and weakly in women (owing to high levels of obesity among the poorest women) (Mendez et al., 2004). The relationships among income, diabetes, and hypertension, however, are nonlinear. In women, plotting diabetes and hypertension prevalence against income forms a U-shaped curve, with the highest rates of diabetes found in the poorest women. Obesity and diabetes are strongly related, especially among poor women. The lack of a strong income gradient in hypertension or diabetes—despite a strong relationship between income and obesity—might be partly attributable to greater adverse effects of obesity among the poor.
- China: A U-shaped relationship exists between socioeconomic status and hypertension in women. This result in part reflects the fact that poorer women tend to have lower body mass indices and to smoke less but engage in less physical activity; the reverse is true for the wealthiest group (Bell, Adair, & Popkin, 2004).

Owing to long and often variable lag times between exposure to risk factors and disease onset, the relationships among poverty, exposure to chronic disease risks, and development of disease remain unclear in LMICs. Over time, accumulation of risk among low-income groups will likely increase as availability and marketing of products associated with a Western lifestyle (e.g., higher-fat, higher-calorie foods) increases in tandem with economic development. In the long term, this trend could worsen associations between chronic diseases and poverty.

Obesity data provide an example. Until the late 1980s, socioeconomic status and obesity tended to be inversely related in high-income populations and directly related in low- and middle-income populations (Sobal & Stunkard, 1989). In other words, in LMICs, obesity was associated with more affluent groups. More recently, however, work from Brazil has shown that over time female obesity shifts toward lower-income groups in economically more developed regions and urban areas (Monteiro, Conde, Lu, & Popkin, 2004). Although low socioeconomic status still confers protection from obesity in low-income nations, once the gross national product (GNP) reaches a value of approximately $2,500 per capita, obesity rates correlate directly with low socioeconomic status (Monteiro et al., 2004; Peña & Bacallao, 2000).

A recent study in the United States and China illustrates this point (Kim, Symons, & Popkin, 2004). Using a composite lifestyle index (LI) that included data on diet, smoking, alcohol, and physical activity, the authors found an inverse relationship exists between socioeconomic status and the LI in the United States, whereas a direct relationship exists in China. This result suggests that policy makers in LMICs should not wait for a social class gradient to appear in the occurrence of chronic disease (or risk factors) before implementing disease-preventive and health-promoting policies.

A growing challenge is the rise of health inequalities among the elderly. With the global population age 60 years and older forecast to more than triple in the next 45 years, the size of this gap will increase substantially (WHO, 2008a). In LMICs, where the proportion of older people is growing fastest, the lack of pension schemes and deteriorating traditional social security arrangements act together to worsen the burden of health inequalities (Commission on Social

Determinants of Health, 2008). Although health inequalities in the elderly tend to be smaller than in younger populations, they remain an important and poorly understood public health challenge (von dem Knesebeck, 2010). Older women, especially widows, are particularly affected because they tend to be poorer and more susceptible to chronic diseases (Women and Gender Equity Knowledge Network [WGEKN], 2007). The *Health and 'The Economy'* chapter provides more information on the relationship between health and the economy.

Impact on Sustainable Development and the Environment

Sustainable development, which in recent decades has linked economic and political development with environmental stewardship, has long impacted chronic disease control strategies and priorities, including the United Nations' recent Sustainable Development Goals. Global research and policy agreements now agree that chronic disease proliferation and environmental threats are inextricably linked.

Sustainable Development and Planetary Health

In 1987, the Bruntland Commission proposed what is now the commonly accepted definition of sustainable development: "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (World Commission on Environment and Development, 1987). Realizing the linkages between human health and the environment, the Commission aimed to place environmental issues on the agenda of the United Nations. It provided the foundation for the 1992 Earth Summit in Brazil, which culminated in Agenda 21 and the Rio Declaration on Environment and Development, and later for the Commission on Sustainable Development (**EXHIBIT 8-5**).

The outputs from UNCED established that human health was both an input and an output of sustainable development. In other words, sustainable development cannot be achieved if there is a high prevalence of chronic disease, whereas human health cannot be maintained without healthy environments (von Schirnding, 2002). Conversely, if left unchecked, drivers of chronic disease can and will impact sustainable environmental growth—for example, through global demand for energy-dense foodstuffs, such as refined grains and red meats, that consume more production resources than vegetables; or use of fossil fuels to power automobiles in lieu of more physically energetic transport choices such as walking or biking.

EXHIBIT 8-5 The United Nations' Summits on Sustainable Development

- **United Nations Conference on Environment and Development** (UNCED: 1992, Rio de Janeiro, Brazil): The primary output was Agenda 21, a nonbinding, voluntary action plan voted on by 178 governments to achieve sustainable development. The UN General Assembly established the Rio Declaration on Environment and Development to coordinate follow-up from the conference.
- **World Summit on Sustainable Development** (2002, Johannesburg, South Africa): The World Summit evaluated progress from the UNCED. The Johannesburg Declaration was the result of the summit, which espoused multilateralism to achieve the objectives of sustainable development.
- **United Nations Conference on Sustainable Development** (2012, Rio de Janeiro, Brazil): The output was an outcome document to facilitate action promoting a sustainable future for the planet for present and future generations. This conference began to establish the Sustainable Development Goals, to build on the Millennium Development Goals.

Chronic diseases have multiple causes and effects, at the level of both individuals and society. Studies of chronic diseases' impact—and the best ways to control it—have debunked several prior assumptions commonly held in global health and development circles (**EXHIBIT 8-6**).

Stakeholder Responses to the Growing Burden of Chronic Disease

The causes of chronic diseases are many and complex, and arise at both individual and population levels. Nevertheless, control of chronic disease is possible, though it requires creative thinking, avoidance of myths and false assumptions, and engagement of multiple stakeholders. A variety of stakeholders have begun to tackle this burden worldwide. This section explores their responses to the growing burden of chronic diseases.

EXHIBIT 8-6 Eight Myths of Chronic Disease Burden and Control

Myth 1: We can wait until infectious diseases are controlled.
Reality: As development progresses, chronic diseases do not smoothly displace acute diseases. Many countries have a double burden of disease; thus we must deal with both and develop the health system accordingly. Further, some infectious diseases are chronic in nature (e.g., HIV/AIDS).
Myth 2: Economic growth will improve all health conditions.
Reality: Economic development can improve health in LMICs, yet economic growth can also exacerbate chronic diseases.
Myth 3: Chronic diseases are diseases of affluence.
Reality: Chronic diseases are not solely diseases of affluence in high-income countries and in most LMICs. Low socioeconomic status leads to cumulative exposure to risk factors, greater comorbidity, and decreased access to quality health care.
Myth 4: Chronic diseases are diseases of the elderly.
Reality: Chronic diseases in LMICs are no longer just diseases of the elderly. Instead, chronic diseases in these countries affect a much higher proportion of people during their prime working years, compared to high-income countries.
Myth 5: Chronic diseases result from freely adopted risks.
Reality: Chronic diseases cannot be blamed solely on the failure of individual responsibility, because the cultural and environmental contexts in any society or community inevitably affect personal choices. Thus governments, industry, and others play a role in their incidence.
Myth 6: Benefits of chronic disease control accrue only to individuals.
Reality: Chronic disease control fosters positive social development and benefits societies economically, thereby benefiting the public as a whole. Like acute disease control, chronic disease control is an appropriate public investment.
Myth 7: Acute, infectious disease models are applicable to chronic diseases.
Reality: Interventions for acute diseases are relatively simple, whereas chronic diseases require a planned, proactive approach to health care, and the active participation of patients, families, and communities.
Myth 8: Treating individuals in the health sector is the only appropriate chronic disease strategy.
Reality: The medical community has focused on using traditional approaches to screen "high-risk" individuals—that is, those persons with a high probability of contracting chronic diseases. Yet, prevention requires a multisectoral commitment in addition to more comprehensive health service interventions for clinical prevention.

EXHIBIT 8-7 Government Levers to Influence Public Health

- **Taxes:** A tax is a financial charge imposed by a government to fund public expenditures. Governments often introduce taxes on products harmful to human health, such as tobacco and alcohol products, and more recently sweetened beverages. Differential taxes—that is, taxes proportionate to the health risk of a product—are particularly effective in deterring unhealthy behaviors and generating additional income for government (Pomeranz, 2016).
- **Access restrictions:** Governments frequently introduce restrictions on the availability of harmful products. For example, minimum age restrictions on the purchase of tobacco and alcohol products limit their use by children and adolescents. Restrictions on purchasing times and locations are also common.
- **Educational campaigns:** Educational campaigns inform a population on a specific health issue, such as physical activity or diet. This approach may also include compulsory labeling on nutrition content in foods.

Government

In many high-income countries, maintaining and improving public health is a government responsibility. Governments can introduce a variety of interventions to influence and motivate the public's health behaviors (**EXHIBIT 8-7**). These are sometimes predicated on paternalistic measures that entail modifying health behaviors at the population level.

One notable example is Finland's North Karelia Project, which represented a response to the high cardiovascular mortality in the country, and which used education, community outreach, and policy and care delivery changes to fight cardiovascular disease in this northern Finnish province. This pilot project began in 1972. Between 1969 and 2005, age-adjusted coronary heart disease mortality rate among the 35- to 64-year-old male population declined in North Karelia by 85% and in all of Finland by 80% (**FIGURE 8-4**, **EXHIBIT 8-8**) (Puska, 2016). The project formally ended after 25 years, though national prevention activities continue.

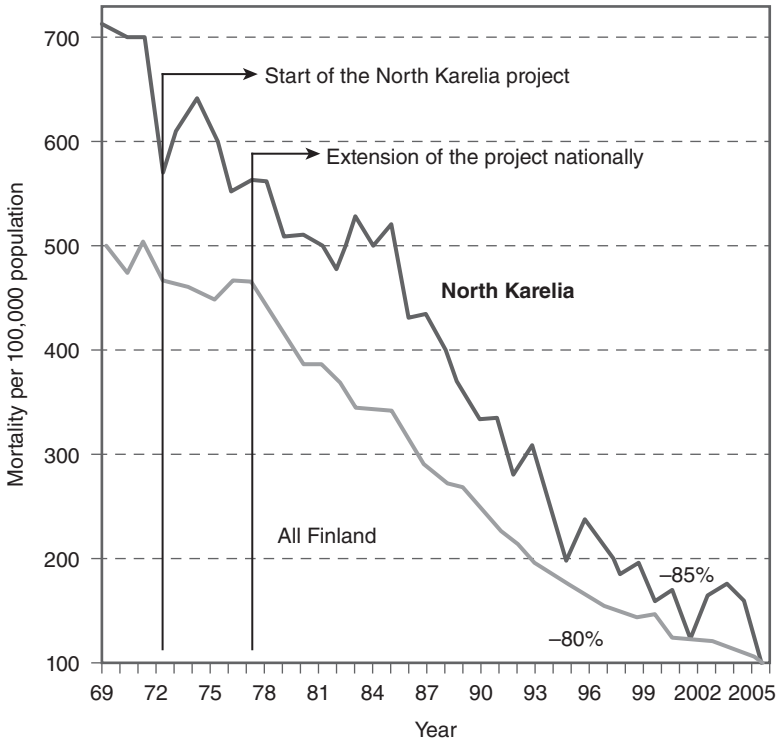


FIGURE 8-4 Age-adjusted mortality rates of coronary heart disease in North Karelia, 1969–2005.

Reproduced from Puska P, Vartiainen E, Laatikainen T, Jousilahti P & Paavola M. (2009). *The North Karelia Project: From North Karelia to national action.* National Institute for Health and Welfare & North Karelia Project Foundation. Retrieved from: https://www.julkari.fi/bitstream/handle/10024/80109/731beafd-b544-42b2-b853-baa87db6a046.pdf?sequence=1.

EXHIBIT 8-8 Why Was the North Karelia Project Successful?

- **Appropriate theory base:** Correct and appropriate public health understanding of the problem that identified major causal risk factors in the population and effective interventions.
- **Flexible intervention:** Ability for interventions to be modified based on practical situations and natural occurrences in the community.
- **Intensive intervention:** Interventions with the potential to mobilize and reach a large number of people.
- **Working with the people:** Collaborative approach with the community that established ownership by the people.
- **Community organization:** Engagement with organizations in the community to contribute to practical objectives.
- **Work with health services:** Health services must be supportive and form a backbone to the local initiatives.
- **Official authority:** Linkage to official administrative structures and health authorities.
- **Limited targets/outcome orientation:** Targeted interventions that were oriented to health outcomes.
- **Positive messages:** Positive framing of health messages to overcome a challenging problem.
- **Bottom-up, top-down:** A blended model that integrated approaches from the top and the bottom.
- **Working with the media:** Partnerships with the media to disseminate information about key activities and results.

Data from Puska, P. (2016). Why did North Karelia-Finland work?: Is it transferrable? *Global Heart, 11*(4), 387–391.

Another country making strides in curbing chronic diseases is Mexico. Diabetes is the leading cause of death and disability in Mexico, and its prevalence increased by 60% between 2000 and 2012. During the same period, hypertension increased by approximately 30%. Despite the large burden of chronic diseases in the country, access to services and effective interventions remains low. The national health system has responded by establishing a chronic disease department within the Ministry of Health, and implementing strategies to improve physical activity, minimize unhealthy diets, and reduce the harmful effects of tobacco and alcohol use.

Mexico also introduced a soda tax of 1 peso per liter on sugar-sweetened beverages in January 2014. Purchases of the taxed beverages decreased by 5.5% in

2014, and by 9.7% in 2015. Households at the lowest socioeconomic level had the largest declines in purchases of taxed beverages. While the tax is a starting point, greater advocacy and interventions are needed to shift the growing burden of chronic diseases in Mexico.

A third country addressing chronic diseases is Ghana, which is testing a community-based approach to cardiovascular disease (CVD) management and control through an international public–private partnership. The project aims to use task-shifting and new technology to improve the Ghana Health Service's response to hypertension (high blood pressure) and other cardiovascular diseases, and to help patients to better manage these conditions (Lamptey et al., 2017). It includes five components:

- Strengthening the health system to better manage hypertension services, including a referral system between the national health system and private drug outlets
- Shifting screening, monitoring, and management services to community health officers and private drug outlets
- Training patients to manage their own hypertension and control other risk factors
- Leveraging technology to improve communication, education, clinical management, adherence, and health information management
- Minimizing out-of-pocket costs by ensuring health coverage by the Ghana National Health Insurance Scheme (National Academies of Sciences, Engineering, and Medicine, 2017)

In 1986, WHO held the First International Conference on Health Promotion in Ottawa, Canada. This conference led to the Ottawa Charter for Health Promotion, which provided a blueprint to achieve "Health for All" by the year 2000 (WHO, 1986). The Ottawa Charter proposed to situate health within the ambit of the individual, in addition to the broader social and structural context that government provides. This framework foreshadowed the importance of behavioral economics by indicating that "the healthier choice [will become] the easier choice" for promoting health.

Behavioral economics has emerged as a complementary approach to government interventions in health (**TABLE 8-15**). Pioneered by Daniel Kahneman and Amos Tversky, behavioral economic theory acknowledges that human decision making can be irrational. For example, individuals give greater weight to payoffs that are closer in the present (termed "present bias") and are more likely to select a smaller reward delivered in the short term than a larger award in the longer term (termed "hyperbolic discounting"). Behavioral economic strategies, aiming to nudge and reward individuals to reject these behaviors, have emerged to correct these decision-making "irrationalities" (Kahneman, 2011). Though government-led, these approaches to behavioral economics to change behaviors are not paternalistic. The U.S. and U.K. governments have embraced behavioral economics by establishing behavioral insight teams that apply research from behavioral economics and psychology to policy making to encourage healthy and other desirable behaviors.

Business

Some companies are starting to invest in the environment and society, realizing that nonfinancial performance impacts their financial bottom line. Moving beyond strategic philanthropy (charitable donations) and corporate social responsibility (sustainability activities), many companies are striving to create shared value (Porter & Kramer, 2011)—that is, business models that generate economic benefit, while addressing societal issues. This approach rests on the premise that companies must generate societal benefit at profit. As explained in **EXHIBIT 8-9**, there are three

TABLE 8-15 Government and Behavioral Economics Approaches to Health Promotion

	Government	Behavioral Economics
Intervention approach	Paternalistic	Nonpaternalistic
Target	Population	Individual
Examples	Directing behaviors through taxes, age and purchasing restrictions, education campaigns	Nudging behaviors through financial or nonfinancial incentives, discounts, or rewards

EXHIBIT 8-9 Classification of Companies Creating Value Beyond the Financial Bottom Line

1. **Companies promoting health:** The company develops products and services that are aligned with promoting better health. Companies pursuing this strategy most often focus on expanding the reach of their portfolio of offerings. Health and life insurers, workplace health providers, and drug and vaccine developers are examples.
2. **Companies undergoing transformations to advance health:** The company offers a mix of products and services that both promote and harm health. Companies adopting this approach work to lower the risks of their unhealthy products and improve the healthiness of their other offerings. Examples include major food and beverage companies.
3. **Companies harmful to health:** The company solely develops products and services that are harmful to human health. There is no intent to transform to develop healthier alternatives. Tobacco and firearms companies are examples.

TABLE 8-16 Corporate Reporting Platforms

Reporting Platform	Description
Dow Jones Sustainability Indexes	Launched in 1999 as the first global sustainability benchmark. Tracks company stock performance based on economic, environmental, and social criteria.
FTSE4Good	Ethical investment stock market indices launched in 2001 by the FTSE Group.
Global Reporting Initiative	Independent organization that helps organizations understand and communicate the impact of business on sustainability issues, including climate change, human rights, and corruption.
International Integrated Reporting Council	Global coalition of regulators, investors, companies, standard setters, the accounting profession, and nongovernmental organizations (NGOs) that establishes integrated reporting practices.
King Report on Corporate Governance	Guidelines for the governance structures and operation of companies in South Africa. Four reports have been issued to date: 1994 (King I), 2002 (King II), 2009 (King III), and 2016 (King IV).
Sustainability Accounting Standards Board	Nonprofit organization incorporated in 2011 to develop and disseminate sustainability accounting standards.
United Nations Global Compact	Voluntary initiative based on CEO commitments to implement universal sustainability principles.

classifications of companies that operate within a shared value lens.

To encourage companies to embrace this concept of shared value, corporate reporting platforms now encourage voluntary disclosures on the extent to which companies address environmental, social, and governance (ESG) indicators (**TABLE 8-16**). This information is often presented in a sustainability or integrated report published by the company.

Realizing the central role that health plays within the workplace, marketplace, and community, companies are beginning to explore how to report on their business's impact on health. For example, companies are investigating ways to indicate the health status of their employees, the extent to which health is considered in their products and services, or how they improve health within the communities where they operate or are located.

In the workplace, many companies have introduced programs to improve the health of employees by targeting major chronic disease risk factors, and thereby minimize their healthcare costs. Recent studies suggest a correlation between a company's investment in workplace health and its financial performance,

due to reduced healthcare costs, increased productivity, and improved financial gains. In a portfolio of U.S. companies that had received a designated workplace health award, the rate of return outperformed the Standard & Poor (S&P) 500 average over a 13-year time frame (Goetzel et al., 2016). This analysis was later replicated in South Africa using Discovery Ltd.'s Healthy Company Index, where the portfolio of healthy companies consistently outperformed the market (Conradie, van der Merwe Smit, & Malan, 2016). This work has also led to partnerships between businesses and community organizations to promote health (Oziransky, Yach, Tsao, Luterek, & Stevens, 2015).

United Nations

In addition to national governments and multinational businesses, international agencies such as the United Nations are increasingly responding to chronic diseases. The United Nations hosted its first High-level Meeting on Non-communicable Diseases in 2011. UN member-states have agreed to convene a third High-level Meeting on chronic diseases in 2018 (Mendis, 2016). Additionally, the Sustainable Development Goals (SDGs), which emerged in 2015 from the Millennium Development Goals (MDGs), now incorporate multiple chronic disease goals and targets.

Sustainable Development Goals

Building on the eight MDGs proposed in 2000 and closed in 2015, the United Nations launched the SDGs. These 17 goals aim to end poverty, protect the planet, and ensure prosperity for all. Member-states agreed to the SDGs on September 25, 2015, and they came into force on January 1, 2016. Each goal has specific targets for 2030 (United Nations, 2016). The *Introduction* to this book lists all 17 goals.

Related to chronic diseases and sustainable development, SDG3 focuses on "good health and well-being." It incorporates targets against infectious and chronic diseases, with the aim of ensuring healthy lives and promoting well-being for all at all ages (GBD SDG Collaborators, 2015). This goal specifically aims to reduce by one-third premature mortality from noncommunicable diseases through prevention and treatment; promote mental health and well-being; provide access to affordable essential medicines and vaccines to all; and strengthen the implementation of the World Health Organization Framework Convention on Tobacco Control in all countries.

Several other SDGs also affect chronic diseases, including SDG11 (sustainable cities and communities), SDG12 (sustainable consumption and production), SDG8 (decent work and economic growth), SDG5 (gender equality), SDG10 (reduced inequalities), SDG4 (inclusive and quality education for all and lifelong learning), and SDG17 (multi-stakeholder partnerships for the goals).

The United Nations has also launched a Global Compact (UNGC) that engages 8,600 businesses and 4,000 nonbusiness organizations from more than 170 countries. These partners have pledged to work to advance responsible and sustainable business practices, and invest in solutions for tackling societal challenges. The UNGC has recently announced plans to advance the SDGs by convening and catalyzing action from within the UN system, and with global businesses in the workplace, marketplace, and community.

World Health Organization

WHO is the dedicated specialized UN agency for health. It engages in chronic disease control both directly and through partnerships with other actors, such as member-states' governments. For example, under the leadership of Gro Harlem Brundtland, the Framework Convention on Tobacco Control (FCTC) was adopted in 2003 as the first international treaty on tobacco control. Since 2003, 177 countries have worked to implement the provisions (WHO, 2003b). More recently, the WHO introduced a Framework for Engagement with Non-State Actors (FENSA), which includes engagement by actors, including the private sector, NGOs, philanthropic foundations, and academia (WHO, 2016e).

Beyond the FCTC, WHO has pioneered a variety of new resolutions related to chronic diseases. In May 2016, the World Health Assembly agreed to a resolution on healthy aging. This agreement entails the creation of a global strategy and action plan on healthy aging between 2016 and 2020, with the overarching aim of having every country commit to actions that support older adults. For example, countries are expected to foster age-friendly environments, provide equitable and sustainable long-term care systems, and implement improved monitoring and research (WHO, 2015b). Stakeholders are further mobilizing to propose a resolution on physical activity to the World Health Assembly.

Lastly, WHO has led commissions related to chronic diseases and their associated risk factors. For example, the Commission on Ending Childhood Obesity was established in 2014. Its mandate was to "review, build upon and address gaps in existing

mandates and strategies in order to prevent infants, children and adolescents from developing obesity."

Donor Agencies

Despite the prevalence of chronic diseases, donor agencies (both governmental and civil society) have to date provided scant funding to prevention and treatment of chronic diseases (Greenberg et al., 2016) for various reasons. Most notably, interventions for chronic disease control are complex and require sustained efforts over multiple years. Their targets are, therefore, not usually achievable within the confines of a grant agreement or electoral cycle, leading to underfunding even in affluent nations. Perhaps for this reason, less than 10% of the U.S. National Institutes of Health's (NIH) $30 billion budget is allocated to behavioral interventions that target major modifiable chronic disease risk factors (Calitz, Pollack, Millard, & Yach, 2015). Similarly, the U.S. Agency for International Development (USAID) has spent almost nothing supporting chronic disease control as a donor agency, for the reasons cited in **EXHIBIT 8-10** (Holmes, 2016). Beyond the NIH and USAID, the total budget from the U.S. government for chronic diseases was $0 in 2013 (Ralston, Reddy, Fuster, & Narula, 2016).

EXHIBIT 8-10 USAID's Reasons for Neglecting Chronic Diseases

- **Our current job is not done.** Current health priorities are challenging enough.
- **Chronic diseases occur only in higher-income countries and populations.** This pervasive myth is still espoused within USAID.
- **Chronic diseases are not an urgent problem.** Chronic diseases are not an issue today, but rather a future health priority.
- **The current focus works well and is well funded.** There is no need to adapt or evolve, as existing health priorities demand a reliable funding stream.
- **Congress says no.** There is a perception that the U.S. Congress does not want USAID to explore new topics.
- **There is no funding.** Funding is in short supply in an era of financial austerity.
- **There are no good solutions and chronic diseases are not USAID's strong point.** Chronic diseases are complex and require a multisectoral approach.
- **USAID is contributing in some way.** USAID's efforts may spill over to chronic diseases.
- **The chronic disease community is easy to ignore.** The chronic disease community is too quiet, pedestrian, and polite to demand more action on this front.

Data from Holmes, P. (2016). USAID: Standing by on NCD. *Global Heart, 11*(4), 425–427.

Foundations and Nongovernmental Organizations

The largest source of global chronic disease funding is foundations and NGOs, which collectively provided 46% of total external support in 2015 (Nugent, 2016). Nonetheless, many remain skeptical about the extent to which these stakeholders can facilitate improvements in health. Much of their work has focused on educational and awareness materials, which are often ineffective in modifying behaviors. Additionally, some argue that behavioral change, and the desire to change, resides at the individual—and not at the institutional—level.

Among foundations, the Robert Wood Johnson Foundation (RWJF) is the largest philanthropic organization in the United States dedicated solely to better health. With an endowment of $9.2 billion, RWJF is working to build a culture of health by placing well-being at the center of every aspect of life. Bloomberg Philanthropies, founded by Michael Bloomberg, focuses on ensuring better and longer lives for the greatest number of people. Together with WHO, Bloomberg Philanthropies has established the Partnership for Healthy Cities, a global network of cities with $5 million in funding to develop and deliver interventions to save lives through the prevention of chronic diseases.

NCD Alliance is a network of 2,000 NGOs focused on chronic diseases. They have emerged as a recognized global leader on chronic disease policy, a convener of civil society movements, and a partner to governments and UN agencies (NCD Alliance, 2016). The Young Professionals Chronic Disease Network (YP-CDN) is another leading example of an organization founded (in 2009) to advocate for chronic diseases. Since its inception, YP-CDN has expanded to more than 7,000 people, and focuses largely on facilitating access to essential medicines in LMICs. Its petitions on essential medicines have led to the addition of 2% of all medicines to the WHO Essential Medicines List (Young Professionals Chronic Disease Network, 2016).

Academic Institutions

In 1913, the Rockefeller Commission's Welch–Rose Report advocated for the development of schools of public health to fight population-level causes of disease. Shortly thereafter in 1916, the Rockefeller

TABLE 8-17 Supporters and Opponents of Selected Policy Issues

Major Risk Factor	Policy Issue	Supporters	Opponents
Tobacco	Excise tax	Finance, World Bank	Tobacco industry
	Advertising bans	N/A	Advertisers, media, libertarians
	Smoke-free areas	Restaurants	Hospitality and restaurant sector
	Agricultural subsidies	Enlightened countries	Farmers, rural voters
	Advertising to children	N/A	Advertisers, multinational food companies, media
Diet, nutrition	Commodity changes, sugar	Fruit, vegetable farmers	Sugar farmers, producers, lobbyists
	Promotion to children	Sports, gyms	Sports, toys, fast-food industries

N/A = not applicable.

Foundation provided a grant to found the first school of public health, at Johns Hopkins University. To date, public health schools have largely focused on infectious diseases. In the United States, chronic diseases remain absent from the curriculum in schools of public health as well as medicine. Many other countries and regions, including Europe (especially the United Kingdom) and Canada, have made more progress in targeting chronic diseases in academia relative to the United States.

Partnerships for Chronic Disease Control: Challenges and Conflicts

Progress in infectious disease control is achievable with additional funding for surveillance, effective treatment, and new drugs and vaccines. It also requires general improvements in housing and related infrastructure (water, sanitation, and energy). None of these investments are considered politically contentious. In contrast, chronic disease prevention and control requires public health partners to urge individuals to stop smoking; eat less fat, sugar, and salt; engage in more physical activity; and eat more fruits, vegetables, nuts, and grains. These messages, and the policies that promote them, often prove controversial or even unacceptable to other stakeholders, such as private industry and its advocacy groups.

TABLE 8-17 summarizes policy issues prone to such conflict, which often requires resolution at a level of government outside of the health department. The table lists traditional opponents to healthy public policies, along with possible supporters of change. As noted earlier, tobacco use, unhealthy diet, and physical inactivity are major contributors to the burden of disease in both high-income countries and LMICs.

For example, the call to stop smoking has unified a variety of opposition groups—the tobacco industry itself, the hospitality and entertainment industries, farmers, advertising companies, and even some media outlets—frequently around false or misleading messages related to the economic impacts of smoking. Opponents of tobacco control have perpetuated myths including the following:

- *Restaurants and pubs will go bankrupt if smoke-free policies are introduced.* The opposite is the truth.
- *Finance departments will lose revenue from smuggling and reduced tax receipts if excise taxes increase.* Although smuggling may have a small impact on revenue, it can be lessened through customs and excise controls.
- *Tobacco farmers in countries such as Zimbabwe and Malawi will be unemployed in a few years as tobacco consumption declines.* Unfortunately from a public health perspective, tobacco consumption rarely drops faster than 2% per year, even in the face of

vigorous tobacco control campaigns. At those rates, and in the face of continued population growth, the demand for leaf tobacco will remain high for many decades. The immediate threat to African tobacco farmers comes more from the introduction of mechanization locally and subsidies in high-income countries, rather than from less smoking.

Similar conflicts have emerged around sugar control policies. During the development of the WHO Global Strategy on Diet, Physical Activity and Health, simple messages regarding moderate sugar consumption faced fierce opposition (A. Waxman, 2004; H. Waxman, 2004). Soda manufacturers have led the global efforts to deny that sugar causes obesity or dental caries, or that specific levels of sugars are desirable. Sugarcane farmers in LMICs are concerned that if new WHO/Food and Agriculture Organization (FAO) guidelines on sugar consumption were applied globally, they would lose their jobs. The economic stakes are greater for sugar than tobacco: The lobbying community is far larger and the evidence base on the economics of sugar use is not as well described.

For several years, proponents of antismoking campaigns prepared for critiques of any efforts to control tobacco use. These included engaging with the FAO, World Bank, International Monetary Fund (IMF), United Nations Children's Fund (UNICEF), and other UN partners, and jointly agreeing that the key tobacco policy goal was demand reduction. This harmony was later reflected in subsequent policy agreements. Similar work has yet to emerge with respect to many food policy issues, which are inherently more complex and involve a wider array of stakeholders. But as with tobacco (and opposite to what some lobbyists maintain), there will be only very modest impacts on the demand for sugar over the next few decades even if the WHO recommendations are fully implemented (LMC International, 2004).

New Initiatives

Access Accelerated. Twenty biopharmaceutical companies have partnered with the World Bank and the Union for International Cancer Control to address gaps in essential medicine access through a new initiative dubbed "Access Accelerated," which addresses key barriers among low-income populations. The initiative combines public, private, and philanthropic endeavors, and is evaluated by Boston University's Global Health department.

Resolve. In 2017, the former director of the U.S. Centers for Disease Control and Prevention (CDC) announced a $225 million initiative to address cardiovascular disease and epidemics through an initiative hosted at Vital Strategies, known as "Resolve." The initiative aims to save up to 100 million lives by scaling up proven interventions, including improved treatment of high blood pressure, elimination of artificial trans fats, and reduction in sodium consumption.

Other Collaborative Networks

Advocates for health promotion and chronic disease prevention have yet to coalesce around common messages or key themes. Different stakeholders support different interests within health. Until recently, public health would not collaborate with business, while policy makers often focused on curative care at the expense of chronic disease prevention.

Businesses are increasingly forming precompetitive partnerships to collectively advocate for better health. One example is Bupa's Chief Medical Officer (CMO) Network. This group convenes world-class clinicians to address the world's largest problems through workplace health initiatives. It has decided to focus on behavior change within the workplace as the first major target area. Other examples include the World Economic Forum's Council on Human-centric Health and the American Heart Association's CEO Roundtable, which productively engages for-profit and nonprofit partners to fight chronic disease.

Future Drivers for Prevention and Management of Chronic Diseases

Although existing chronic disease control work has been limited to date, and remains fraught by complex political and economic conflicts, multiple initiatives have demonstrated effective chronic disease control strategies. These tend to involve several common health system strategies, centered on the judicious use of scarce resources and on the structuring of care around prevention in addition to cure.

Management of Chronic Diseases by Healthcare Systems

Country governments, NGOs, businesses, and other international stakeholders have multiple tools to target the *population-level* determinants of chronic diseases (such as rising consumption of tobacco and sugar-sweetened beverages) using *population-level* interventions such as taxes and public education

campaigns. Nonetheless, local and national health systems must also respond to chronic diseases—through both prevention and treatment—at the level of *individual patients* and their communities. In lower- and higher-income countries, these interventions are still evolving. Most take aim at preventing or treating the core lifestyle risk factors for chronic disease (poor diet, tobacco use, sedentary activity), treating and preventing diseases that emerge from these conditions (hypertension, obesity), and treating and preventing the chronic conditions that emerge in turn from those disorders (congestive heart failure, stroke). Successful programs have taken many forms, but display several common features:

- The use of non-physician healthcare workers (NPHWs), including community health workers (CHWs)
- Leveraging or expanding existing programs for other diseases, such as HIV, to adapt to new conditions such as cardiovascular disease
- Using models of care that address multiple chronic disease simultaneously, such as hypertension and diabetes

These approaches use scarce resources effectively in LMICs, but are increasingly being applied to chronic disease care in high-income countries as well. Further research is needed on the most efficient way to design and evaluate these interventions.

In LMICs, there is a profound shortage of physicians, nurses, and other qualified health providers for patients with both chronic and acute conditions (Kar, Thakur, Jain, & Kumar, 2008; WHO, 2006b). In high-income countries, these providers are more abundant, but are often disproportionately involved in specialist care rather than the primary care activities most germane to prevention of chronic disease risk factors, and tend to concentrate in urban areas (Bodenheimer & Smith, 2013; Goodyear-Smith & Janes, 2008). As a result, local health systems have used staff with less formal training to complete tasks normally shouldered by these providers, a process called task-shifting (Callaghan, Ford, & Schneider, 2010; Federspiel et al., 2015; Joshi et al., 2014).

Task-shifting is not a new concept: Nonphysicians functioned as health officers in nineteenth century France, and so-called barefoot doctors provided widespread medical care to rural China in the mid-twentieth century. This work inspired the universal primary care movement of the late 1970s (Heller, 1978; Sidel, 1972) and many of the programs that grew out of it (Black et al., 2017; Mann, Eble, Frost, Premkumar, & Boone, 2010). However, the concept rose to prominence in the mid-2000s as a vehicle for the control of HIV/AIDS in lower-income countries, when the uncontrolled epidemic vastly outstripped the human resources available for its containment ("The 3 by 5 Initiative," n.d.; WHO, 2008b). By 2010, a systematic review of task-shifting approaches to the control of HIV in Africa found 51 separate studies with measurable outcomes (Callaghan et al., 2010). These largely demonstrated that NPHWs can increase access to medication and other treatment, at comparable cost-effectiveness to physician care models and with no change in quality of care.

Based on these data, delivery programs have increasingly used NPHWs for the screening and control of chronic diseases, with similar results. For example, studies in India and Pakistan have demonstrated that nonphysician health workers can screen for persons at high risk of cardiovascular disease with good accuracy, agreeing with physicians in 89% of cases (Abegunde et al., 2007; Joshi et al., 2012). Further work has corroborated this result in South Africa, Guatemala, Mexico, and Bangladesh (Gaziano et al., 2015; Gaziano et al., 2013), and shown its feasibility for other chronic conditions such as depression (Adams, Almond, Ringo, Shangali, & Sikkema, 2012) and oral and cervical cancer (Gajalakshmi, Krishnamurthi, Ananth, & Shanta, 1996; Swai et al., 2005; Warnakulasuriya et al., 1984). For some conditions, such as cardiovascular disease, simplified screening algorithms have facilitated this process, both by avoiding blood tests and other often-unavailable investigations, and by approximating complex decision matrices with a few key criteria that NPHWs can apply even with incomplete knowledge of medical physiology (Fernandez-Alvira et al., 2017; Gaziano, Young, Fitzmaurice, Atwood, & Gaziano, 2008; WHO, 2017a).

NPHWs can also provide treatment for chronic conditions using related algorithms. These include prescribing medications for conditions including hypertension (Kengne, Awah, Fezeu, Sobngwi, & Mbanya, 2009; Ogedegbe et al., 2014; Vedanthan et al., 2017), diabetes (Kengne et al., 2009; Labhardt, Balo, Ndam, Grimm, & Manga, 2010), and asthma (Kengne et al., 2008), as well as counseling for conditions such as depression (Chibanda et al., 2011; Patel et al., 2010) and tobacco cessation (Siddiqi et al., 2013). Increasingly, WHO and other international organizations have developed algorithms that leverage task-shifting (as well as task-sharing, in which doctors aid NPHWs in providing care) for cost-effective integrated management of chronic diseases (WHO, 2004, 2017a). In part as a result of these successes, models using CHWs and other NPHWs to help patients control chronic conditions are now increasingly common in

high-income countries as well. One example is City Health Works (2017) in New York City.

The optimal role of NPHWs in chronic disease care is incompletely understood, and likely depends on the context of the local health system. Systematic reviews have demonstrated these workers' efficacy (Joshi et al., 2014; Mutamba, van Ginneken, Paintain, Wandiembe, & Schellenberg, 2013; Khetan et al., 2017), but also document such barriers as skepticism toward permitting NPHWs to prescribe medications, lack of sufficient supplies at the NPHW clinic level, limited formal NPHW training, and difficulty tracking patient records (Joshi et al., 2014). Conversely, health systems that provide short, practical NPHW trainings; that permit these workers to have sufficient autonomy; and that ensure they have reliable supplies for practice have reported more success (Joshi et al., 2014). NPHWs are an essential tool for chronic disease care, but their role must be properly clarified and adapted to the context of the local care delivery system, in both high-income countries and lower-income countries.

Leveraging or Expanding Programs for Other Conditions

Task-shifting and NPHW-led care rose to recent prominence during the mid-2000s when the HIV/AIDS pandemic in LMICs was at its height and medications and human resources were severely inadequate to cope with the disease burden imposed by that pandemic. In part due to improved clinical outcomes in persons living with HIV and AIDS (PLWHA), chronic conditions have become more prevalent in this population as they age (Kwarisiima et al., 2016; Rabkin, Kruk, & El-Sadr, 2012). Moreover, HIV infection and treatment are associated with many chronic conditions, such as cardiovascular disease, stroke, and multiple types of cancer (Narayan et al., 2014). Additionally, the mechanics of modern HIV/AIDS care—which now usually involves universal screening, lifestyle counseling, and medical treatment for all persons with a positive test—closely resemble validated approaches to the control of chronic conditions such as hypertension and diabetes (Chamie et al., 2012).

As a result, HIV/AIDS programs in LMICs are increasingly leveraging their resources to concurrently screen for chronic conditions in both HIV-positive and HIV-negative persons using physicians and NPHWs. The Sustainable East Africa Research in Community Health (SEARCH) study, for example, is a massive randomized trial of 300,000 persons in Uganda and Kenya aimed at establishing whether screening entire communities for HIV/AIDS—and ensuring adequate treatment of all PLWHA—can arrest the spread of HIV (SEARCH, n.d.). SEARCH has also succeeded in screening tens of thousands of people for hypertension and diabetes in the communities, and consistently linked those persons with a positive test to follow-up care (Chamie et al., 2012; Kotwani et al., 2014). Other studies have also shown the efficacy of this model in other contexts (Janssens et al., 2007) and for other conditions such as depression (Adams et al., 2012). AMPATH's model, for example, has focused on HIV and chronic diseases such as Burkitt's lymphoma since its inception in the 1990s, and now provides NPHW-led hypertension care (Vedanthan et al., 2014). Although there are limited examples, recent meta-analyses have demonstrated that these programs are largely effective (Haldane et al., 2018). Predictors of success included the use of multidisciplinary teams such as pharmacists and other NPHWs as well as engagement with community leaders and members (Haldane et al., 2018).

Programs for chronic conditions can also be integrated into other existing care structures such as programs related to maternal and child health. As an example, a study is under way in northern Ghana that will leverage the Community-Based Health Planning and Services (CHPS) program, which has used community health nurses to decrease maternal and under-5 child mortality through education and referral, for the screening and treatment of risk factors for cardiovascular disease (Phillips, 2016–2018).

Integrating Multiple Chronic Conditions

Achieving control of chronic conditions in resource-poor settings requires careful parsimony—for example, using nonphysicians when doctors are unavailable, or using existing health programs rather than building new ones. Further efficiency is achievable by focusing on more than one chronic condition using the same cadre of staff and resources. Cardiovascular epidemiologic data demonstrate that it is less effective to treat persons with each major cardiovascular disease risk factor separately (such as high blood pressure and smoking) than to calculate the overall risk of an individual based on these factors and provide integrated care for all such persons (WHO, 2007). WHO's (2017a) HEARTS initiative is aimed at providing such care, although controversy remains over whether it has addressed all elements of heart disease care (Kishore, Heller & Vasan, 2018). A predecessor of the HEARTS protocol was implemented successfully in China and Nigeria (Mendis et al., 2010), and studies are under way in Jordan and Uganda, among other sites, to implement this approach (Collins et al., 2017; Sandy Gove, personal communication, 2017).

Although models like HEARTS for the integrated care of diabetes, hypertension, and other cardiovascular disease risk factors have become more common (Chamie et al., 2012; Coleman, Gill, & Wilkinson, 1998; Labhardt et al., 2010), programs that integrate cardiovascular disease care with care of other chronic conditions such as asthma or epilepsy remain rare (Kengne et al., 2008; Labhardt et al., 2010). In high-income countries, where chronic diseases already constitute nearly the entire disease burden, persons who suffer from multiple concurrent chronic conditions often experience poor health outcomes and struggle to access consistent preventive care. In addition to CHW-led models, new research aims to build primary care systems focused on more effective management.

Characteristics of Effective Systems of Care for Chronic Diseases

Health care for patients with chronic conditions requires a fundamental change in perspective from the familiar acute care model. The magnitude of this challenge is eased somewhat by the fact that chronic conditions share many common features. Whereas biomedical management changes depending on the unique features of the specific disease, the general components of care organization and delivery for patients with chronic conditions are essentially the same. These components include a well-defined care plan, patient self-management, scheduled follow-up appointments, monitoring of outcome and adherence, and stepwise treatment protocols. Collectively, these approaches represent a significant shift in healthcare practices. The differences between typical current and desired future approaches are described in the following subsections.

Patient-Centered Care

Patient-centered care recognizes the patient as a person; fully informs patients about the risks and benefits of treatment options; tailors decision making in response to individual patients' values, needs, and expressed preferences; shares power and responsibility among patients and providers; and develops patients' abilities to participate in their care. Across its multiple meanings, research shows that patient-centered care is crucial for obtaining good outcomes for chronic conditions.

Several experiences within LMICs have demonstrated the utility of patient-centered care across diverse cultures and resource contexts. WHO's Integrated Management of Adult Illness (IMAI) general principles of chronic care, for example, focus on equipping first-level healthcare workers to provide patient-centered health care. Specifically, the guidelines and related training materials prepare such workers to solicit patients' concerns and preferences, work in collaboration with patients to decide specific goals and treatment plans, and support patients in their daily efforts regarding prevention, medication adherence, and self-management.

Early results indicate that this approach is understandable and usable by first-level health workers. To date, physicians, nurses, and lay providers have been trained in this approach in Burkina Faso, Burundi, Ethiopia, Sudan, and Uganda. In Shanghai, China, a community-based chronic disease self-management program was shown to improve health status and reduce hospitalizations among patients with hypertension, heart disease, chronic lung disease, arthritis, stroke, and diabetes. Participants learned to take responsibility for the day-to-day management of their disease and the physical and emotional problems caused by their disease. The program was led by lay people with chronic conditions, who followed a detailed leader's manual throughout the program (Fu et al., 2003).

Emphasis on Primary Health Care

In LMICs, patients with chronic conditions present and need to be managed mainly at the primary healthcare level. This represents a departure from the approach used in healthcare systems that are driven by tertiary-care, specialty settings. Oman has successfully made the shift to a decentralized primary healthcare system, with health programs and activities in this country now being coordinated with the regional health services via referrals and linkages (WHO, 2002a). Similarly, the health policy of the Islamic Republic of Iran is based on primary health care, with particular emphasis on the expansion of health networks and programs in rural areas.

Population-Based Care

Health care for chronic conditions is most effective when policies, plans, and practices prioritize the health of a defined population rather than the single unit of a patient seeking care. A population focus implies that healthcare systems assess and monitor the health of communities, emphasize prevention and promote healthy behavior, assure universal access to appropriate and cost-effective services, and contribute

to the evidence base for effective treatments and systems of care.

Cuba's family doctor program is a notable application of population-based care. Each family doctor is responsible for the general health of the entire population in a small, defined area. Physicians are expected to provide preventive, maternal, and curative services to children and adults, and to monitor all patients with chronic conditions. They live in the communities that they serve, often residing in the same apartment block as their patients. In addition to engaging in medical consultations, Cuba's family doctors play an active role in promoting health among the communities they serve. They provide informal advice and counseling to community members, and they run regular prevention and self-management groups concerning a range of issues. They are also expected to set a positive example for their patients in the conduct of their day-to-day lives (Warman, 2001).

Proactive Care

Proactive care anticipates patients' needs rather than relying on a patient-initiated, often acute care–focused interaction. In rural South Africa, a proactive noncommunicable disease management program for hypertension, diabetes, asthma, and epilepsy was established within primary health care. This program emphasized planned care: Clinic-held treatment cards and registries were introduced, and diagnostic and management protocols were followed, which included regular, planned follow-up with a clinic nurse. Using this proactive care approach, nurses were able to achieve good disease control among most of the patient population—68% of patients with hypertension, 82% of those with diabetes, and 84% of those with asthma (Coleman et al., 1998).

A Model of Care: Innovative Care for Chronic Conditions

The WHO has developed a model based on these approaches. The Innovative Care for Chronic Conditions (ICCC) framework provides a roadmap for decision makers who want to improve their health system's capacity to manage chronic conditions (WHO, 2002b) (**FIGURE 8-5**). This framework is composed of fundamental components within the levels of patient interactions, organization of health care, community, and

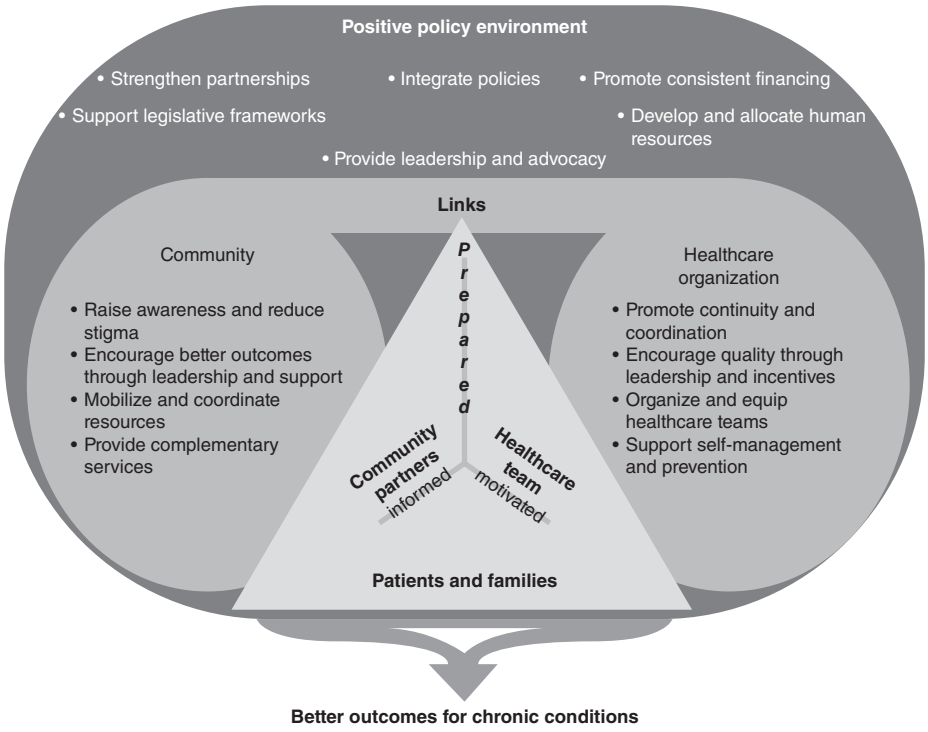


FIGURE 8-5 The World Health Organization's Innovative Care for Chronic Conditions framework.

Reprinted from World Health Organization (WHO). (2002). *Innovative care for chronic conditions: Building blocks for action.* Retrieved from http://www.who.int/chp/knowledge/publications/icccglobalreport.pdf

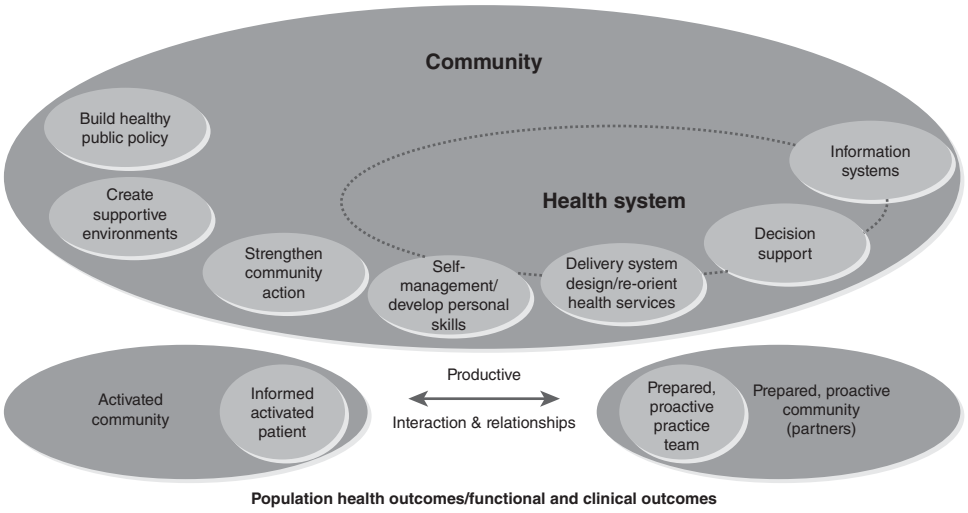


FIGURE 8-6 The expanded chronic care model.

Reproduced with permission from Barr, V. J., Robinson, S., Marin-Link, B., Underhill, L., Dotts, A., Ravensdale, D. et al. (2003). The expanded Chronic Care Model: An integration of concepts and strategies from population health promotion and the Chronic Care Model. *Hospital Quarterly, 7,* 73–82.

policy. These components are described as building blocks that can be used to help decision makers progressively create or redesign a healthcare system to expand its capacity to manage long-term health problems. Although the framework does not prescribe specific changes that must be tailored to unique needs and resources, it highlights the need for comprehensive system design or change in the requirements for effective care.

Barr et al. (2003) have also proposed an expanded chronic care model (**FIGURE 8-6**) that emphasizes the inclusion of elements of the population health promotion field so that broad-based prevention efforts, recognition of the social determinants of health, and enhanced community participation can all be part of the work of health system teams as they address the prevention and control of chronic diseases. **TABLE 8-18** compares the components of the expanded chronic care model with those of the chronic care model developed by Wagner and colleagues (1996).

Comprehensive Care Applied to Clinical Prevention

The preponderance of evidence suggests that effective healthcare strategies for reducing risk do exist, but tend to be weakly implemented (Coffield et al., 2001). Many professional competencies for delivering effective clinical prevention are outside the scope and culture of clinical medicine, so healthcare professionals frequently have little or no training in the skills required to improve care (Glasgow, Orleans, & Wagner, 2001).

Many opportunities for better integration of the treatment of chronic diseases and the prevention of risk factor behaviors remain largely untapped. For example, stronger support for smoking cessation among patients with TB would save lives in the long term, and smoking cessation among patients with congenital heart disease is the single most effective intervention for reducing mortality in these patients who smoke. As demonstrated by the 36% reduction in the relative risk of mortality among patients with congenital heart disease who quit tobacco use, smoking is at least as important as other secondary prevention measures such as use of statins (a 29% reduction), aspirin (15%), beta blockers (23%), or angiotensin-converting enzyme (ACE) inhibitors (23%) (Critchley & Capewell, 2003). Smoking cessation is also a priority for people with mental disorders. In one of the major studies on comorbidity between tobacco use and depression, researchers found that people with mental disorders are almost twice as likely to smoke as individuals without such disorders. Further, they found that people with a mental disorder had consumed 44.3% of all cigarettes smoked by a nationally representative sample in the previous days (Lasser et al., 2000).

Another challenge is that effective clinical prevention services must extend beyond the mere provision of information to patients. The era of exhortation by healthcare professionals to "eat better" or "drop some weight" is long past: Modern, evidence-based interventions emphasize shared decision making and collaborative goal setting among providers and patients (Serdula, Khan, & Dietz, 2003). The clearer and more personalized

TABLE 8-18 Comparison of the Chronic Care Model with the Expanded Chronic Care Model

Components of the Chronic Care Model		Components of the Expanded Chronic Care Model		Examples
Health system—organization of health care	Program planning that includes measurable goals for better care of chronic illness			
Self-management support	Emphasis on the importance of the central role that patients have in managing their own care	Self-management/develop personal skills	Enhancing skills and capacities for personal health and wellness	■ Smoking prevention and cessation programs ■ Seniors' walking programs
Decision support	Integration of evidence-based guidelines into daily clinical practice	Decision support	Integration of strategies for facilitating the community's abilities to stay healthy	■ Development of health promotion and prevention "best practice" guidelines
Delivery system design	Focus on teamwork and an expanded scope of practice to support chronic care	Delivery system design/reorient health services	Expansion of mandate to support individuals and communities in a more holistic way	■ Advocacy on behalf of (and with) vulnerable populations
Clinical information systems	Developing information systems based on patient populations to provide relevant client data	Information system	Creation of broadly based information systems to include community data beyond the healthcare system	■ Emphasis in quality improvement on health and quality of life outcomes, not just clinical outcomes ■ Use of broad community needs assessments that take into account: • Poverty rates • Availability of public transportation • Violent crime rate
Community resources and policies	Developing partnerships with community organizations that support and meet patients' needs	Build healthy public policy Create supportive environments Strengthen community action	Development and implementation of policies designed to improve population health Generating living and employing conditions that are safe, stimulating, satisfying, and enjoyable Working with community groups to set priorities and achieve goals that enhance the health of the community	■ Advocating for/developing: • Smoking bylaws • Walking trails • Reductions in the price of whole-wheat flour • Maintaining older people in their homes for as long as possible • Work toward the development of well-lit streets and bicycle paths • Supporting the community in addressing the need for safe affordable housing

Reproduced with permission from Barr, V. J., Robinson, S., Marin-Link, B., Underhill, L., Dotts, A., Ravensdale, D. et al. (2003). The expanded Chronic Care Model: An integration of concepts and strategies from population health promotion and the Chronic Care Model. *Hospital Quarterly, 7*, 73–82.

the goal, the better (Estabrooks, Glasgow, & Dzewaltowski, 2003). Skill building to overcome barriers, self-monitoring, personalized feedback, and systematic links to community resources such as peer support groups are other important elements for success (Steptoe et al., 2003). Many healthcare settings deliver these kinds of interventions in group formats, which enhances the efficiency of healthcare professionals and provides the added element of social support (Noel & Pugh, 2002).

Comprehensive Care Applied to Adherence

The ICCC framework recognizes the importance of treatment adherence as a primary determinant of the effectiveness of treatment. Good adherence confers both health and economic benefits. Adherence has been associated with improved blood pressure control (Luscher, Vetter, Sigenthaler, & Vetter, 1985) and lessened complications of hypertension (Morisky et al., 1983; Psaty, Koepsell, Wagner, Lo Gerfo, & Inui, 1990). Despite the clear importance of treatment adherence, a number of rigorous reviews have found that in high-income countries, adherence among patients with chronic diseases averages only 50%; it is even lower in LMICs. In Gambia, China, and the United States, for example, only 27%, 43%, and 51%, respectively, of patients adhere to their medication regimen for high blood pressure. Similar patterns have been reported for other conditions, such as depression (range of 40% to 70%), asthma (43% for acute treatments and 28% for maintenance), and HIV/AIDS (range of 37% to 83%) (WHO, 2003a).

Adherence is a complex behavioral process that is influenced by five interacting dimensions: social and economic factors, healthcare system factors, condition-related factors, therapy-related factors, and patient-related factors. Because each dimension plays an important role in determining adherence rates (WHO, 2003a), all of them should be considered when designing interventions to improve outcomes. The most effective interventions have been shown to be multilevel, targeting more than one factor with more than one intervention (Dickinson, Wilkie, & Harris, 1999).

Contemporary perspectives have pointed out the importance of conceptualizing adherence as the active, voluntary involvement of the patient in the management of his or her disease, including a mutually agreed-upon course of treatment and sharing of responsibility between the patient and healthcare providers (Flood & Chiang, 2001). According to these perspectives, adherence is an active, responsible, and flexible process of self-management, in which the person strives to achieve good health by working in close collaboration with healthcare staff instead of simply following rigidly prescribed rules.

Summary of Effective Health Care

Reviews of interventions to improve health care for chronic conditions have demonstrated the importance of using multifaceted approaches as opposed to "magic bullet" or "single lever" interventions (**TABLE 8-19**) (Grimshaw et al., 2001; Renders et al., 2002; Wagner et al., 2001). Models of integrated, coordinated care, such as the ICCC framework, capture this complexity in an organized way.

Several key concepts have emerged from research within this area. First, it is necessary to work across multiple levels in a coordinated fashion to effect meaningful change in health care for chronic diseases. Second, organized systems of care—not just individual healthcare workers—are essential in producing positive outcomes for chronic disease. Third, it is crucial to work across the disease continuum in a comprehensive way. Comprehensive care for chronic conditions must span the full range of phases from clinical prevention, to treatment, to rehabilitation, to palliation (see **EXHIBIT 8-11**).

Healthcare Access and Quality

WHO has introduced a range of tools to support the management of chronic diseases that aim to impact healthcare access and quality. These have included the Package of Essential Medicines and Technologies for Non-Communicable Disease Interventions (PEN) and the HEARTS protocol, whose initial emphasis is on cardiovascular health. The HEARTS protocol includes the following elements:

H: healthy lifestyle

E: evidence-based treatment protocols

A: access to essential medicines and technologies

R: risk-based management

T: task-shifting and team-based care

S: systems for monitoring (WHO & CDC, 2017)

These guidelines, toolkits, and roadmaps are being co-developed for use at four levels: community, facility, district, and national, and are meant to guide decision making in the face of a paucity of domestic financing.

A growing body of work on implementation science in the public and private sectors has emerged to

TABLE 8-19 Stepwise Policy and Program Targets for National Prevention and Control of Chronic Diseases

Resource Level	Population Approaches		Individual High-Risk Approaches
	National Level	Community Level	
Step 1: Core	The WHO Framework Convention on Tobacco Control (FCTC) is ratified in every country. Tobacco control legislation consistent with the elements of the FCTC is enacted and enforced. A national nutrition and physical activity policy consistent with the global strategy is developed and endorsed at the cabinet level, including laws. Health impact assessment of public policy is carried out; priority areas include transport, urban planning, taxation, trade, and agriculture.	Local infrastructure plans include the provisions for and maintenance of accessible and safe sites for physical activity (e.g., parks and pedestrian-only areas). Health-promoting community projects include participatory actions to cope with the environmental factors that increase individuals' risk of chronic diseases—inactivity, unhealthy diet, and tobacco and alcohol use. Active health promotion programs focusing on chronic diseases are implemented in different settings (e.g., villages, schools, workplaces) and explicitly aim to reach poor communities.	Context-specific guidelines for chronic disease prevention and control have been adopted and are used at all healthcare levels. A sustainable, accessible, and affordable supply of appropriate medication is assured for priority chronic diseases. A system exists for the consistent, high-quality application of clinical guidelines and for the clinical audit of services offered. A proactive follow-up system for patients with diabetes and hypertension is in operation.
Step 2: Expanded	Tobacco legislation provides for incremental increases in taxes on tobacco, and a proportion of the revenue is earmarked for health promotion. Food standards legislation is enacted and enforced; it includes nutrition labeling. Sustained, well-designed, national programs (counter advertising) are in place to promote nonsmoking lifestyles, consumption of fruits and vegetables, and physical activity. Country standards are established that regulate marketing of unhealthy food to children.	Sustained, well-designed programs are in place to promote tobacco-free lifestyles (e.g., smoke-free public places, smoke-free sports). Healthy diets (e.g., low-cost, low-fat foods; fresh fruit and vegetables). Physical activity (e.g., "movement") in different domains (occupational and leisure).	Systems are in place for selective and targeted prevention aimed at high-risk populations, based on absolute levels of risk. Publicly financed "quit-line" for smokers; weight control line.
Step 3: Optimal	Policies shown to work for chronic disease prevention and control are implemented. There is policy coherence between agricultural systems and chronic diseases. Country standards are established that regulate marketing of unhealthy food to children.	Recreational and fitness centers are available for community use.	Opportunistic screening, case-finding, and management programs are implemented. Self-management groups are fostered for tobacco cessation and overweight reduction. Appropriate diagnostic and therapeutic interventions are implemented.

Modified from World Health Organization (WHO). (2003). *World health report, 2003*. Retrieved from http://apps.who.int/iris/bitstream/10665/42789/1/9241562439.pdf

EXHIBIT 8-11 Case Study: Hypertension

Two of us (DJH and SPK) have argued that an approach to medication access used for HIV/AIDS mobilization should be leveraged for hypertension, the leading risk for cardiovascular disease (Heller & Kishore, 2017). Leveraging lessons learned from HIV/AIDS and infectious disease programs, we believe it is crucial to focus on the gap—and outline steps to close it.

> Antihypertensive medicines, called thiazides, can close the treatment gap for less than $20 per person per year. Over a decade, they could save almost 5 million lives worldwide, at a cost under $9 billion per year: significantly less than the estimates that galvanized the global HIV community 15 years ago. Thiazides are the initial medication of choice for blood pressure control in nearly all settings. Although developed in 1958, they remain a drug of first resort, according to the International Society for Hypertension (ISH), European Society for Cardiology, and others (James et al., 2014; Mancia et al., 2013; WHO, 2007). They are also generic, costing a dollar or less per year in developing countries (Partners in Health, Bukhman, & Kidder, 2011). Although they require annual blood tests and doctor visits to screen for (rare) changes in kidney function, groups such as Partners in Health (2011) have treated patients with hydrochlorothiazide for $4–13 per year in settings like Rwanda.
>
> As in the early days of HIV control, conventional wisdom holds that universal hypertension treatment is prohibitively expensive. In 2007, WHO and ISH published guidelines arguing that only high-risk persons should receive antihypertensive medications, because "treatment of [hypertensive] patients with very low cardiovascular disease risk may be cost-effective only if inexpensive antihypertensive drugs are used" (WHO, 2007). Subsequent studies concurred (Gaziano, 2008; Ilesanmi, Ige, & Adebiyi, 2012), with one case finding many patients spent 10% or more of their annual income on antihypertensive medications (Ilesanmi et al., 2012) Yet recent data suggest universal treatment is now cost-effective and a better value than HIV/AIDS treatment at the dawn of the Global Fund and PEPFAR. Partners in Health (2011) pays $0.3 cents per thiazide tablet in Rwanda or $1.03 per patient-year of treatment. Patients requiring further treatment receive amlodipine, whose annual cost is $4.69. Each patient also receives one to two clinic visits per year at $3 each, as well as laboratory tests costing $0.17 per check. Conservatively, assuming that all patients require two medications, clinic visits, and laboratory tests per year, these results correspond to a per-patient annual cost of $12.90. Mills et al. (2016) report that about 32% of all adults in LMICs have hypertension—approximately 1.039 billion people. Some 11.6% have chronic kidney disease, according to a recent systematic analysis (Mills et al., 2015), or other severe hypertension (9.1% in Partners in Health's cohort) requiring treatment beyond thiazides and amlodipine (Partners in Health, 2011). Of the remaining 830 million, 29% are currently treated (Mills et al., 2016), leaving 589 million persons in need of treatment worldwide.

Reproduced from Heller D. J., Kishore S. P. (2017). Closing the blood pressure gap: An affordable proposal to save lives worldwide. *BMJ Global Health 2*:e000429, with permission from BMJ Publishing Group Ltd.

champion appropriate models. For example, *The Lancet*'s Commission on NCDs and Injuries provides tools for national ministries to execute better care on chronic diseases (*The Lancet* NCDI Poverty Commission, 2017). Moreover, while antiretroviral drugs (ARVs) for HIV/AIDS can cost $3,500 or more annually, access to low-cost, first-line treatments for hypertension such as thiazides could be achieved for less than $20 per person (Partners in Health, Bukhman, & Kidder, 2013).

A Model of Care: Innovative Care for Chronic Conditions

New models exist on how to best tackle chronic conditions in resource-poor communities. As an example, AMPATH in Eldoret, Kenya, has pioneered an innovative model for caring for patients with chronic illnesses (**FIGURE 8-7**). This model is based on the HIV model that is shifting toward population health, to include income, food security, clinical care delivery for chronic diseases, as well as community partnerships and social networks that are durable and long-lasting.

This shift toward population health is being mirrored by other academic initiatives such as global health partnerships amplified by the research outputs of the Fogarty Institute for Innovation and other global health institutes across the globe. A key emphasis of this model is team-based care and task-shifting to ensure that nurse- and community health worker-driven models are scaled appropriately, including in rural settings. Adherence and linkage to care are hallmarks of the approach. These models are patient centered, population based, data driven, and proactive (rather than reactive).

Cost-Effectiveness

Chronic disease interventions, including medications, often have low costs and no patent protection. Many of these interventions are also listed on WHO's Essential Medicines List, which guides national purchasing,

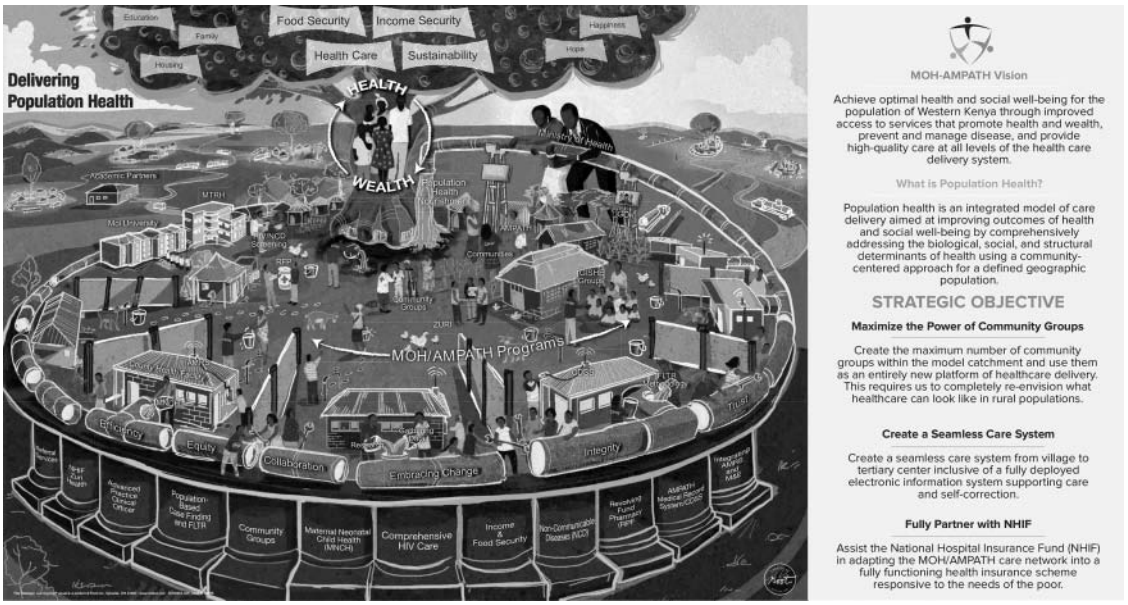


FIGURE 8-7 AMPATH's framework for population health.

The Root Learning Map® visual was created by Root Inc and provided courtesy of AMPATH.

procurement, and distribution at districts and facilities. With the exception of oncologic products (e.g., trastazumab for *HER2*-positive breast cancers), these tools are cost-effective per established WHO standard (cost ÷ DALY is less than 3 times the gross national income per capita).

One fast-growing movement is coalescing around the use of a multidrug regimen in a single fixed-dose combination—known as the polypill—for use for cardiovascular disease. This tool could be rapidly deployed for primary and secondary prevention to promote adherence and deliver clinical results across meaningful population sizes at low cost.

Reorientation of Health Services to Address Chronic Disease

Many lives continue to be lost prematurely because of inadequate treatment and long-term management of chronic diseases, even though simple and inexpensive approaches to address these diseases exist. Even in high-income countries, the full potential of these interventions is not realized. The situation in both poorer countries and poor communities within rich countries is even less satisfactory. In most countries, effective means of preventing, treating, and providing palliative care for cancer exist, but are not broadly implemented. There are many opportunities for coordinated risk reduction, care, and long-term management of chronic disease. For example, smoking cessation is a priority for all patients who smoke; dietary and physical activity information and skill building should be provided to most patients in virtually all healthcare settings. Unfortunately, few efforts have been made to explicitly target poor communities with such interventions.

Considerable progress has been made in improving access to, and reducing the prices of, antiretroviral agents for HIV/AIDS, drugs to treat TB, and several vaccines. Similar progress has yet to be made for essential drugs that are required to improve survival for treatable cancers, diabetes, and cardiovascular disease. A patient with heart disease in a poor nation has the same right to effective drug treatment as a patient with malaria, tuberculosis, or HIV/AIDS. NGOs have yet to advocate as effectively for better access to chronic disease treatment as they have for selected infectious diseases, despite the huge savings in lives and suffering that would result from broader access to such health care.

Continued strengthening of certain aspects of infectious disease control, particularly those related to chronic infectious diseases such as TB and HIV/AIDS, will in turn benefit the control of cardiovascular disease, diabetes, and cancer. The same transformation of healthcare systems is required to address prevention and long-term disease management for both infectious and noninfectious chronic diseases. In sub-Saharan African countries, an opportunity exists to ensure that the new platforms for health services delivery that are being built to expand access to treatment for HIV/AIDS also address noninfectious chronic diseases. The marginal increased investments required to provide this more comprehensive infrastructure would, in all likelihood, yield substantial gains for public health among poor communities whose members already suffer from cardiovascular disease, diabetes, and cancer.

Summary

It is often stated that "An ounce of prevention is worth a pound of cure." Yet too often, treatment is valued over prevention. This emphasis is maintained despite evidence suggesting the cost-effectiveness of prevention interventions compared to treatments for chronic diseases. One major driving factor underlying this phenomenon is the time value of prevention versus treatment. Prevention saves statistical figures in the future, yet treatment saves real people today. In too many instances, this factor leads to treatment being prioritized over prevention.

This chapter has presented the urgent case for acting on chronic diseases today and not tomorrow. It defined and outlined the global prevalence of chronic diseases, their health and economic impact, and their four key risk factors; described stakeholder efforts to date in chronic disease prevention; and proposed stakeholder actions that can further control the rising prevalence of chronic diseases. To fully realize these objectives and to place these actions and activities in context, **FIGURE 8-8** presents a graphic of major milestones in chronic disease prevention. It includes future predictions related to chronic diseases through the year 2030, when the United Nations measures progress against the SDGs.

To assist students planning careers related to chronic diseases, we have also developed a roadmap on potential future developments in health. This roadmap (**EXHIBIT 8-12**) lists multiple often-uncharted paths for action and advocacy against chronic conditions in the context of the SDGs.

Discussion Questions

1. What are the leading risk factors that contribute to death and disability globally? How do these risk factors compare between LMICs versus high-income countries?
2. Given current trends in risk factors, what are likely to be the major causes of death and disability in LMICs versus high-income countries over the next decade?
3. Which policies and actions taken at national and international levels could influence trends in chronic disease and their associated risk factors? Which indicators might facilitate their implementation?
4. Describe how globalization could be positively harnessed for chronic disease prevention. In doing so, consider how public–private partnerships involving major multinational corporations could play a more effective role in promoting health.

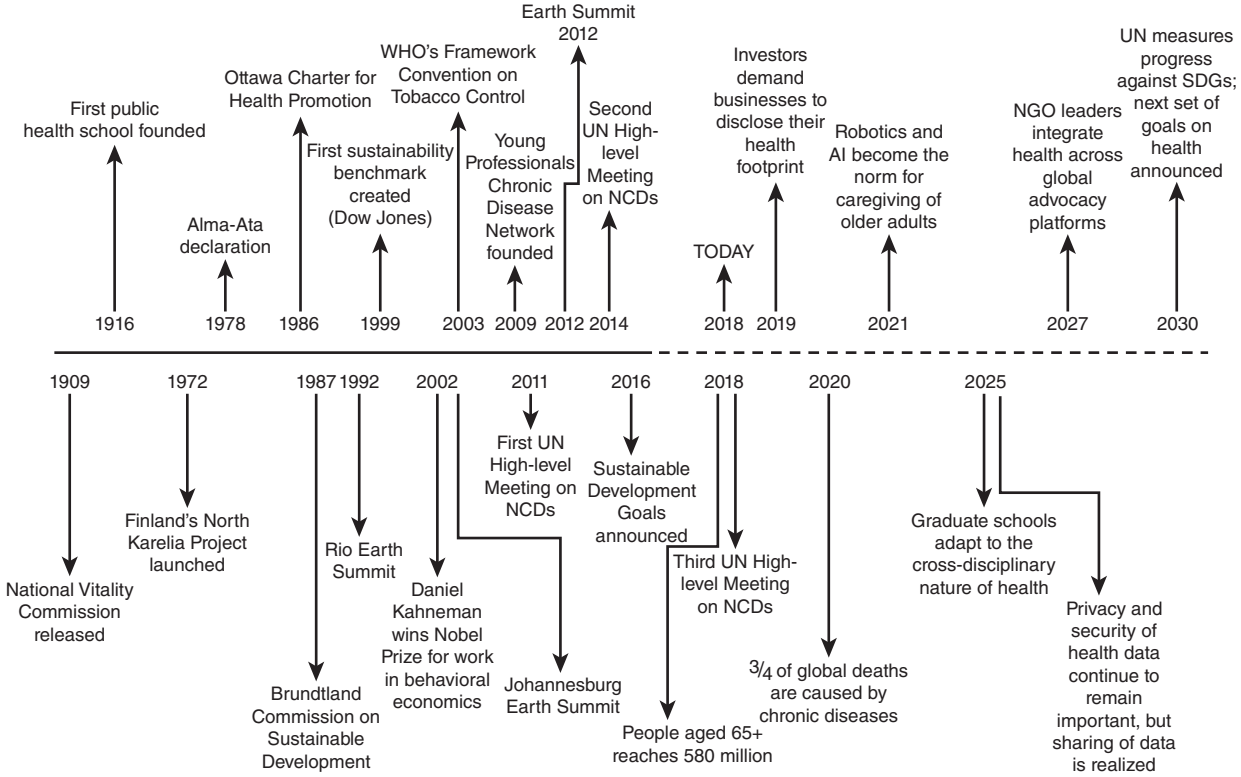


FIGURE 8-8 Milestones in chronic disease prevention, 1900–2030.

EXHIBIT 8-12 Future Roadmap for Health

To encourage interest in chronic diseases and to offer ideas on emerging trends, this list summarizes predictions about global trends related to health and well-being. It serves to inform actions to advance the Sustainable Development Goals (SDGs) and advocate for chronic disease prevention in coming decades.

- **Demographic changes:** Life expectancy now exceeds 70 years in many countries; rural workers are moving to cities to search for economic opportunities; and climate and political instability drives migration toward safer lands. These are intensifying pressures to develop healthy and vibrant communities.
- **Epidemiology structures:** Declines in undernutrition and maternal deaths have been complemented by rises in overweight/obesity and poor mental health. Infectious disease agents—most recently Ebola and Zika—will continue to threaten the resiliency of global populations. Strategies that encourage positive, individual behavioral change, and that prevent infectious disease outbreaks, will be critical to creating healthy populations and ensuring the sustainability of healthcare systems.
- **Health inequalities:** Health inequalities and inequities between different social groups are widening in many countries. Individuals' risk of poor health increases with declines in their socioeconomic position. Obesity rates, for example, are expected to rise in coming decades, especially among more disadvantaged groups.
- **Affordable health services:** Government austerity and rising healthcare costs threaten the sustainability of health systems. This has led to increased demand for affordable health services. Provisions for universal health coverage, investments into personalized treatments, and workplace health programs can reduce healthcare costs by governments and employers.
- **Digital technologies:** Rapid advances in technology are enabling individuals to quantify their health status, and can improve the well-being of disparate populations. These technologies empower individuals to better prevent, control, and treat disease, though associated privacy and security concerns must be proactively addressed to ensure their uptake among all populations.
- **Media and their consumers:** Media, including social media, serve to spark action and investment by documenting events with relevance to health, making consumers increasingly more vocal in their demands for creating healthier societies.

References

Abegunde, D. O., Shengelia, B., Luyten, A., Cameron, A., Celletti, F., Nishtar, S., & Mendis, S. (2007). Can non-physician health-care workers assess and manage cardiovascular risk in primary care? *Bulletin of the World Health Organization, 85*(6), 432–440.

Aboderin, I., Kalache, A., Ben-Shlomo, Y., Lynch, J. W., Yajnik, C. D., Kuh, D., & Yach, D. (2001). *Life course perspectives on coronary heart disease, stroke and diabetes: Key issues and implications for policy and research*. Geneva, Switzerland: World Health Organization.

Adams, J. L., Almond, M. L., Ringo, E. J., Shangali, W. H., & Sikkema, K. J. (2012). Feasibility of nurse-led antidepressant medication management of depression in an HIV clinic in Tanzania. *International Journal of Psychiatry in Medicine, 43*(2), 105–117.

Barr, V. J., Robinson, S., Marin-Link, B., Underhill, L., Dotts, A., Ravensdale, D., & Salivaras, S. (2003). The expanded Chronic Care Model: An integration of concepts and strategies from population health promotion and the Chronic Care Model. *Hospital Quarterly, 7*, 73–82.

Batty, G. D., & Leon, D. A. (2002). Socioeconomic position and coronary heart disease risk factors in children and young people. *European Journal of Public Health, 12*, 263–272.

Bell, A. C., Adair, A. S., & Popkin, B. M. (2004). Understanding the role of mediating risk factors and proxy effects in the association between socioeconomic status and untreated hypertension. *Social Science and Medicine, 59*, 275–283.

Black, R. E., Taylor, C. E., Arole, S., Bang, A., Bhutta, Z. A., Chowdhury, A. M. R., & Perry, H. B. (2017). Comprehensive review of the evidence regarding the effectiveness of community-based primary health care in improving maternal, neonatal and child health: 8. Summary and recommendations of the Expert Panel. *Journal of Global Health, 7*(1), 010908.

Bloom, D. E., Cafiero, E. T., Jané-Llopis, E., Abrahams-Gessel, S., Bloom, L. R., Fathima, S., . . . Weinstein, C. (2011). *The global economic burden of noncommunicable diseases*. Geneva, Switzerland: World Economic Forum.

Bodenheimer, T. S., & Smith, M. D. (2013). Primary care: Proposed solutions to the physician shortage without training more physicians. *Health Affairs (Millwood), 32*(11), 1881–1886.

Brands, A., & Yach, D. (2001). *Noncommunicable diseases and gender*. Geneva, Switzerland: World Health Organization.

Calitz, C., Pollack, K. M., Millard, C., & Yach, D. (2015). National Institutes of Health funding for behavioral interventions to prevent chronic diseases. *American Journal of Preventive Medicine, 48*(4), 462–471.

Callaghan, M., Ford, N., & Schneider, H. (2010). A systematic review of task-shifting for HIV treatment and care in Africa. *Human Resources for Health, 8*, 8.

Canto, J. G., & Iskandrian, A. E. (2003). Major risk factors for cardiovascular disease. Debunking the "only 50%" myth. *Journal of the American Medical Association, 290*, 947–949.

Cappuccio, F. P., Plange-Rhule, J., Phillips, R. O., & Eastwood, J. B. (2000). Prevention of hypertension and stroke in Africa. *The Lancet, 356*, 677–678.

Case, A., & Deaton, A. (2015). Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. *Proceedings of the National Academy of Sciences, 112*(49), 15078–15083.

Chamie, G., Kwarisiima, D., Clark, T. D., Kabami, J., Jain, V., Geng, E., & Charlebois, E. D. (2012). Leveraging rapid community-based HIV testing campaigns for non-communicable diseases in rural Uganda. *PLoS One, 7*(8), e43400.

Chibanda, D., Mesu, P., Kajawu, L., Cowan, F., Araya, R., & Abas, M. A. (2011). Problem-solving therapy for depression and common mental disorders in Zimbabwe: Piloting a task-shifting primary mental health care intervention in a population with a high prevalence of people living with HIV. *BMC Public Health, 11*, 828.

City Health Works. (2017). Home page. Retrieved from http://cityhealthworks.com/

Clark, G. L., & Hebb, T. (2002). *Understanding pension fund corporate engagement in a global arena*. Paper prepared for the seminar Understanding Pension Fund Corporate Engagement in a Global Arena, November 24–26, 2002, Oxford University, Oxford, UK.

Clinton, C., & Sridhar, D. (2017). Who pays for cooperation in global health? A comparative analysis of WHO, the World Bank, the Global Fund to fight HIV/AIDs, tuberculosis and malaria, and Gavi, the Vaccine Alliance. *The Lancet, 390*(10091), 324–332.

Coffield, A. B., Maciosek, M. V., McGinnis, J. M., Harris, J. R., Caldwell, M. B., Teutsch, S. M., . . . Haddix, A. (2001). Priorities among recommended clinical preventive services. *American Journal of Preventive Medicine, 21*(1), 1–9.

Coleman, R., Gill, G., & Wilkinson, D. (1998). Noncommunicable disease management in resource-poor settings: A primary care model from rural South Africa. *Bulletin of the World Health Organization, 76*(6), 633–640.

Collins, D. R. J., Jobanputra, K., Frost, T., Muhammed, S., Ward, A., Shafei, A. A., & Heneghan, C. (2017). Cardiovascular disease risk and prevention amongst Syrian refugees: Mixed methods study of Medecins Sans Frontieres programme in Jordan. *Conflict and Health, 11*, 14.

Commission on Social Determinants of Health. (2008). *Closing the gap in a generation: Health equity through action on the social determinants of health*. Geneva, Switzerland: World Health Organization.

Conradie, C. S., van der Merwe Smit, E., & Malan, D. P. (2016). Corporate health and wellness and the financial bottom line. *Journal of Occupational and Environmental Medicine, 58*(2), e45–e53.

Critchley, J., & Capewell, S. (2003). Mortality risk reduction associated with smoking cessation in patients with coronary heart diseases: A systematic review. *Journal of the American Medical Association, 290*, 86–97.

Dickinson, D., Wilkie, P., & Harris, M. (1999). Taking medicines: Concordance is not compliance. *British Medical Journal, 319*(7212), 787.

Dieleman, J. L., Baral, R., Birger, M., Bui, A. L., Bulchis, A., Chapin, A., & Murray, C. J. L. (2016). US spending on personal health care and public health, 1996–2013. *Journal of the American Medical Association, 316*(24), 2627–2646.

Duncan, B. D., Schmidt, M. I., Achutti, A. C., Polanczyk, C. A., Benia, L. R., & Maia, A. A. (1993). Socioeconomic distribution of noncommunicable disease risk factors in urban Brazil: The case of Pôrto Alegre. *Bulletin of the Pan American Health Organization, 27*(4), 337–349.

Ehrlich, R. I., White, N., Norman, R., Laubscher, R., Steyn, K., Lombard, C., & Bradshaw, D. (2004). Predictors of chronic bronchitis in South African adults. *International Journal of Tuberculosis and Lung Disease, 8*(3), 369–376.

Estabrooks, P. A., Glasgow, R. E., & Dzewaltowski, D. A. (2003). Physical activity promotion through primary care. *Journal of the American Medical Association, 289*(22), 2913.

Federspiel, F., Mukhopadhyay, S., Milsom, P., Scott, J. W., Riesel, J. N., & Meara, J. G. (2015). Global surgical and anaesthetic task shifting: A systematic literature review and survey. *The Lancet, 385*(suppl 2), S46.

Fernandez-Alvira, J. M., Fuster, V., Pocock, S., Sanz, J., Fernandez-Friera, L., Laclaustra, M., & Bueno, H. (2017). Predicting subclinical atherosclerosis in low-risk individuals: Ideal cardiovascular health score and Fuster-BEWAT score. *Journal of the American College of Cardiology, 70*(20), 2463–2473.

Flood, R. G., & Chiang, V. W. (2001). Rate and prediction of infection in children with diabetic ketoacidosis. *American Journal of Emergency Medicine, 19*, 270–273.

Frenk, J., Bobadilla, J. L., Sepúlveda, J., & Cervantes, M. L. (1989). Health transition in middle-income countries: New challenges for health care. *Health Policy and Planning, 4*(1), 29–39.

Fries, J. F., Bruce, B., & Chakravarty, E. (2011). Compression of morbidity 1980–2011: A focused review of paradigms and progress. *Journal of Aging Research, 2011*, 1–10.

Fu, D., Fu, H., McGowan, P., Shen, Y., Zhu, L., Yang, H., . . . Wei, Z. (2003). Implementation and quantitative evaluation of chronic disease self-management programme in Shanghai, China: Randomized controlled trial. *Bulletin of the World Health Organization, 81*(3), 174–182.

Gajalakshmi, C. K., Krishnamurthi, S., Ananth, R., & Shanta, V. (1996). Cervical cancer screening in Tamilnadu, India: A feasibility study of training the village health nurse. *Cancer Causes Control, 7*(5), 520–524.

Gajalakshmi, V., Peto, R., Kanaka, T. S., & Jha, P. (2003). Smoking and mortality from tuberculosis and other diseases in India. *The Lancet, 362*, 507–515.

Gaziano, T. A. (2008). Economic burden and the cost-effectiveness of treatment of cardiovascular diseases in Africa. *Heart, 94*(2), 140–144.

Gaziano, T. A., Abrahams-Gessel, S., Denman, C. A., Montano, C. M., Khanam, M., Puoane, T., & Levitt, N. S. (2015). An assessment of community health workers' ability to screen for cardiovascular disease risk with a simple, non-invasive risk assessment instrument in Bangladesh, Guatemala, Mexico, and South Africa: An observational study. *The Lancet Global Health, 3*(9), e556–e563.

Gaziano, T. A., & Pagidipati, N. (2013). Scaling up chronic disease prevention interventions in lower- and middle-income countries. *Annual Review of Public Health, 34*(1), 317–335.

Gaziano, T. A., Pandya, A., Steyn, K., Levitt, N., Mollentze, W., Joubert, G., & Laubscher, R. (2013). Comparative assessment of absolute cardiovascular disease risk characterization from non-laboratory-based risk assessment in South African populations. *BMC Medicine, 11*, 170.

Gaziano, T. A., Young, C. R., Fitzmaurice, G., Atwood, S., & Gaziano, J. M. (2008). Laboratory-based versus non-laboratory-based

method for assessment of cardiovascular disease risk: The NHANES I follow-up study cohort. *The Lancet, 371*(9616), 923–931.

Glasgow, R. E., Orleans, C. T., & Wagner, E. H. (2001). Does the chronic care model serve also as a template for improving prevention? *Milbank Quarterly, 79*(4), 579–612.

Global Burden of Disease (GBD) Collaborators. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388*(10053), 1439–1544.

GBD 2015 SDG Collaborators. (2016). Measuring the health-related Sustainable Development Goals in 188 countries: A baseline analysis from the Global Burden of Disease Study 2015. *The Lancet, 388*, 1813–1850.

Goetzel, R. Z., Fabius, R., Fabius, D., Roemer, E. C., Thornton, N., Kelly, R. K., & Pelletier, K. R. (2016). The stock performance of C. Everett Koop Award winners compared with the Standard & Poor's 500 index. *Journal of Occupational and Environmental Medicine, 58*(1), 9–15.

Goldin, J. G., Louw, S. J., & Joubert, G. (1996). Spirometry of healthy adult South African men Part I. Interrelationship between socio-environmental factors and "race" as determinant of spirometry. *South African Medical Journal, 86*(7), 820–826.

Goodyear-Smith, F., & Janes, R. (2008). New Zealand rural primary health care workforce in 2005: More than just a doctor shortage. *Australian Journal of Rural Health, 16*(1), 40–46.

Greenberg, H., Leeder, S. R., & Raymond, S. U. (2016). And Why So Great a "No?": The Donor and Academic Communities' Failure to Confront Global Chronic Disease. *Global Heart, 11*(4), 381–385.

Grimshaw, J. M., Shirran, L., Thomas, R., Mowatt, G., Fraser, C., Bero, L., . . . O'Brian, M. A. (2001). Changing provider behavior: An overview of systematic reviews of interventions. *Medical Care, 39*, 112–145.

Haldane, V., Legido-Quigley, H., Chuah, F. L. H., Sigfrid, L., Murphy, G., Ong, S. E., & Perel, P. (2018). Integrating cardiovascular diseases, hypertension, and diabetes with HIV services: A systematic review. *AIDS Care, 30*(1), 103–115.

Heller, D. J., & Kishore, S. P. (2017). Closing the blood pressure gap: An affordable proposal to save lives worldwide. *British Medical Journal Global Health, 2*(3), 1–4.

Heller, P. S. (2003). *Who will pay? Coping with aging societies, climate change, and other long-term fiscal challenges*. Washington, DC: International Monetary Fund.

Heller, R. (1978). Officiers de sante: The second-class doctors of nineteenth-century France. *Medical History, 22*(1), 25–43.

Holmes, P. (2016). USAID: Standing by on NCD. *Global Heart, 11*(4), 425–427.

Horton, R. (2015). Offline: Chronic diseases: The social justice issue of our time. *The Lancet, 386*(10011), 2378.

Horton, R. (2016). Offline: The reflection of ourselves we choose to ignore. *The Lancet, 387*(10018), 520.

Ilesanmi, O. S., Ige, O. K., & Adebiyi, A. O. (2012). The managed hypertensive: The costs of blood pressure control in a Nigerian town. *Pan African Medical Journal, 12*(1).

Institute of Medicine. (2010). *Preventing the global epidemic of cardiovascular disease: Meeting the challenges in developing countries*. Washington, DC: National Academies Press.

International Diabetes Federation. (2003). *Diabetes atlas* (2nd ed.). Brussels, Belgium: Author.

James, P. A., Oparil, S., Carter, B. L., Cushman, W. C., Dennison-Himmelfarb, C., Handler, J., & Ortiz, E. (2014). 2014 evidence-based guideline for the management of high blood pressure in adults: Report from the panel members appointed to the Eighth Joint National Committee (JNC 8). *Journal of the American Medical Association, 311*(5), 507–520.

Janssens, B., Van Damme, W., Raleigh, B., Gupta, J., Khem, S., Soy Ty, K., & Zachariah, R. (2007). Offering integrated care for HIV/AIDS, diabetes and hypertension within chronic disease clinics in Cambodia. *Bulletin of the World Health Organization, 85*(11), 880–885.

Jha, P., & Chaloupka, F. J. (Eds.). (2000). *Tobacco control in developing countries*. New York, NY: Oxford University Press.

Joshi, R., Alim, M., Kengne, A. P., Jan, S., Maulik, P. K., Peiris, D., & Patel, A. A. (2014). Task shifting for non-communicable disease management in low and middle income countries: A systematic review. *PLoS One, 9*(8), e103754.

Joshi, R., Chow, C. K., Raju, P. K., Raju, K. R., Gottumukkala, A. K., Reddy, K. S., & Neal, B. (2012). The Rural Andhra Pradesh Cardiovascular Prevention Study (RAPCAPS): A cluster randomized trial. *Journal of American College of Cardiology, 59*(13), 1188.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kar, S. S., Thakur, J. S., Jain, S., & Kumar, R. (2008). Cardiovascular disease risk management in a primary health care setting of north India. *Indian Heart Journal, 60*(1), 19–25.

Kengne, A. P., Awah, P. K., Fezeu, L. L., Sobngwi, E., & Mbanya, J. C. (2009). Primary health care for hypertension by nurses in rural and urban sub-Saharan Africa. *Journal of Clinical Hypertension, 11*(10), 564–572.

Kengne, A. P., Fezeu, L., Sobngwi, E., Awah, P. K., Aspray, T. J., Unwin, N. C., & Mbanya, J. C. (2009). Type 2 diabetes management in nurse-led primary healthcare settings in urban and rural Cameroon. *Primary Care Diabetes Journal, 3*(3), 181–188.

Kengne, A. P., Sobngwi, E., Fezeu, L. L., Awah, P. K., Dongmo, S., & Mbanya, J. C. (2008). Nurse-led care for asthma at primary level in rural sub-Saharan Africa: The experience of Bafut in Cameroon. *Journal of Asthma, 45*(6), 437–443.

Khetan, A. K., Purushothaman, R., Chami, T., Hejjaji, V., Mohan, S. K. M., Josephson, R. A., & Webel, A. R. (2017). The effectiveness of community health workers in CVD prevention in LMIC. *Global Heart, 12*(3), 233–243.

Kim, S., Symons, M., & Popkin, B. M. (2004). Contrasting socioeconomic profiles related to healthier lifestyles in China and the United States. *American Journal of Epidemiology, 159*(2), 184–191.

Kimmis, J., Gottchalk, R., Armendariz, E., & Griffith-Jones, S. (2002). *Making the case for UK pension fund investment in developing country assets*. Unpublished manuscript, Institute for Development Studies, University of Sussex, Brighton, UK.

Kishore, S. P., Heller. D. J., & Vasan, A. (2018). Beyond hypertension: Integrated cardiovascular care as a path to comprehensive primary care. *Bulletin of the World Health Organization, 96*, 219–221.

Knuiman, J. T., Hermus, R. J. J., & Hautvast, J. G. A. J. (1980). Serum total and high-density lipoprotein (HDL) cholesterol concentrations in rural and urban boys from 16 countries. *Atherosclerosis, 36*, 529–537.

Kogevinas, M., Pearce, N., Susser, M., & Boffetta, P. (Eds.). (1997). *Social inequalities and cancer* (IARC Scientific Publications No. 138). Lyon, France: IARC.

Kotwani, P., Balzer, L., Kwarisiima, D., Clark, T. D., Kabami, J., Byonanebye, D., & Charlebois, E. D. (2014). Evaluating linkage to care for hypertension after community-based screening in

rural Uganda. *Tropical Medicine and International Health, 19*(4), 459–468.

Kwarisiima, D., Balzer, L., Heller, D., Kotwani, P., Chamie, G., Clark, T., & Kamya, M. R. (2016). Population-based assessment of hypertension epidemiology and risk factors among HIV-positive and general populations in rural Uganda. *PLoS One, 11*(5), e0156309.

Labhardt, N. D., Balo, J. R., Ndam, M., Grimm, J. J., & Manga, E. (2010). Task shifting to non-physician clinicians for integrated management of hypertension and diabetes in rural Cameroon: A programme assessment at two years. *BMC Health Services Research, 10*, 339.

Lamptey, P., Laar, A., Adler, A. J., Dirks, R., Caldwell, A., Prieto-Merino, D., . . . Perel, P. (2017). Evaluation of a community-based hypertension improvement program (ComHIP) in Ghana: Data from a baseline survey. *BMC Public Health, 17* (1), 368.

Lasser, K., Boyd, J. W., Woolhandler, S., Himmelstein, D. U., McCormick, D., & Bor, D. H. (2000). Smoking and mental health: A population-based prevalence study. *Journal of the American Medical Association, 284*, 2606–2610.

Leeder, S., Raymond, S., Greenberg, H., Liu, H., & Esson, K. (2004). *A race against time: The challenge of cardiovascular disease in developing countries*. New York, NY: Trustees of Columbia University.

Levitt, N. S., Steyn, K., de Wet, T., Morrell, C., Edwards, R., Ellison, G. T. H., & Cameron, N. (1999). An inverse relationship between blood pressure and birth weight among 5 year old children from Soweto, South Africa. *Journal of Epidemiology and Community Health, 53*, 264–268.

LMC International. (2004, May). *Implications of the WHO/FAO expert consultative report on the sugar market*. Unpublished manuscript, LMC International, London.

Luscher, T. F., Vetter, H., Sigenthaler, W., & Vetter, W. (1985). Compliance in hypertension: Facts and concepts. *Journal of Hypertension, 3*, 3–9.

Mackenbach, J. P., Cavelaars, A. E. J. M., Kunst, A. E., & Groenhof, F. (2000). Socio-economic inequalities in cardiovascular disease mortality: An international study. *European Heart Journal, 21*, 1141–1151.

Mancia, G., Fagard, R., Narkiewicz, K., Redon, J., Zanchetti, A., Bohm, M., . . . Zannad, F. (2013). 2013 ESH/ESC guidelines for the management of arterial hypertension: The Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). *Journal of Hypertension, 31*(7), 1281–1357.

Mann, V., Eble, A., Frost, C., Premkumar, R., & Boone, P. (2010). Retrospective comparative evaluation of the lasting impact of a community-based primary health care programme on under-5 mortality in villages around Jamkhed, India. *Bulletin of the World Health Organization, 88*(10), 727–736.

Marmot, M. G., Adelstein, A. M., Robinson, N., & Rose, G. A. (1978). Changing social-class distribution of heart disease. *British Medical Journal, 2*(6145), 1109–1112.

Mayosi, B. M., Flisher, A. J., Lalloo, U. G., Sitas, F., Tollman, S. M., & Bradshaw, D. (2009). The burden of non-communicable diseases in South Africa. *The Lancet, 374*(9693), 934–947.

Mendez, M. A, Cooper, R. S., Luke, A., Wilks, R., Bennett, F., & Forrester, T. (2004). Higher income is more strongly associated with obesity than with obesity-related metabolic disorders in Jamaican adults. *International Journal of Obesity, 28*(4), 543–550.

Mendis, S. (2016). Combating chronic diseases: The role of the World Health Organization. *Global Heart, 11*(4), 413–415.

Mendis, S., Johnston, S. C., Fan, W., Oladapo, O., Cameron, A., & Faramawi, M. F. (2010). Cardiovascular risk management and its impact on hypertension control in primary care in low-resource settings: A cluster-randomized trial. *Bulletin of the World Health Organization, 88*(6), 412–419.

Mills, K. T., Bundy, J. D., Kelly, T. N., Reed, J. E., Kearney, P. M., Reynolds, K., & He, J. (2016). Global disparities of hypertension prevalence and control. *Circulation, 134*(6), 441–450.

Mills, K. T., Xu, Y., Zhang, W., Bundy, J. D., Chen, C. S., Kelly, T. N., . . . & He, J. (2015). A systematic analysis of worldwide population-based data on the global burden of chronic kidney disease in 2010. *Kidney International, 88*(5), 950–957.

Misra, A., Pandey, R. M., Devi, J. R., Sharma, R., Vikram, N. K., & Khanna, N. (2001). High prevalence of diabetes, obesity and dyslipidaemia in urban slum populations in northern India. *International Journal of Obesity, 25*, 1–8.

Mnookin, S. (2016). Out of the shadows: Making mental health a global development priority. Retrieved from http://documents.worldbank.org/curated/en/270131468187759113/Out-of-the-shadows-making-mental-health-a-global-development-priority

Monteiro, C. A., Conde, W. L., Lu, B., & Popkin, B. M. (2004, June 22). Obesity and health inequities in the developing world. *International Journal of Obesity,* 1–6. [Epub ahead of print].

Morisky, D. E., Levine, D. M., Green, L. W., Shapiro, S., Russell, R. P., & Smith, C. R. (1983). Five year blood pressure control and mortality following health education for hypertensive patients. *American Journal of Public Health, 73*, 153–162.

Mutamba, B. B., van Ginneken, N., Paintain, L. S., Wandiembe, S., & Schellenberg, D. (2013). Roles and effectiveness of lay community health workers in the prevention of mental, neurological and substance use disorders in low and middle income countries: A systematic review. *BMC health services research, 13*(1), 412.

Narayan, K. M., Miotti, P. G., Anand, N. P., Kline, L. M., Harmston, C., Gulakowski, R., 3rd, & Vermund, S. H. (2014). HIV and noncommunicable disease comorbidities in the era of antiretroviral therapy: A vital agenda for research in low- and middle-income country settings. *Journal of Acquired Immune Deficiency Syndromes, 67*(suppl 1), S2–S7.

National Academies of Sciences, Engineering, and Medicine. (2017). *Promoting cardiovascular health and preventing cancer. National Academies of Sciences, Engineering, and Medicine report on global health and the future role of the United States*. Washington, DC: National Academies Press.

National Institute on Aging. (2011). Global aging. Retrieved from https://www.nia.nih.gov/research/publication/global-health-and-aging/living-longer

Nature. (2016). Urogenital diseases. Retrieved from http://www.nature.com/subjects/urogenital-diseases

NCD Alliance. (2016). About the NCD Alliance. Retrieved from https://ncdalliance.org/who-we-are/about-ncd-alliance

Nissinen, A., Bothig, S., Granroth, H., & Lopez, A. D. (1988). Hypertension in developing countries. *World Health Statistics Quarterly, 41*, 141–154.

Noel, P. H., & Pugh, J. A. (2002). Management of overweight and obese adults. *British Medical Journal, 325*(7367), 757–761.

Nugent, R. (2016). A chronology of global assistance funding for NCD. *Global Heart, 11*(4), 371–374.

Ogedegbe, G., Plange-Rhule, J., Gyamfi, J., Chaplin, W., Ntim, M., Apusiga, K., & Cooper, R. (2014). A cluster-randomized trial

of task shifting and blood pressure control in Ghana: Study protocol. *Implementation Science, 9*, 73.

Omada Health. (2016). Medicare will cover diabetes prevention program for at-risk seniors. Retrieved from https://www.omadahealth.com/news/medicare-will-cover-diabetes-prevention-program-for-at-risk-seniors

Oziransky, V., Yach, D., Tsao, T. Y., Luterek, A., & Stevens, D. (2015). *Beyond the four walls: Why community is critical to workforce health*. New York, NY: Vitality Institute.

Partners in Health, Bukhman, G., & Kidder, A. (2011). *The PIH guide to chronic care integration for endemic non-communicable diseases*. Boston, MA: Partners in Health.

Partners in Health, Bukhman, G., & Kidder, A. (2013). *The PIH guide to chronic care integration for endemic non-communicable diseases*. Boston, MA: Partners in Health.

Patel, V., Weiss, H. A., Chowdhary, N., Naik, S., Pednekar, S., Chatterjee, S., & Kirkwood, B. R. (2010). Effectiveness of an intervention led by lay health counsellors for depressive and anxiety disorders in primary care in Goa, India (MANAS): A cluster randomised controlled trial. *The Lancet, 376*(9758), 2086–2095.

Peña, M., & Bacallao, J. (2000). Obesity among the poor: An emerging problem in Latin America and the Caribbean. In M. Peña & J. Bacallao (Eds.), *Obesity and poverty: A new public health challenge* (Scientific Publication No. 576, pp. 3–10). Washington, DC: Pan American Health Organization.

Phillips, J. F. (2016–2018). Adapting the WHO cardiovascular disease risk management package to the Ghanaian Community-Based Health Planning and Services (CHPS) primary care model: An implementation pilot study. Retrieved from https://www.fic.nih.gov/Grants/Search/Pages/ncd-research-TW010452.aspx

Pomeranz, J. L. (2016). Valuing federal taxation policies to prevent disease and raise revenue. *American Journal of Preventive Medicine, 51*(4), 518–521.

Porter, M. E., & Kramer, M. R. (2011). Creating shared value. *Harvard Business Review, 89*(1/2), 62–77.

Poulter, N., Khaw, K. T., Hopwood, B. E., Mugambi, M., Peart, W. S., Rose, G., & Sever, P. S. (1990). The Kenyan Luo migration study: Observations on the initiation of a rise in blood pressure. *British Medical Journal, 300*, 967–972.

Psaty, B. M., Koepsell, T. D., Wagner, E. H., Lo Gerfo, J. P., & Inui, T. S. (1990). The relative risk of incident coronary heart diseases associated with recently stopping the use of β-blockers. *Journal of the American Medical Association, 73*, 1653–1657.

Puska, P. (2016). Why did North Karelia–Finland work?: Is it transferrable? *Global Heart, 11*(4), 387–391.

Rabkin, M., Kruk, M. E., & El-Sadr, W. M. (2012). HIV, aging and continuity care: Strengthening health systems to support services for noncommunicable diseases in low-income countries. *AIDS, 26*(suppl 1), S77–S83.

Ralston, J., Reddy, K. S., Fuster, V., & Narula, J. (2016). Cardiovascular diseases on the global agenda: The United Nations high level meeting, sustainable development goals, and the way forward. *Global Heart, 11*(4), 375–379.

Ramachandran, A., Snehalatha, C., Vijay, V., & King, H. (2002). Impact of poverty on the prevalence of diabetes and its complications in urban southern India. *Diabetic Medicine, 19*, 130–135.

Renders, C. M., Valk, G. D., Griffin, S. M., Wagner, E. H., Eijk, J. T., & Assendelft, W. J. (2002). Interventions to improve the management of diabetes in primary care, outpatient and community settings: A systematic review. *Diabetes Care, 24*, 1821–1833.

Salomon, J. A., Wang, H., Freeman, M. K., Vos, T., Flaxman, A. D., Lopez, A. D., & Murray, C. J. (2013). Healthy life expectancy for 187 countries, 1990–2010: A systematic analysis for the Global Burden Disease Study 2010. *The Lancet, 380*(9859), 2144–2162.

Seftel, H., Asvat, M. S., Joffe, B. I., Raal, F. J., Panz, V. R., & Vermaak, W. J. H. (1993). Selected risk factors for coronary heart disease in male scholars from the major South African population groups. *South African Medical Journal, 83*, 891–897.

Serdula, M. K., Khan, L. K., & Dietz, W. H. (2003). Weight loss counseling revisited. *Journal of the American Medical Association, 289*(14), 1747–1750.

Siddiqi, K., Khan, A., Ahmad, M., Dogar, O., Kanaan, M., Newell, J. N., & Thomson, H. (2013). Action to Stop Smoking in Suspected Tuberculosis (ASSIST) in Pakistan: A cluster randomized, controlled trial. *Annals of Internal Medicine, 158*(9), 667–675.

Sidel, V. W. (1972). The barefoot doctors of the People's Republic of China. *New England Journal of Medicine, 286*(24), 1292–1300.

Siegel, K. R., Ali, M. K., Srinivasiah, A., Nugent, R. A., & Narayan, K. M. V. (2014). Do we produce enough fruits and vegetables to meet global health need? *PLoS One, 9*(8), e104059.

Singh, R. B., Beegom, R., Mehta, A. S., Niaz, M. A., De, A. K., Haque, M., . . . Haque, M. S. (1999). Social class, coronary risk factors and undernutrition: A double burden of diseases, in women during transition, in five Indian cities. *International Journal of Cardiology, 69*(2), 139–147.

Singhal, A., & Lucas, A. (2004). Early origins of cardiovascular disease: Is there a unifying hypothesis? *The Lancet, 363*, 1642–1645.

Sobal, J., & Stunkard, A. J. (1989). Socioeconomic status and obesity: A review of the literature. *Psychological Bulletin, 105*(2), 260–275.

Steptoe, A., Perkins-Porras, L., McKay, C., Rink, E., Hilton, S., & Cappuccio, F. P. (2003). Behavioural counselling to increase consumption of fruit and vegetables in low income adults: Randomised trial. *British Medical Journal, 326*(7394), 855.

Stewart, B. W., & Kleihaus, P. (Eds.). (2003). *World cancer report*. Lyon, France: IARC Press.

Steyn, K., de Wet, T., Richter, L., Cameron, N., Levitt, N. S., & Morrell, C. (2000). Cardiovascular disease risk factors in five-year-old urban South African children: The Birth to Ten study. *South African Medical Journal, 90*(7), 719–726.

Stuckler, D. (2011). *Sick societies: Responding to the global challenge of chronic disease*. Oxford, UK: Oxford University Press.

Sustainable East Africa Research in Community Health (SEARCH). (n.d.). Retrieved from https://clinicaltrials.gov/ct2/show/NCT01864603

Swai, A. B., McLarty, D. G., Kitange, H. M., Kilima, P. M., Tatalla, S., Keen, N., . . . Alberti, K. G. (1993). Low prevalence of risk factors for coronary heart disease in rural Tanzania. *International Journal of Epidemiology, 22*(4), 651–659.

Swai, E. S., French, N. P., Beauchamp, G., Fitzpatrick, J. L., Bryant, M. J., Kambarage, D., & Ogden, N. H. (2005). A longitudinal study of sero-conversion to tick-borne pathogens in smallholder dairy youngstock in Tanzania. *Veterinary Parasitology, 131*(1–2), 129–137.

The 3 by 5 initiative. (n.d.). Integrated Management of Adolescent and Adult Illness (IMAI) modules. Retrieved from http://www.who.int/3by5/publications/documents/imai/en/

The Lancet NCDI Poverty Commission. (2017). *The Lancet*. Retrieved from http://www.ncdipoverty.org

United Nations. (2016). Sustainable Development Goals. Retrieved from http://www.un.org/sustainabledevelopment/sustainable-development-goals

U.S. Department of Health and Human Services (DHHS), National Institutes of Health, & National Institute of Diabetes and Digestive and Kidney Diseases. (2008). Diabetes Prevention Program (DPP). Retrieved from https://www.niddk.nih.gov/about-niddk/research-areas/diabetes/diabetes-prevention-program-dpp

U.S. Department of Health and Human Services (DHHS), Office of the Surgeon General. (2016). *Facing addiction in America: The Surgeon General's report on alcohol, drugs, and health*. Washington, DC: National Institutes of Health.

Vedanthan, R., Bernabe-Ortiz, A., Herasme, O. I., Joshi, R., Lopez-Jaramillo, P., Thrift, A. G., & Fuster, V. (2017). Innovative approaches to hypertension control in low- and middle-income countries. *Cardiology Clinics, 35*(1), 99–115.

Vedanthan, R., Kamano, J. H., Naanyu, V., Delong, A. K., Were, M. C., Finkelstein, E. A., & Fuster, V. (2014). Optimizing linkage and retention to hypertension care in rural Kenya (LARK hypertension study): Study protocol for a randomized controlled trial. *Trials, 15*, 143.

Vikram, V. M., Pandey, R. M., Misra, A., Sharma, R., Devi, J. R., & Khanna, N. (2003). Non-obese (body mass index, 25 kg/m^2) Asian Indians with normal waist circumference have high cardiovascular risk. *Nutrition, 19*(6), 503–509.

Vitality Institute. (2014). Investing in prevention: A national imperative. Retrieved from http://thevitalityinstitute.org/commission

Vitality Institute. (2016). *Vitality screening and vaccination recommendations*. New York, NY: Health Assessment, Monitoring, and Intervention Board.

von dem Knesebeck O. (2010, May 18). Health inequalities in ageing societies. *International Journal of Public Health, 55*(6), 523–524.

von Schirnding, Y. (2002). Health and sustainable development: Can we rise to the challenge? *The Lancet, 360*(9333), 632–637.

Wagner, E. H., Glasgow, R. E., Davis, C., Bonomie, A. E., Provost, L., McCulloch, D., & Sixta, C. (2001). Quality improvement in chronic illness care: A collaborative approach. *Joint Commission Journal on Quality Improvement, 27*, 63–80.

Warman, A. (2001). Living the revolution: Cuban health workers. *Journal of Clinical Nursing, 10*(3), 311–319.

Warnakulasuriya, K. A., Ekanayake, A. N., Sivayoham, S., Stjernsward, J., Pindborg, J. J., Sobin, L. H., & Perera, K. S. (1984). Utilization of primary health care workers for early detection of oral cancer and precancer cases in Sri Lanka. *Bulletin of the World Health Organization, 62*(2), 243–250.

Waxman, A. (2004). The WHO global strategy on diet, physical activity and health: The controversy on sugar. *Development, 47*(2), 75–83.

Waxman, H. A. (2004). Politics of international health in the Bush administration. *Development, 47*(2), 24–30.

Women and Gender Equity Knowledge Network (WGEKN). (2007). *Unequal, unfair, ineffective and inefficient gender inequity in health: Why it exists and how we can change it. Final report of the Women and Gender Equity Knowledge Network of the Commission on Social Determinants of Health*. Geneva, Switzerland: World Health Organization.

Wong, M. D., Shapiro, M. F., Boscardin, W. J., & Ettner, S. L. (2002). Contribution of major diseases to disparities in mortality. *New England Journal of Medicine, 347*(20), 1585–1592.

World Commission on Environment and Development. (1987). *Our common future*. Oxford, UK: Oxford University Press. Retrieved from http://www.un-documents.net/wced-ocf.htm

World Health Organization (WHO). (1986). Ottawa charter for health promotion. Retrieved from http://www.who.int/healthpromotion/conferences/previous/ottawa/en

World Health Organization (WHO). (2002a). *Innovative care for chronic conditions: Building blocks for action*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2002b). *WHO report on the consultative meeting on primary health care policy review in the Eastern Mediterranean Region, Muscat, Oman, 28–30 January 2002*. Cairo, Egypt: Author.

World Health Organization (WHO). (2003a). *Investing in mental health*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2003b). WHO Framework Convention on Tobacco Control. Retrieved from http://www.who.int/fctc/text_download/en

World Health Organization (WHO). (2004). General principles of good chronic care. Retrieved from http://www.who.int/hiv/pub/imai/generalprinciples082004.pdf?ua=1

World Health Organization (WHO). (2006a). Neurological disorders: Public health challenges. Retrieved from http://www.who.int/mental_health/neurology/neurodiso/en

World Health Organization (WHO). (2006b). The world health report 2006: Working together for health. Retrieved from http://www.who.int/whr/2006/en/

World Health Organization (WHO). (2007). Prevention of cardiovascular disease. Retrieved from http://www.who.int/cardiovascular_diseases/publications/Prevention_of_Cardiovascular_Disease/en/

World Health Organization (WHO). (2008a). The global burden of disease: 2004 update. Retrieved from http://www.who.int/healthinfo/global_burden_disease/2004_report_update/en/

World Health Organization (WHO). (2008b). Task shifting: Rational redistribution of tasks among health workforce teams: Global recommendations and guidelines. First Global Conference on Task Shifting. Retrieved from http://www.who.int/healthsystems/task_shifting/en/

World Health Organization (WHO). (2014). Mental health: A state of well-being. Retrieved from http://www.who.int/features/factfiles/mental_health/en

World Health Organization (WHO). (2015a). 10 facts on ageing and the life course. Retrieved from http://www.who.int/features/factfiles/ageing/ageing_facts/en

World Health Organization (WHO). (2015b). World report on ageing and health. Retrieved from http://apps.who.int/iris/bitstream/10665/186463/1/9789240694811_eng.pdf?ua=1

World Health Organization (WHO). (2016a). Cancer. Retrieved from http://www.who.int/cancer/en

World Health Organization (WHO). (2016b). Cardiovascular diseases (CVDs). Retrieved from http://www.who.int/mediacentre/factsheets/fs317/en

World Health Organization (WHO). (2016c). Chronic respiratory diseases. Retrieved from http://www.who.int/respiratory/en

World Health Organization (WHO). (2016d). Diabetes. Retrieved from http://www.who.int/topics/diabetes_mellitus/en

World Health Organization (WHO). (2016e). *WHO's engagement with non-state actors*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2017a). HEARTS technical package for cardiovascular disease management in primary health care. Retrieved from http://www.who.int/cardiovascular_diseases/hearts/en/

World Health Organization (WHO). (2017b). Noncommunicable diseases. Retrieved from http://www.who.int/mediacentre/factsheets/fs355/en

World Health Organization (WHO). (2017c). Tobacco. Retrieved from http://www.who.int/mediacentre/factsheets/fs339/en

World Health Organization (WHO) & Centers for Disease Control and Prevention (CDC). (2017). Global Hearts Initiatives: Working together to beat cardiovascular diseases. Retrieved from http://www.who.int/cardiovascular_diseases/global-hearts/en

Yach, D., Khan, M., Bradley, D., Hargrove, R., Kehoe, S., & Mensah, G. (2010). The role and challenges of the food industry in addressing chronic disease. *Globalization and Health, 6*, 10.

Yach, D., & Raviglione, M. (2004). TB and tobacco. *Evidence -Based Health Care, 8*, 28.

Young Professionals Chronic Disease Network. (2016). Essential medicines. Retrieved from http://www.ncdaction.org/essential_medicines

CHAPTER 9

Unintentional Injuries and Violence

Margaret Peden, Berit Kieselbach, Alexander Butchart, and Adnan A. Hyder

Introduction

Injuries are no "accident" and violence does not "just happen." Recognition that a scientific approach to the prevention and control of injuries and violence can be effective has significantly raised these conditions' profile on the public health agenda. Twenty-plus years ago, injuries and violence hardly merited mention in the curricula of most schools of public health, whereas today their prevention and control is an integral part of a well-balanced training program in public health in most high-income countries and in increasing numbers of low- and middle-income countries (LMICs).

Injuries have traditionally been defined as physical damage to a person caused by an acute transfer of energy (mechanical, thermal, electrical, chemical, or radiation energy) or by the sudden absence of heat or oxygen. It is now recognized, however, that this definition is too narrow, but rather should be broadened to include impacts that result in psychological harm, maldevelopment, or deprivation (Krug, Dahlberg, Mercy, Zwi, & Lozano, 2002).

Injuries can be grouped in various ways, but most commonly are categorized with reference to the presumed underlying intent—that is, as injuries of unintentional, intentional, and undetermined intent. Unintentional injuries are those where there is no evidence of predetermined intent—sometimes referred to as "accidents"—which erroneously gives the impression they are not amenable to prevention. They include injuries sustained as a result of road traffic crashes, poisonings, falls, burns, and drowning, as well as occupational and sports injuries.

Intentional injuries result from violence, defined as "the intentional use of physical force or power, threatened or actual, against oneself, another person, or against a group or community, that either results in or has a high likelihood of resulting in injury, death, psychological harm, maldevelopment or deprivation" (Krug et al., 2002). In addition to intentional injuries, victims of violence are at increased risk of a wide range of psychological and behavioral problems, including depression, alcohol abuse, smoking, anxiety, and suicidal behavior, as well as reproductive health problems such as unwanted pregnancy, sexually transmitted diseases, and sexual dysfunction (Krug et al., 2002). Violence can be categorized as self-directed violence (violence that a person inflicts upon himself or herself), interpersonal violence (violence inflicted by another individual or a small group of individuals), or collective violence (violence inflicted by larger groups such as states, organized political groups, religious groups, militia groups,

and terrorist organizations). *Self-directed* violence includes completed suicides, attempted suicides, suicidal ideation, suicidal behaviors, and self-harm. *Interpersonal* violence includes child maltreatment, intimate-partner violence, and elder abuse—all forms that occur largely between family members—as well as youth violence and sexual violence, which occur largely in the community between people who are unrelated and may or may not be acquainted. *Collective* violence can be subdivided into social, political, and economic violence.

It is not always clear whether an injury has occurred from violence or an unintended cause. In such cases, the injury is classified as being of *undetermined* intent, although for legal or social reasons, it may well be classified as an unintentional injury.

This chapter considers unintentional injuries and violence, including issues related to a number of cause-specific injuries. The latter include those injuries that are routinely analyzed and for which statistics are published by the World Health Organization (WHO); individually, these injuries account for the greatest injury burden in terms of mortality and disability-adjusted life-years (DALYs). They include road traffic injuries, poisonings, falls, burns, and drowning, as well as self-directed violence, interpersonal violence (including child maltreatment, youth violence, intimate-partner violence, sexual violence, and elder abuse), and collective violence, including war.

The first section provides an overview of the global burden of injuries and violence. It includes discussion of data sources and the challenges of obtaining accurate data on these types of events, especially in LMICs. The next section outlines known and potential risk factors for unintentional injuries and violence. Included here is information on the extent to which knowledge about risk factors in one setting can be applied to other settings, and especially the extent to which knowledge about risk factors can be transferred from high-income settings to low- and middle-income settings.

Current evidence about effective interventions to prevent and control injuries and violence is then summarized. Again, the extent to which evidence of effectiveness from high-income countries can be extrapolated to LMICs is considered.

The last section considers the opportunities and challenges that exist in advancing the injury and violence prevention agenda. Issues addressed include the ongoing need for advocacy for prevention, the essential role of research, the significance of a workforce trained to deal with these issues, and the role of national and international organizations in promoting the injury and violence prevention agenda, including international societies.

The Global Burden of Unintentional Injuries and Violence

Data Sources

Data on unintentional injuries and violence may be obtained from sources both within and outside the health sector. Health sector data sources include the usual data sources for other health conditions, such as health information systems, vital registration systems, and hospital discharge data, in addition to sources that are more specific to injuries and violence, such as ambulance data and trauma registries. Sources outside the health sector cover a wide spectrum of resources, depending on the type and nature of the injuries or violence. Commonly used non-health-sector data sources include police data, transport-sector data, legal records, and insurance company claims. Innovative data sources for injuries have also been used, such as newspapers and consumer reports. For both unintentional injuries and violence, population-based surveys are an essential complement to service-based sources of information, especially in settings where access to services such as health care and the police is limited, and where victims may be reluctant or unable to seek help, such as in cases of child maltreatment, intimate-partner violence, and elder abuse. This diversity of data sources makes the field of injuries and violence both unique and challenging—unique, in terms of the intersectoral nature of the information, and challenging, because the biases and nature of data from each source need to be understood.

Comprehensive global data on unintentional injuries and violence are available from the WHO as part of that organization's Global Health Estimates (GHE).[1] Attempts to obtain consistent and internationally comparable data have been made by WHO since 1990. WHO data do have some limitations: Notably, the information relies on death certificate data submitted to WHO from governments around the world. Also, in specific instances, it may not cover the entire country or might be incomplete;

[1] http://www.who.int/gho/mortality_burden_disease/causes_death/en/

for example, some countries do not classify all subcategories of unintentional injuries and violence. In addition, data on burns include only fire burns (and not scalds), while data on drownings do not include drownings due to floods. Nevertheless, its aggregation of all unintentional injuries and violence into one unified system of information makes the WHO database most useful for public health purposes, as it is also publicly available.

The Global Burden of Disease (GBD) database from the Institute for Health Metrics and Evaluation is another valuable source of data. These data are collected by a consortium of researchers from more than 130 countries on premature death and disability from more than 300 diseases and injuries. The data are updated regularly, and each update revises all previous estimates for all years (GBD 2015 DALYs and HALE Collaborators, 2016).

For specific types of injury, alternative data sources are available, and they frequently provide estimates that differ from those provided by WHO. The most comprehensive global road safety database is regularly compiled by WHO into a series of global status reports on road safety with its associated database published in the Global Health Observatory.[2] Examples of regional databases include the International Road and Traffic Accident Database, which collates data from 32 Organisation for Economic Co-operation and Development (OECD) member countries, and OISEVI, which collates data from 20 Latin American countries.[3] The road traffic death estimates from various databases may vary depending on the source of data—for example, vital registration (death certificates) versus police report forms, or whether these data are reported, adjusted, or estimated

WHO regularly compiles data on homicide, prevalence of child maltreatment, and intimate-partner and sexual violence in the Global Health Observatory (GHO),[4] which is a gateway to health-related statistics providing data and analyses on a set of global health priorities.

Comprehensive data on national efforts to address interpersonal violence, including the existence of national surveys, policies, laws, prevention programs, and services for victims of violence, are regularly published in WHO's Global Status Reports (Butchart & Mikton, 2014).[5]

The United Nations' Organization on Drugs and Crime publishes the Global Study on Homicide every other year; this report includes homicide rates, trends, mechanisms of homicide, and information about who is most at risk obtained from either criminal justice or public health systems.[6]

The quantity and quality of data for different health outcomes from unintentional injuries and violence vary. Generally, more and better-quality data are available for deaths than for morbidity and disability. Data on the types and severity of nonfatal health outcomes are important, yet challenging to obtain, especially from LMICs. Data on nonfatal health outcomes have primarily been derived from high-income countries, although in the past decade significant information on injuries and violence has emerged from several LMICs. In general, however, the state of routine health information in LMICs has been fragile, especially in regions such as sub-Saharan Africa and South Asia. Thus, it is not surprising that these regions have little tradition of developing specific information sources for unintentional injuries and violence. Notably, population-based studies from these regions frequently suggest that the unintentional injury and violence burden is higher than that reported in national official statistics, indicating that these conditions are significantly underreported in official estimates.

Estimates of Unintentional Injuries and Violence Mortality and Disability

In the following section, data from WHO's Global Health Estimates as well as the most recent global status reports on road safety[7] and violence prevention[5] are presented to demonstrate the global burden of injuries and violence. Additionally, specific studies or sources have been quoted, to highlight recent or unique information from LMICs.

Almost 5 million deaths occurred from all unintentional injuries and violence worldwide in 2015, of which nearly 90% were in LMICs. Approximately 27% of these deaths were caused by road traffic injuries, with suicide and homicide accounting for 16% and 9.5% of deaths, respectively (**TABLE 9-1**). Notably, 14% of deaths were classified as "other" types of unintentional injury deaths, which include injuries such as animal bites, insect bites, and unspecified

[2] http://www.who.int/gho/en/
[3] http://www.oisevi.org/
[4] http://www.who.int/gho/violence/en/
[5] http://www.who.int/violence_injury_prevention/violence/status_report/2014/en/
[6] https://www.unodc.org/gsh
[7] http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/

TABLE 9-1 Distribution of Global Deaths from Injury, 2015

Cause of Injury	Number of Deaths	Proportion of All Injury Deaths (%)
Unintentional Injuries		
Road traffic injuries	1,342,265	27.2
Poisoning	107,705	2.2
Falls	646,271	13.1
Fire, heat, and hot substance burns	180,051	3.6
Drowning	359,717	7.3
Exposure to mechanical forces	201,581	4.1
Natural disasters	14,088	0.3
Other unintentional injuries	674,907	13.7
Intentional Injuries		
Self-harm	788,089	16.0
Interpersonal violence	467,619	9.5
Collective violence and legal intervention	156,271	3.2

Reproduced from World Health Organization (WHO). (2016). *Global health estimates 2015: Deaths by cause, age, sex, by country and by region, 2000–2015*. Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html

injuries. Not surprisingly, injury-related death rates were highest for road traffic injuries, followed by suicide, "other" unintentional injuries, falls, and homicide (**TABLE 9-2**). Somewhat similar patterns were observed for burden of disease rates, which include nonfatal health outcomes (using DALYs); however, falls had higher rates of nonfatal health outcomes due to the high morbidity from such injuries (**TABLE 9-3**). Importantly, the rates were highest for all "other" unintentional injuries, which likely reflects the challenges inherent in coding the specific causes of nonfatal unintentional injuries.

Unintentional Injuries

In 2015, more than 3.5 million unintentional injury deaths occurred, which were collectively responsible for 6.3% of the global mortality burden. The vast majority (90%) of unintentional injury deaths occurred in individuals in LMICs. The worldwide crude death rate for unintentional injuries was 48 deaths per 100,000 population; the highest rate was found in the African region (73.4 per 100,000 population) and the lowest rate in the Americas (36.1 per 100,000 population). Unintentional injuries were also responsible for more than 212 million DALYs in 2015, with 91% occurring in LMICs and the rates being highest in those regions that include large numbers of LMICs.

Road traffic injuries killed more than 1.3 million people in 2015, making these types of injuries the tenth leading cause of death worldwide. Road traffic injuries accounted for 2.4% of global mortality and resulted in a heavy death toll among people in all age categories. According to the Global Health

TABLE 9-2 Global Injury Death Rates per 100,000 Population, 2015

Type of Injury	Number of Deaths (000s)	Proportion of Total Deaths (%)	Crude Death Rate (per 100,000 population)	Age-Adjusted Death Rate (per 100,000 population)
Unintentional Injuries	3,527	6.3	48.0	47.2
Road traffic injuries	1,342	2.4	18.3	18.2
Poisoning	108	0.2	1.5	1.4
Falls	646	1.2	8.8	8.5
Fire, heat, and hot substances burns	180	0.3	2.5	2.4
Drowning	360	0.6	4.9	4.9
Exposure to mechanical forces	202	0.4	2.7	2.7
Natural disasters	14	0.0	0.2	0.2
Other unintentional injuries	675	1.2	9.2	9.0
Intentional Injuries	1,412	2.5	19.2	19.2
Self-harm	788	1.4	10.7	10.7
Interpersonal violence	468	0.8	6.4	6.4
Collective violence and legal intervention	156	0.3	2.1	2.1

Reproduced from World Health Organization (WHO). (2016). *Global health estimates 2015: Deaths by cause, age, sex, by country and by region, 2000–2015*. Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.htm

Estimates, deaths from road traffic injuries are projected to increase in the future, becoming the seventh leading cause by 2030.[8] Globally, more than 1 million of those persons killed from road traffic injuries during 2013 were from LMICs, corresponding to 24 deaths per 100,000 population in low-income countries and 18.4 per 100,000 in middle-income countries (WHO, 2015a). The absolute number of fatalities and the mortality rate resulting from road traffic injuries vary considerably across countries and regions, ranging from 84,143 fatalities (9.3 per 100,000 population) from this cause in the European region to 323,296 deaths in the Western Pacific region (17.3 per 100,000 population) (**FIGURE 9-1**). Although all age groups are affected by fatalities resulting from road traffic injuries, young adults—particularly males, who account for 75% of deaths—are at greatest risk of loss of life. Because this age group corresponds to the most economically productive segment of the population, road traffic injuries have serious implications for national economies (**FIGURE 9-2**).

Road traffic injuries are frequently defined in terms of the types of road users involved. Typically, data specify whether the users were the occupants of four-wheeled motorized vehicles or pedestrians, cyclists, or riders of motorized two-wheeled vehicles, with the latter being described as the *vulnerable road*

[8] http://www.who.int/healthinfo/global_burden_disease/projections/en/

TABLE 9-3 Global Injury DALYs by Sex, 2015

Type of Injury	All	Males	Females
Unintentional Injuries	212,206,821	141,651,044	70,555,777
Road traffic injuries	76,020,322	57,247,296	18,773,027
Poisoning	6,558,469	3,708,519	2,849,950
Falls	31,509,233	18,661,115	12,848,119
Fire, heat, and hot substances burns	12,041,246	5,457,124	6,584,123
Drowning	22,655,470	15,346,187	7,309,284
Exposure to mechanical forces	15,197,575	10,504,673	4,692,902
Natural disasters	1,461,462	753,538	707,923
Other unintentional injuries	46,763,042	29,972,593	16,790,449
Intentional Injuries	78,459,927	55,488,761	22,971,166
Self-harm	37,672,153	23,897,972	13,774,181
Interpersonal violence	27,766,947	22,546,056	5,220,891
Collective violence and legal intervention	13,020,827	9,044,733	3,976,094

Reproduced from World Health Organization (WHO). (2016). *Global health estimates 2015: Deaths by cause, age, sex, by country and by region, 2000–2015*. Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.htm

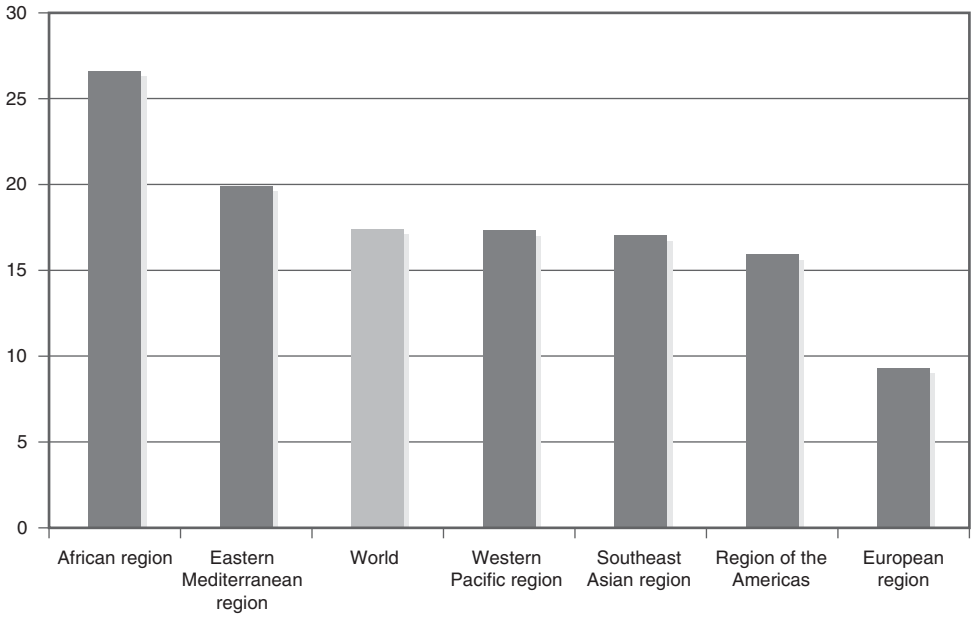


FIGURE 9-1 Road traffic fatality rates per 100,000 population, by WHO region, 2013.

Reprinted from World Health Organization (WHO). (2015). *Global status report on road safety, 2015*. Geneva, Switzerland: World Health Organization.

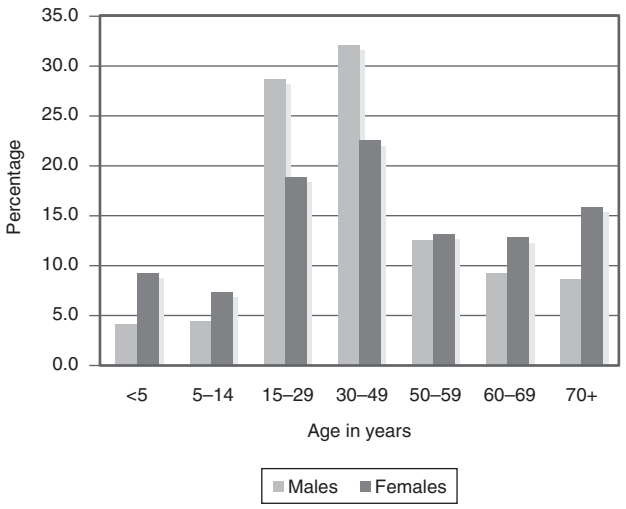


FIGURE 9-2 Global road traffic deaths by age, 2015.

Reproduced from World Health Organization (WHO). (2016). *Global health estimates 2015: Deaths by cause, age, sex, by country and by region, 2000–2015*. Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.htm

users. A recent survey undertaken by WHO (2015a) indicated that vulnerable road users account for almost half of all global deaths and contribute disproportionately to deaths in LMICs (**TABLE 9-4**).

According to WHO estimates, 107,705 *poisoning* deaths occurred in 2015. More than 90% of these deaths occurred in LMICs.

Falls caused more than 646,000 deaths worldwide in 2015, with the majority (83%) of these fatalities occurring in LMICs, and the male-to-female ratio being quite similar but with a distinctive skewing toward older age ranges (**FIGURE 9-3**). These numbers translate to an age-adjusted mortality rate of 8.5 deaths per 100,000 population globally and a loss of approximately 31 million DALYs. Clearly, falls make an important contribution to morbidity and disability on a global basis.

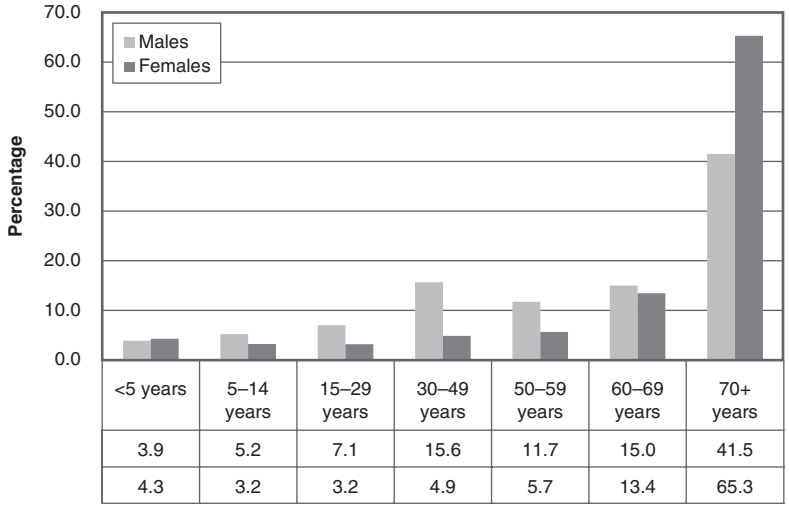
WHO registers data only for *burns* that result from fire, heat, and hot substances; it does not tally data for scalds (water burns)—an important limitation of this data source. More than 180,000 deaths were caused by fires in 2015, resulting in more than 12 million DALYs lost. Unlike other types of injuries, more females than males died from fire-related burns (male-to-female ratio of 0.8:1.00).

Drowning is the process of experiencing respiratory impairment from submersion or immersion in liquid leading to death or morbidity (Bierens, 2006). In 2015, an estimated 360,000 people drowned. More than half of all drownings occurred in the Southeast Asia and Western Pacific regions, and drowning rates in LMICs are more than three times higher than those in high-income countries (WHO, 2014a). Furthermore, more than half of all drowning deaths are among those younger than 25 years, with this type of event being ranked in the top 10 causes of death in most regions (**FIGURE 9-4**). It is important to note that these data include only "accidental drowning and submersion"; they exclude drowning due to floods (cataclysms), boating, and water transport.

TABLE 9-4 Road Traffic Deaths by Type of Road User, by WHO Region, 2013

Region/Area	Cyclists (%)	Pedestrians (%)	Motorized Two- or Three-Wheelers (%)	Car Occupants (%)	Other (%)
Africa	4	39	7	40	11
Americas	3	22	20	35	21
Eastern Mediterranean	3	27	11	45	14
Europe	4	26	9	51	10
Southeast Asia	3	13	34	16	34
Western Pacific	7	23	34	22	14
World	4	22	23	31	21

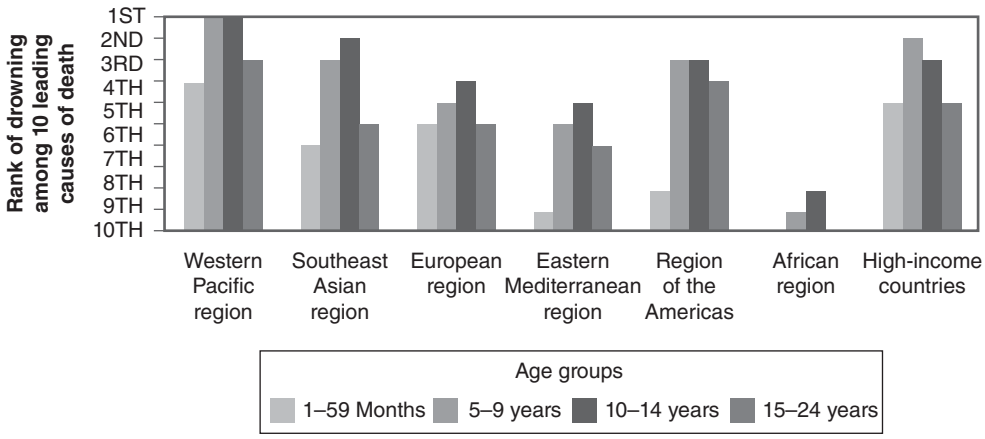
Modified from World Health Organization (WHO). (2015). *Global status report on road safety, 2015*. Geneva, Switzerland: World Health Organization.



<5 years	5–14 years	15–29 years	30–49 years	50–59 years	60–69 years	70+ years
3.9	5.2	7.1	15.6	11.7	15.0	41.5
4.3	3.2	3.2	4.9	5.7	13.4	65.3

FIGURE 9-3 Global falls by age and sex, 2015.

Reproduced from World Health Organization (WHO). (2016). *Global health estimates 2015: Deaths by cause, age, sex, by country and by region, 2000–2015*. Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.htm



Note: Data for all high-income countries appears as 'high-income countries'. All WHO regions provided show ranking for only the LMICs within those regions.

FIGURE 9-4 Rank of drowning among 10 leading causes of death by region and age group.

Reprinted from World Health Organization (WHO). (2014). *Global report on drowning, 2014*. Geneva, Switzerland: World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/10665/143893/1/9789241564786_eng.pdf

Intentional Injuries

In 2015, almost 1.5 million deaths occurred due to violence, which equals 2.5% of the global mortality burden. Most of these deaths occur in LMICs (86%). Regions where the most violent deaths occurred in 2015 included the Eastern Mediterranean region (30.6 per 100,000), the Americas (28.3 per 100,000), and the African region (20.7 per 100,000). The high rate in the Eastern Mediterranean region was mainly driven by the large number of deaths related to the conflict in Syria. Suicide kills more than 788,089 people annually, which makes it the 17th leading cause of death worldwide. Interpersonal violence accounts for more than 467,619 deaths annually, affecting young men in particular (**TABLE 9-5**). Such violence is highly concentrated in terms of geographic area, and in some countries it is the leading cause of death in young people. Collective violence is responsible for 156,271 deaths annually, with approximately one-third of these deaths being due to interpersonal violence. Beyond deaths, violence causes significant impacts on disability and quality of life.

Self-directed violence caused the deaths of more than 788,000 people globally in 2015 and was associated with more than 37 million DALYs lost. These injuries include suicides, attempted suicides, self-destructive behaviors, and self-mutilation. Although data on self-directed injuries are challenging to obtain, case studies from several parts of the world indicate that they are being increasingly documented globally (**EXHIBIT 9-1**). *Preventing Suicide: A Global Imperative* (WHO, 2014b), the first global report on suicide of this kind, calls for more investments in data collection

TABLE 9-5 Intentional Injuries by Sex, 2015

	Both Sexes	Males	Females
Intentional injuries	1,411,979	997,018	414,961
Self-harm	788,089	504,216	283,873
Interpersonal violence	467,619	380,347	87,272
Collective violence and legal intervention	156,271	112,455	43,816

Reproduced from World Health Organization (WHO). (2016). *Global health estimates 2015: Deaths by cause, age, sex, by country and by region, 2000–2015*. Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.htm

EXHIBIT 9-1 Suicides in South Asia: Case Study from Pakistan

No official data on suicides are collected in Pakistan, a conservative South Asian Islamic country with traditionally low suicide rates. As a result, national rates of suicide are neither known nor reported to WHO. Nevertheless, an accumulating body of anecdotal evidence suggests that suicide rates have been gradually increasing in Pakistan over the past few years. To date, both suicide and attempted suicide have been under-studied and under-researched subjects in this country. A general lack of trained mental health researchers in the country is partly to blame for this situation. Furthermore, a lack of interest by available researchers may be related to the generally held belief that suicide and attempted suicide are rare events in Muslim countries and, therefore, unworthy of scientific study. Other reasons behind the paucity of data may be the difficulty in gaining access to such data, as there are many legal, social, and cultural issues related to suicide and attempted suicide. Both suicide and attempted suicide are illegal acts, socially and religiously condemned, making research in this area difficult.

Police data from Sindh—one of the four provinces of Pakistan—provide a unique historical picture of trends of suicide over a 15-year span (1985–1999). During this period, there were 2,568 reported suicides (71% men, 39% women; ratio 1.8:1.00). The lowest number was 90 in 1987, and the highest number was 360 in 1999. Poisoning by organophosphates was the most common method of suicide, followed by hanging.

A more recent review of 23 studies found young age, male gender, marriage, and low socioeconomic status to be reported risk factors for suicides. These data, although limited in scope, provide evidence of an important public health issue of suicides in Pakistan. Firearms, hanging, and organophosphorus poisoning were the most common methods of suicide. There is urgent need for further research on suicide in countries like Pakistan, so that interventions for suicide prevention can then be planned.

Sources

Khan, M. M., & Hyder, A. A. (2006). Suicides in the developing world: Case study from Pakistan. *Suicide & Life-Threatening Behavior, 36*(1), 76–81.

Khan, M. M., & Prince, M. (2003). Beyond rates: The tragedy of suicide in Pakistan. *Tropical Doctor, 33*, 67–69.

Shahid, M., & Hyder, A. A. (2008). Deliberate self-harm and suicides: A review from Pakistan. *International Journal of Injury Prevention & Safety Promotion, 15*(4), 233–241.

on self-harm, while at the same time scaling up interventions to prevent suicide.

Interpersonal violence disproportionately affected LMICs, where more than 90% of all interpersonal violence-related deaths were estimated to have occurred. The estimated rate of violent death in LMICs was 7.1 per 100,000 population in 2015, compared to 2.6 per 100,000 population in high-income countries.

The extent of *child maltreatment* is difficult to gauge because most of this abuse goes unreported. Nevertheless, child maltreatment is a global problem with serious lifelong consequences for many of its victims. In studies of self-reported child maltreatment (Stoltenborgh, van Ijzendoorn, Euser, & Bakermans-Kranenburg, 2011; Euser et al., 2015), the data reveal that 76/1,000 boys and 180/1,000 girls experienced sexual abuse, 226/1,000 children reported physical abuse, and 363/1,000 children reported emotional abuse; 163/1,000 children reported physical neglect, and 184/1,000 children endured emotional neglect. In 2015, almost 31,000 children younger than 15 years died from homicide, although this number probably underestimates the true extent of the problem. A significant proportion of deaths due to child maltreatment are missed because the maltreatment is not apparent to medical practitioners—that is, these deaths appear to be a result of unintentional injuries such as falls or burning. In conflict and refugee

settings, girls and boys may be at particular risk of sexual violence, exploitation, and abuse by combatants, security forces, members of their communities, aid workers, and others.

Youth violence is defined as violence committed by or against individuals between the ages of 10 and 29 years. In 2012, an estimated 205,303 youth homicides took place globally—equivalent to 562 children and young people age 10 to 29 years being killed every day. Youth homicide rates vary dramatically between countries. In some countries of Latin America, the Caribbean, and sub-Saharan Africa, youth homicide rates are 100 or more times higher than the rates in Western Europe and the Western Pacific (WHO, 2015b). In all countries, young males are both the main perpetrators and the main victims of homicide, and female rates of youth homicide are almost everywhere much lower than male rates. For every young person killed by violence, an estimated 20 to 40 more sustain injuries that require hospital treatment (Mercy, Butchart, Farrington, & Cerda, 2002).

Physical fighting and bullying are also common among young people. A global analysis of data from the Global School-Based Student Health Survey found that depending on the region, between 31.4% and 46.7% of adolescents age 13 to 15 have been involved in physical fighting in the past year. Between 2.1% and 5.1% of adolescents report they engaged in frequent physical fighting (more than 12 times) in the past year (WHO, 2015b).

A study of bullying in school-age children in 40 countries (mainly in Europe) found that exposure to bullying varies across countries, with estimated exposure rates ranging from 8.6% to 45.2% among boys, and from 4.8% to 35.8% among girls. Adolescents in Baltic countries reported higher rates of bullying and victimization, whereas their peers in northern European countries reported the lowest prevalence. Boys reported higher rates of bullying in all countries. Rates of victimization were higher for girls in 29 of 40 countries (Craig et al., 2009).

Intimate-partner violence is defined as "any behavior within an intimate relationship that causes physical, psychological or sexual harm to those in the relationship." Such behaviors include "acts of physical aggression—such as slapping, hitting, kicking, and beating; psychological abuse—such as intimidation, constant belittlement, and humiliation; forced intercourse and other forms of sexual coercion; and various controlling behaviors—such as isolating a person from their family and friends, monitoring their movements, and restricting their access to information or assistance" (Heise & García-Moreno, 2002; García-Moreno, Jansen, Ellsberg, Heise, & Watts, 2005). Globally, according to population-based self-report surveys, 30% of ever-partnered women have experienced intimate-partner violence at least once in their life (García-Moreno et al., 2013), and women often become victims of intimate-partner violence repeatedly. The highest prevalence of such violence appears to occur in the WHO African, Eastern Mediterranean, and Southeast Asian regions, where approximately 37% of women have reported they were exposed to intimate-partner violence at some point in their lives.

Sexual violence includes rape (by either an intimate partner or a stranger), attempted rape, gang rape, and other forms of forced sexual acts in any setting (e.g., home, work). Globally, 30% of women have been exposed to intimate-partner violence including sexual violence at least once in their lifetime. Physical and sexual violence by intimate partners overlap to a large extent, and surveys often do not distinguish between physical intimate-partner violence and sexual intimate-partner violence. Global prevalence estimates indicate that 7.2% of women have experienced sexual violence by someone other than their intimate partner (García-Moreno et al., 2013).

Abuse of the elderly includes physical, sexual, emotional, and psychological acts or neglect directed toward older people. It is an important type of violence, especially given the prediction that the global population of persons older than 65 years of age will reach more than 1 billion in the first half of the 21st century (Wolf, Daichman, & Bennett, 2002). More than 80% of these older people will reside in LMICs, where they will represent more than 12% of the population. Abuse of the elderly occurs in both home and institutional settings and is difficult to measure; thus, there is a persistent lack of global statistics on this issue. Globally, 15.7% of older people have experienced abuse: 11.6% of older people reported they experienced psychological abuse, 6.8% reported financial abuse, 4.2% reported they had been a victim of neglect, 2.6% had experienced physical abuse, and 0.9% of older people experienced sexual abuse (Yon, Mikton, Gassoumis, & Wilber, 2017).

Collective violence—including armed conflict with and between states—was estimated by WHO to cause more than 156,271 deaths worldwide in 2015. Men are more than two and a half times more likely to die from collective violence as compared to women. This estimate includes only deaths caused directly by war; it does not include deaths caused indirectly by war, such as those that result from disruption of the health system or those that occur after the end of war.

Estimates of the Economic and Social Costs of Injury and Violence

Loss of life and health is just one dimension in explaining why injuries and violence are an important public health problem. Other important dimensions include these conditions' social and economic toll on individuals, families, and societies. Death of loved ones, lifelong disabilities resulting in unemployment, and the staggering costs of medical care for victims are some of the other effects of injuries and violence that need to be described. In addition, factors such as lost productivity, insurance claims, income-replacement activities, and family consequences are important. Unfortunately, such information is not readily available and is rarely collected in large parts of the world; as a consequence, our knowledge of the economic and social impacts of injuries and violence remains limited. This section, while not presenting a comprehensive review of the economic and social costs of all injuries and violence, provides examples of the breadth and—importantly—the limitations of what has been documented for a few specific injuries.

Costs associated with *unintentional injuries* have been studied in some high-income countries. In one study of unintentional injuries that occurred in home settings in the United States, the total societal cost was estimated at $217 billion; falls accounted for the largest proportion (42%) of this cost (Zaloshnja, Miller, Lawrence, & Romano, 2005). In another study, focusing on childhood unintentional injuries, the total cost, which included loss of future work and quality of life, was $347 billion per year, or $17,000 per child injured (Danseco, Miller, & Spicer, 2000).

A few cost-focused studies have addressed unintentional child injuries in LMICs. For example, analysis of more than 6,000 children younger than 15 years of age who were hospitalized for unintentional injuries in China showed a mean institutional cost of $166 and a mean length of stay exceeding 17 days per injury case (Jiang et al., 2010). Further, the cost associated with poisoning averaged $53 per case; for scald burns, this cost was $198 per case. A community-level analysis in Vietnam showed that the total annual cost of unintentional child injuries was more than $235,000—equivalent to the income of nearly 1,800 people in the country (Thanh, Hang, Chuc, & Lindholm, 2003) (**EXHIBIT 9-2**).

Despite the global significance of *road traffic injuries* in terms of mortality and disability, the economic and social consequences of such injuries have been documented only in the past decade, and most countries in the world report not having undertaken costing exercises (WHO, 2009). Work performed by Transport Research Laboratory (TRL), based on road crash costs from 21 high-income and LMICs, found that the average annual cost of road crashes was approximately 1% of the gross national product (GNP) in LMICs, 1.5% of GNP in countries in economic transition, and 2% in highly motorized countries. Based on these data, the researchers suggested that the annual burden of the economic costs of road traffic injuries globally

EXHIBIT 9-2 Impact of Helmet Laws in Vietnam

On December 15, 2007, a new law mandating the use of helmets among motorcycle riders and passengers took effect in Vietnam. Strong enforcement and severe financial penalties, accompanied by public education through media campaigns including print, radio, and television, were established to support the law. As a result of these efforts, the helmet-wearing rate increased from 30% to 93% (Nguyen et al., 2013) and remained greater than 90% for both riders and passengers even one year after passage of the law (Pervin et al., 2009).

The increased helmet usage resulted in an 18% reduction in the risk of death and a 16% reduction in the risk of head injury among motorcyclists in the country (Passmore et al., 2010). It also saved individuals approximately $18 million in direct acute care costs and averted $29 million in individual income losses in the year following the introduction of the law. The combination of benefits suggests that similar policy changes could be beneficial in countries with high motorized two-wheeler usage but where helmet-wearing rates are low (Olson et al., 2016).

Sources

Nguyen, H. T., Passmore, J., Cuong, P. V., & Nguyen, N. P. (2013). Measuring compliance with Viet Nam's mandatory motorcycle helmet legislation. *International Journal of Injury Control and Safety Promotion, 20*, 192–196.

Olson, Z., Staples, J. A., Mock, C., Nguyen, N. P., Bachani, A. M., Nugent, R., & Verguet, S. (2016). Helmet regulation in Vietnam: Impact on health, equity and medical impoverishment. *Injury Prevention, 22*(4), 233–238.

Passmore, J., Tu, N. T., Luong, M. A., Chinh, N. D., & Nam, N. P. (2010). Impact of mandatory motorcycle helmet wearing legislation on head injuries in Viet Nam: Results of a preliminary analysis. *Traffic Injury Prevention, 11*(2), 202–206.

Pervin, A., Passmore, J., Sidik, M., McKinley, T., Nguyen, T. H., & Nguyen, P. N. (2009). Viet Nam's mandatory motorcycle helmet law and its impact on children. *Bulletin of the World Health Organization, 87*(5), 369–373.

amounted to $518 billion in 2000, with the annual costs of road traffic injury in LMICs being approximately $65 billion, exceeding the total annual amount received by these countries in development assistance at that time (Jacobs, Aeron-Thomas, & Astrop, 2000). These data have more recently been reviewed by the World Bank and the global estimate lost to road traffic deaths and injuries is now estimated to be approximately 3% of gross domestic product (GDP) (Dahdah & McMahon, 2008).

Further work has attempted to document and compare the impact of road traffic injuries on low-income households in Bangladesh and India (Aeron-Thomas, Jacobs, Sexton, Gururaj, & Rahmann, 2004; Prinja, Jagnoor, Chauhan, Aggarwal, & Ivers, 2015). In both of these countries, males who provided the majority of the household income were the most common victims of road traffic fatalities; their deaths reduced household income and food consumption for the victim's family. The poor were also found to spend a much greater proportion of their income on funeral and/or medical costs compared to the nonpoor; as a consequence, road traffic injury was a contributor in tipping households into poverty.

Interpersonal violence is expensive. Estimates of the cost of violence in the United States put this figure at 3.3% of GDP. In England and Wales, the total costs from violence—including homicide, wounding, and sexual assault—amount to an estimated $40.2 billion annually. Although interpersonal violence disproportionately affects LMICs, studies addressing the economic effects of this violence in LMICs remain scarce (Waters, Hyder, Rajkotia, et al., 2004). Estimates from LMICs indicate that the overall costs of violence are substantial (**EXHIBIT 9-3**).

Comparisons with high-income countries are complicated by the fact that economic losses related to productivity tend to be undervalued in lower-income countries because these losses are typically based on foregone wages and income. For example, a single *homicide* is calculated to cost, on average, $15,319 in South Africa, $602,000 in Australia, $829,000 in New Zealand, and more than $2 million in the United States (all monetary values have been converted to 2001 U.S. dollars to enable comparisons and to adjust for inflation and varying exchange rates).

Many of the studies detailing the costs of specific types of violence have been conducted in the United States. The economic burden due to newly identified cases of child maltreatment in one year was estimated at $124 billion (Fang et al., 2012). If other types of violence were included, this number would rise further. A costing study from East Asia and the Pacific estimated the associated treatment costs to mitigate negative health consequences of child maltreatment at 1.4% to 2.5% of the GDP of this region (Fang et al., 2015).

Evidence abounds that the public sector—and thus society in general—picks up much of the tab for interpersonal violence. Several studies in the United States have shown that 56% to 80% of the costs of caring for victims of gunshot and stabbing injuries are either directly paid by public financing or not paid at all; in the latter case, the costs are absorbed by government and society in the form of uncompensated care financing and overall higher payment rates. In LMICs, it is also probable that society absorbs much of the costs of violence, through direct public expenditures and negative effects on investment and economic growth.

EXHIBIT 9-3 Estimating the Economic Costs of Violence

In 2008, WHO, the U.S. Centers for Disease Control and Prevention, and the Small Arms Survey developed economic costing guidelines to assess the direct and indirect economic burden of interpersonal and self-directed violence (Butchart et al., 2008). These guidelines were subsequently tested to assess the costs of fatal and nonfatal interpersonal and self-directed violence in Brazil, Jamaica, and Thailand. In Brazil during 2004, the direct medical cost of injuries due to violence (R$519 million or ~$152 million) amounted to 0.4% of the total health budget, while loss of productivity due to violence-related injuries (R$15.5 billion or ~$4.54 billion) accounted for approximately 12% of all health expenditures, or 1.2% of GDP. In Jamaica, the direct medical costs of interpersonal violence in 2006 totaled J$2.1 billion (~$16.9 million)—the vast majority of which was concentrated among young males. Indirect medical costs were 10 times higher, exceeding J$27.5 billion (~$222 million). Direct medical costs accounted for approximately 160% of Jamaica's total health expenditure, while the combined direct and indirect impacts were equivalent to 4% of GDP. In Thailand for 2005, the direct medical cost of violence-related injuries accounted for approximately 4% of that country's total health budget, while the loss of productivity due to violence-related injuries accounted for approximately 0.4% of GDP.

Source: Butchart, A., Brown., D., Khanh-Huynh, A., Corso, P., Florquin, N., & Muggah, R. (2008). *Manual for estimating the economic costs of injuries due to interpersonal and self-directed violence*. Geneva, Switzerland: World Health Organization.

Studies documenting the economic effects of interpersonal violence have used a broad range of categories to distinguish the various types of costs associated with this condition. Those estimating indirect costs—including the opportunity cost of time, lost productivity, and reduced quality of life—provide higher cost estimates than studies that limit the costs of violence to direct costs alone. Other key methodological issues include the economic values assigned to human life, lost productive time, and psychological distress. The rates at which future costs and benefits are discounted, in accounting terms, also vary across studies.

Risk Factors for Unintentional Injuries and Violence

As with most diseases, the causes of unintentional injuries and violence are considered to be multifactorial. The traditional epidemiologic paradigm of host (including biological and behavioral), vector, and environmental factors, which in combination contribute to the incidence of disease, has been readily adapted and applied in determining the causes of unintentional injury. This paradigm has been extended, however, to consider each factor in relation to the timing of the injury occurrence—that is, factors operating prior to, during, and following the injury that might be associated with both the incidence and the severity of the injury (Haddon, 1968). In determining the causes of violence, especially interpersonal violence, a somewhat different model (described as an ecological model), has more commonly been utilized, which focuses on the interplay between individual, relationship, community, and societal factors (Krug et al., 2002).

In the past few decades, the evidence base identifying risk factors for unintentional injuries and violence has expanded dramatically, as the numbers of both injury researchers and research institutions have increased. The application of public health research methods, commonly used in identifying risk factors for other causes of death and disability, to the problems of unintentional injuries and violence has undoubtedly contributed to the growth in this knowledge. For example, case-control and cohort studies now likely contribute as much to the evidence base in the injury field as they do for other leading causes of death and disability, such as cancer and heart disease.

Most of this research has been undertaken in high-income countries, in large part because of the preponderance of researchers and research institutions in these areas. Nevertheless, the evidence base identifying risk factors for unintentional injuries and violence in LMICs is growing, along with the recognition that certain types of injuries and forms of violence are unique to these countries. Thus, while some risk factors may be common across a wide range of settings (e.g., alcohol misuse), other risk factors are unique to the environments in which they occur (e.g., the significance of water wells in increasing the risks of drowning in LMICs, or the specificity of suicide by self-immolation among women in some Asian societies).

This section summarizes known and potential risk factors for the leading cause-specific injuries. The discussion here highlights the extent to which knowledge has been obtained exclusively in high-income countries, the extent to which this knowledge might be transferred to LMICs, and the areas where the evidence base is still minimal.

Risk Factors for Road Traffic Injuries

Not surprisingly, given the significant burden of road traffic injuries, much is known about risk factors for such injuries, as outlined in detail in *World Report on Road Traffic Injury Prevention* (Peden et al., 2004) and updated most recently in the 2015 *Global Status Report on Road Safety* (WHO, 2015a). The 2004 world report describes road traffic injury risk in terms of four functions: factors affecting exposure to risk, factors influencing crash involvement, factors influencing crash injury severity, and factors influencing the severity of post-crash injuries

Factors Influencing Exposure to Risk

Increasing motorization is, without doubt, one of the main factors contributing to the increase in road traffic injuries worldwide and especially in LMICs. Motorization rates rise with income (Kopits & Cropper, 2003). Thus, in a growing number of LMICs whose economies are experiencing growth, there has been a corresponding increase in the number of motor vehicles used in those countries. Growth rates for vehicle ownership are highest in the world's emerging economies: In 2016, there were a record 72 million new passenger cars on the world's roads and around half of these were produced in middle-income countries (Organization of Motor Vehicle Manufacturers, 2016).

Unfortunately, in some LMICs, traffic growth has included proliferation of less safe forms of travel (i.e., powered two-wheeled vehicles and rickshaws and other three-wheeled auto-rickshaws), resulting in concurrent increases in related injuries (OECD, 2015; WHO, 2015a, 2017a). This growth in powered two-wheeled vehicles is not unique to LMICs, however:

High-income countries are adopting these vehicles in larger numbers as they try to find solutions to the problems of growing traffic congestion. In London, for example, due to the increased use of such vehicles, deaths and injuries among motorized two-wheeler users are increasing (Transport for London, 2016).

Projected demographic changes in high-income countries over the next 20 to 30 years are likely to result in greater numbers of people older than 65 years being exposed to traffic risks and, given their greater physical vulnerability, greater numbers sustaining injury. By comparison, in LMICs, increasing economic growth is fueling both the aspirations of younger people to drive motor vehicles and the need to travel greater distances to work. As a consequence, in many LMICs, this vulnerable road user group will continue to be the predominant population involved in road crashes. In fact, road traffic injuries have been identified as the leading cause of death among adolescents (Sheehan et al., 2017; WHO, 2017b).

Transport, land use, and road network planning have all been shown to be important in determining exposure to injury risk. While many of the technical aspects of planning, highway design, traffic engineering, and traffic management have been the hallmarks of transport systems, such planning systems are frequently absent in LMICs, as evidenced by the fact that only 31 countries surveyed for the global status report had a fully funded national road safety strategy (WHO, 2015a). The necessity for such planning is probably even greater in LMICs than in high-income countries, given the extremely diverse and multiple modes of traffic (both motorized and nonmotorized) seen in the former countries.

Factors Influencing Crash Involvement

The overwhelming influences of speed and alcohol consumption on the risk of crash involvement have been confirmed primarily in studies undertaken in high-income countries, but also in an increasing number of studies undertaken in LMICs. A number of other host-related factors, as outlined in this section, have also been postulated as increasing the risk of crash involvement, although evidence to support their involvement remains limited and is mostly restricted to studies undertaken in high-income countries.

Very good evidence from high-income countries shows a strong relationship between increasing vehicle speeds and increasing risk of crash, both for motor vehicle occupants and for vulnerable road users, particularly pedestrians (Global Road Safety Partnership, 2008; WHO, 2017c, 2017d). This relationship is likely to hold true in LMICs as well. Indeed, data obtained from routinely collected police reports (Sobngwi-Tambekou, Bhatti, Kounga, Salmi, & Lagarde, 2010; Zhao et al., 2009), observational studies (Bachani et al., 2017), and case-control studies (Donroe, Tincopa, Gilman, Brugge, & Moore, 2008) in a number of LMICs show that speed is a leading causal factor in road traffic crashes.

The observation that alcohol is associated with an increased risk of road crashes has been confirmed in many studies of victims and perpetrators conducted both in high-income countries and in LMICs (Global Road Safety Partnership, 2007; Peden et al., 2004). In particular, a survey of studies conducted in LMICs found that alcohol was present in the blood of between 4% and 69% of injured drivers, 18% to 90% of crash-injured pedestrians, and 10% to 28% of injured motorcyclists (Global Road Safety Partnership, 2007). These findings are reinforced by case-control and case cross-over studies that provide clear evidence of alcohol's association with increased rates of road traffic injuries (Odero & Zwi, 1997; Woratanarat et al., 2009).

Other factors that have been shown to increase the risks of road crashes in a number of high-income countries include fatigue, use of hand-held mobile telephones (WHO, 2011), use of illicit drugs (WHO, 2016a), and inadequate visibility of vulnerable road users (Peden et al., 2004); these factors are equally likely to increase these risks in a number of LMICs. Indeed, studies in China show twofold increased risks of car crash associated with driver chronic sleepiness (Liu et al., 2003). In addition, surveys of commercial and public road transport in a number of African countries have shown that drivers often work unduly long hours and go to work when exhausted (Nafukho & Khayesi, 2002), consistent with the overall evidence on the relationship between fatigue and work-related traffic crashes (Robb, Sultana, Ameratunga, & Jackson, 2008).

Clearly, road-related and vehicle-related risk factors may increase the risk of crash involvement. Specific factors related to road planning that have been suggested as risk factors for crashes include through-traffic passing through residential areas, conflicts between pedestrians and vehicles near schools located on busy roads, lack of segregation of pedestrians and high-speed traffic, lack of median barriers to prevent dangerous overtaking on single-carriage roads, and lack of barriers to prevent pedestrian access onto high-speed roads. Studies examining the risks associated with each of these factors are sorely lacking, especially in LMICs, but recent surveys conducted by the International Road Assessment Programme (iRAP, 2015) have revealed that affordable road improvements, such as footpaths, safety barriers, bicycles lanes, and paved shoulders, if applied at

a large scale, could prevent tens of millions of deaths and serious injuries. Although vehicle-related factors clearly have the potential to increase risks of injury, over the last 50 years substantial safety improvements have been made, particularly by vehicles produced in Europe, Japan, and the United States. Even so, 40 countries around the world implement only minimum regulatory standards (WHO, 2015a)—most in emerging markets. The lack of effective vehicle standards contributes to the higher rates of collisions in these countries (Jacobs et al., 2000).

Factors Influencing Injury Severity

In-vehicle crash protection is undoubtedly a factor relating to crash injury severity. While significant improvements have been made to private vehicles in the past decades, many of these advances have not yet made their way into vehicles in LMICs (Odero, Garner, & Zwi, 1997) or are de-specified to keep vehicle costs low. However, the issue of crash protection that reduces injury severity to vulnerable road users is probably of greater relevance, especially in LMICs, where vulnerable road users are predominant. Few countries, whether high- or low-income countries, have established requirements to protect vulnerable road users by means of crash-worthy designs for the front of cars or buses (Mohan, 2002).

A significant risk factor for increased injury severity among motorized two-wheeler users is non-use or inappropriate use of motorcycle helmets (Liu et al., 2008; WHO, 2017a). Likewise, a risk factor for increased injury severity in bicyclists is non-use of helmets (Thompson, Rivara, & Thompson, 2003). Failure to use seat belts is also a significant risk factor associated with injury severity in vehicle occupants, especially in many LMICs that lack any requirement for seat belts to be present in vehicles or used by drivers and passengers (FIA Foundation for the Automobile and Society, 2009; WHO, 2015a).

Studies in high-income countries suggest that roadside hazards, such as trees, poles, and road signs, may contribute to between 18% and 42% of road crash fatalities (Kloeden, McLean, Baldock, & Cockington, 1998). The extent to which this risk applies to LMICs has not been determined.

Factors Influencing Severity of Post-Crash Injuries

Both the availability and the quality of prehospital and in-hospital care are major influences on the outcomes for patients with injuries sustained in a traffic crash. Comparisons between high-income countries and LMICs show clear differences in the proportions of injured individuals who die before reaching a hospital, in large part reflecting the limited access to prehospital medical services in LMICs (Mock, Jurkovich, nii-Amon-Kotei, Arreola-Risa, & Maier, 1998). Factors that determine survival and outcome include early availability of care, the time interval between the injury and the patient's arrival at a definitive-care hospital, referral based on triage, and availability of physical and human resources.

Risk Factors for Poisonings

The global literature on unintentional poisonings includes significant information on occupational-related poisonings, especially pesticide poisonings, and a growing body of information on environmental poisoning, especially lead poisoning. The particular focus of this section, however, is on risk factors for other unintentional poisonings; occupational and environmental poisoning are covered elsewhere in this text under the topic of environmental health.

The literature in this area almost exclusively considers risk factors for poisonings in young children (Peden et al., 2008), even though the majority of poisonings occur in adults, for whom therapeutic errors and adverse reactions to medications may well be significant issues. In high-income countries, product accessibility—in terms of both safe packaging and storage—has long been recognized as the key risk factor in unintentional poisoning (Peden et al., 2008; Shannon, 2000). Storage issues also appear to be risk factors for poisoning in LMICs, including the numbers of used storage containers in the residence, the use of nonstandard containers for storage (e.g., the use of cola bottles for the storage of kerosene), and the storage of poisons at ground level (Peden et al., 2008).

As highlighted in the *World Report on Child Injury Prevention* (Peden et al., 2008), several case-control studies conducted in LMICs show evidence of a number of sociodemographic risk factors for unintentional poisoning, including young age of parents, residential mobility, and limited adult supervision of children (Azizi, Zulkifli, & Kasim, 1993; Soori, 2001). Previous poisoning may also be a risk factor (Soori, 2001).

Risk Factors for Fall-Related Injuries

Fall-Related Injuries in Older People (Including Hip Fractures)

Risk factors for fall-related injuries in older people are generally considered in terms of risk factors for falling (both intrinsic and extrinsic risk factors), risk factors associated with the severity of the impact following the fall, and risk factors associated with low levels of

bone mineral density, insofar as almost all fall-related injuries in older people involve broken bones. The last group of risk factors include low bone density; poor nutritional status and low body mass index (BMI); low calcium intake; comorbid conditions, such as hypertension and diabetes; poor performance in activities of daily living (ADLs and instrumental ADLs [IADLs]); low levels of engagement in physical activity; poor cognitive function; poor perceived health status; poor vision; environmental factors affecting balance or gait; family history of hip fracture; and alcohol consumption (Hippisley-Cox & Coupland, 2009; Robbins et al., 2007). Falls from beds as well as other environmental risk factors related to flooring, lighting, furniture, and fittings such as hand rails are also associated risks (Ambrose, Paul, & Hausdorff, 2013; Lord, Sherrington, Menz, & Close, 2007).

While much of the research in this area has been undertaken in high-income countries, reports from LMICs show that these studies' findings are consistent across most countries (Chew, Yong, Mas Ayu, & Tajunisah, 2010; Coutinho, Bloch, & Rodrigues, 2009). A few studies, however, have identified other factors that have not previously been identified and may be more relevant in the context of LMICs. For example, studies in Thailand have suggested that factors associated with poor socioeconomic status may be risk factors, such as lack of electricity in the house and living in Thai-style houses or huts (Jitapunkul, Yuktananandana, & Parkpian, 2001).

Falls related to the construction industry and on farms are generally more prevalent among working-age adults in low- and middle-income settings (Dandona et al., 2010; Haslam et al., 2005).

Fall-Related Injuries in Younger People

The *World Report on Child Injury Prevention* (Peden et al., 2008) has systematically reviewed the available evidence on risk factors for fall-related injuries in younger people. Much of the research has sought to identify factors associated with falls from heights. In high-income countries, such falls occur from balconies and apartment windows, beds/bunks and nursery equipment (including baby walkers), and playground equipment. In contrast, in LMICs, such falls are more likely to occur from rooftops, trees, and animals, such as in those countries where camel racing is a popular activity. Poverty appears to be a consistent underlying risk factor for falls, while the absence of protective rails/guards or similar equipment is associated with increased risk. Severity of injury appears to be associated with the height of the fall and the nature of the surface on which the child falls.

Risk Factors for Burn-Related Injuries

Burn-related injuries sustained as a result of fires account for the majority of burn-related deaths, although hot water burns are a significant cause of morbidity, especially in LMICs.

In high-income countries, much of our knowledge about risk factors for fire-related injuries has come from cross-sectional studies and only a few case-control studies. In these studies, a number of risk factors have been consistently identified. Risk factors that cannot be modified include age (both younger and older groups), gender, race, or those difficult to modify such as poverty and disability; in contrast, risk factors that are amenable to prevention include place of residence, type of residence (such as mobile homes and homes without a smoke detector or telephone), smoking, and alcohol use as well as unsafe equipment, fireworks, and cooking areas (Peden et al., 2008; Warda, Tenenbein, & Moffatt, 1999). By comparison, few controlled studies have examined risk factors for hot water burns, and most discussion in the literature focuses on the lack of temperature controls for hot water systems or taps (Jaye, Simpson, & Langley, 2001).

Studies undertaken in LMICs have similarly mostly focused on risk factors for burn-related injuries, with few examining risk factors for hot water–related injuries. In the majority of cases, these studies have addressed risk factors for burns in children (Peden et al., 2008). Many of the identified risk factors are related to poverty, such as low family income, lack of water supply, and overcrowding. Lack of parental supervision and history of burns in siblings are risk factors as well. By comparison, maternal education appears to be consistently protective. Not surprisingly, other protective factors relate to the knowledge about burn risks, presence of smoke detectors, and ready access to first aid services.

Risk Factors for Drowning

In high-income countries, the majority of drowning incidents are associated with recreation or leisure activities including swimming pools (Brenner, 2003). In contrast, in most LMICs, these events are associated with everyday activities near bodies of water, including rivers, water wells, or buckets (Hyder, et al., 2003; Kobusingye, Guwatudde, & Lett, 2001; WHO, 2014a).

Only a few studies, in either high- or low-income countries, have examined risk factors for drowning in adults. Nevertheless, evidence consistently shows that a significant proportion of adult drowning incidents are associated with positive blood alcohol levels

(Carlini-Cotrim & da Matta Chasin, 2000; Driscoll, Harrison, & Steenkamp, 2004) as well as flood disasters and traveling on water (WHO, 2014a, 2017e).

Most studies in this area—although there have been few properly controlled studies—have focused on risk factors for drowning in children, as reported in *World Report on Child Injury Prevention* (Peden et al., 2008) and the *Global Report on Drowning* (WHO, 2014a). Children living in rural areas and close to uncovered or unprotected bodies of water appear to be at greatest risk, with lack of parental supervision being noted as a factor across most studies.

Risk Factors for Self-Directed Violence

The report *Preventing Suicide: A Global Imperative* (WHO, 2014b) provides a comprehensive overview of known risk factors for suicidal behavior based on the findings of an extensive body of research, primarily in high-income countries, that has examined the role of psychiatric, biological, social, and environmental factors as well as factors related to an individual's life history.

Case-control studies, using psychological autopsies (i.e., information gathered after death from relatives, healthcare professionals, and medical records), have played an important role in identifying psychiatric risk factors for suicide, as have longitudinal studies. Depression is perhaps the leading psychiatric condition associated with increased risk of suicidal behaviors. Given that this type of mental illness is relatively common, it is not surprising that a large proportion of all suicides are believed to be related to depression. Other conditions associated with increased risk for self-directed violence include schizophrenia, anxiety disorders of conduct and personality, impulsivity, and a sense of hopelessness. Alcohol and drug abuse also play significant roles, although the close relationship between the latter and depression makes it difficult to tease out the independent contributions of these conditions. Most certainly, another important risk factor is previous suicidal attempt, with some studies suggesting that the risk for persons who have attempted suicide in the past could be as high as 20 to 30 times the risk seen in the general population.

Among the biological and medical markers that have been identified as risk factors, family history of suicide is one of the strongest markers, suggesting the possibility of a genetic predisposition; this contention has been supported by twin studies. Other evidence in support of a biological basis for suicide comes from studies of neurobiological processes, particularly those that have examined serotonin levels in psychiatric patients. Such investigations suggest that altered serotonin levels may, in part, be linked to greater risks of suicide. Suicide may also be the consequence of severe and painful illness, although the extent to which any such relationship is independent of psychiatric illness has not been determined as yet.

Certain negative life events may act as precipitating factors for suicide for some individuals, including personal loss (whether through divorce, separation, or death), interpersonal conflict (including bullying), a broken or disturbed relationship, and legal or work-related problems. In particular, studies have shown a higher risk of suicide attempts among victims of violence between intimate partners due to unresolved conflicts, as well as among individuals with a history of physical or sexual abuse in childhood. While in many countries, marriage and having children do appear to be protective factors against suicide, those who marry early may not be equally protected. Studies undertaken in Pakistan and China also suggest that married women, especially older married women, may not necessarily be protected from this risk (Khan & Reza, 1998; Yip, 1998). In contrast, individuals who are socially isolated appear to be at increased risk for suicide, including homosexual adolescents and the elderly.

Social and environmental factors that are believed to increase the risk of suicidal behavior include availability of the means of suicide; a person's place of residence, employment, or immigration status; affiliation with a religion; and economic conditions. Numerous studies, undertaken in a wide range of countries, provide clear evidence that both the ready availability of means of suicide and the lethality of those methods influence the incidence of suicide attempts and completions. Likewise, studies conducted in high-, middle-, and low-income countries show that rural residence is an important risk factor for suicide, possibly related to issues of social isolation and increased accessibility of means of suicide. While some exceptions exist, there does appear to be a consistent relationship between greater religious involvement and lower risk of suicide. Both the economic prosperity of a community and personal economic circumstances appear to be related to risk of suicide, as observed in various studies.

Risk Factors for Interpersonal Violence

The quantity and nature of research on the risk factors for interpersonal violence vary across the categories of child maltreatment, intimate-partner violence, youth violence, sexual violence, and elder abuse, as does the extent to which the research has focused on risk factors associated with perpetrators versus risk factors associated with victims.

While certain risk factors are specific to these different types of interpersonal-directed violence and to perpetrators or victims, a number of common risk factors have been observed across at least three of these different categories and, to some extent, in both high-income and LMICs. In terms of individual risk factors, these characteristics include being male and young, abusing alcohol and other drugs, and being a victim of child abuse or neglect. Family risk factors include low socioeconomic status of the household, marital discord, and parental conflict involving use of violence. Community risk factors include low social capital in the community, high crime levels in the community of residence, and poor access to or inadequate medical care and situational factors. Lastly, societal risk factors include rapid social change (leading to the breakdown of traditional value and social support networks), economic inequality, poverty, weak economic safety nets, poor rule of law and high levels of corruption, a culture of violence, gender inequalities, high firearms availability, and punitive responses to perpetrators and conflict/post-conflict scenarios (Butchart, Phinney, Check, & Villaveces, 2004; Mercy, Butchart, Rosenberg, Dahlberg, & Harvey, 2008; Rosenberg, Butchart, Mercy, Narasimhan, & Waters, 2006).

A number of risk factors for *child maltreatment* have been identified through research in various countries, but not all areas of the world; thus, one cannot assume that they apply to all social and cultural contexts. Nevertheless, the list presented here does provide an overview of risk factors that should be considered when attempting to understand the causes of child maltreatment in any context (WHO & International Society for Prevention of Child Abuse and Neglect, 2006). Child maltreatment risk factors include characteristics and attributes of the child, and of the parent and caregiver. Of course, while they may have risk factors, children themselves are the victims and are never to blame for the maltreatment. Age is a major risk factor among children who suffer maltreatment. Children younger than 4 years of age are at greatest risk of severe injury and death from maltreatment, but across all forms of child maltreatment combined, adolescents are at greatest risk. Other risk factors in children include being an unwanted infant or failing to fulfill parental expectations; having special needs, crying persistently, or having abnormal physical features; and demonstrating symptoms of mental ill health, or personality or temperament traits that are perceived as problematic.

Parental or caregiver factors that can increase the risk of child maltreatment include difficulty bonding with a newborn or not nurturing the child; having been maltreated as a child; lacking awareness of child development or having unrealistic expectations; inflicting inappropriate, excessive, or violent punishment; lacking self-control when upset or angry; misusing alcohol or drugs, including during pregnancy; involvement in criminal activity; being depressed or exhibiting feelings of low self-esteem or inadequacy; and experiencing financial difficulties. Factors within family, friend, intimate-partner, and peer relationships that may increase the risk of child maltreatment include lack of parent–child attachment; physical, developmental, or mental health problems of a family member; family breakdown or violence; being isolated in the community or lacking a support network; breakdown of support for childrearing from the extended family; and discrimination against the family.

Characteristics of communities and societies that can increase the risk of child maltreatment include gender and social inequality; social, economic, health, and education policies that lead to poor living standards or to socioeconomic inequality or instability; lack of adequate housing or services to support families and institutions; high levels of unemployment or poverty; rapid rates of in- and out-migration to neighborhoods; the easy availability of alcohol or a local drug trade; inadequate policies and programs to prevent child maltreatment, child pornography, child prostitution, and child labor; and social and cultural norms that demand rigid gender roles, support the use of corporal punishment or severe physical punishment of children, and diminish the status of the child in parent–child relationships.

Risk factors associated with *youth violence* exist at the level of the individual, relationship, community, and society (Mercy et al., 2002). Individual-level factors include personality and behavioral traits such as hyperactivity; impulsiveness; poor behavioral control; attention problems; history of early aggressive behavior; early involvement with drugs, alcohol, and tobacco; antisocial beliefs and attitudes; low intelligence; low commitment to school and school failure; residence in a single-parent household; experiencing parental separation or divorce at a young age; and exposure to violence and conflict in the family. Additionally, studies show that drunkenness is an important immediate situational factor that can precipitate youth violence.

Factors within family, friend, intimate-partner, and peer relationships that may increase the risk of youth violence include poor monitoring and supervision of children by parents; harsh, lax, or inconsistent disciplinary practices; a low level of attachment between parents and children; low parental involvement in activities; parental substance abuse or criminality; poor family functioning and a low level of family cohesion; low family income; and witnessing

violence or experiencing abuse as a child. Associating with delinquent peers has also been linked to violence in young people.

At community and societal levels, poor social cohesion within a community has been linked to higher rates of youth violence, as have rapid urbanization and unemployment leading to the social dislocation of young people and the erosion of informal social controls. Gangs and a local supply of guns and drugs represent a potent mixture that increases the likelihood of youth violence. Factors such as extreme and highly visible socioeconomic inequalities, rapid social and demographic changes in the youth population, and urbanization can also increase the risk of youth violence. In addition, the quality of a country's governance—its laws and the extent to which they are enforced, as well as policies for social protection—has important implications for such violence.

Intimate-partner violence and *sexual violence* share a number of common risk factors, as described in the WHO and London School of Hygiene & Tropical Medicine (2010) document titled *Preventing Intimate Partner and Sexual Violence: Taking Action and Generating Evidence*. Women and men with lower levels of education are at increased risk of victimization and perpetration of intimate-partner violence, respectively. Childhood sexual abuse is strongly associated with the perpetration of intimate-partner violence and sexual violence by men as well as the victimization of women in the form of intimate-partner and sexual violence. A diagnosis of antisocial personality disorder is a strong risk factor for the perpetration of both intimate-partner and sexual violence. Alcohol abuse is consistently found to be associated with the perpetration of both types of violence as well. Males who have multiple sexual partners or who are suspected by their partners of infidelity are more likely to initiate both intimate-partner and sexual violence. In addition, attitudes that are accepting of violence are strongly associated with perpetration and victimization in the form of such violence (WHO & London School of Hygiene & Tropical Medicine, 2010).

Risk factors that are specific to intimate-partner violence include a past history of violence, which for both perpetrators and victims is a strong risk factor for future intimate-partner violence. Marital discord and dissatisfaction are strongly associated with both perpetration of and victimization by this kind of violence. Likewise, ineffective court/police responses are associated with intimate-partner violence (WHO & London School of Hygiene & Tropical Medicine, 2010).

Risk factor research on *elder abuse* is scant. Although a variety of risk factors have been hypothesized to exist, the evidence cited to support many of these hypotheses—such as the significance of caregiver stress—is not convincing. One of the most consistent findings, however, is the important role of social isolation. Additionally, while little empirical evidence has been documented to support this contention, cultural norms relating to ageism and sexism are thought to play an important role in elder abuse (Krug et al., 2002).

Risk Factors for Collective Violence

As outlined in greater detail in *World Report on Violence and Health* (Krug et al., 2002), the Carnegie Commission on Preventing Deadly Conflict (Carnegie Commission, 1997) has identified a range of factors that—either alone or in combination—may precipitate the risk of political violence. In brief, these factors include political and economic factors (e.g., lack of democratic processes, unequal access to power, distribution of resources and access to resources, and control over both key natural resources and drug production or trading), societal and community factors (e.g., inequalities between groups, fueling of group fanaticism, and ready availability of weapons), and demographic factors (particularly rapid demographic change).

Interventions to Prevent Unintentional Injuries and Violence

Interventions to prevent unintentional injuries have traditionally been considered in terms of the *three E's*: education, enforcement, and engineering. Such measures are also described within the framework of the Haddon matrix—that is, in terms of their focus on preventing the occurrence of the injury event, versus minimizing the severity of injury at the time of the injury event, versus minimizing the severity of injury following the injury event. By comparison, interventions aimed at reducing violence have addressed individual, relationship, community, and societal approaches, mirroring the ecological model outlined in the previous section.

While randomized controlled trials, whether they involve individuals or communities, represent the gold standard by which the effectiveness of preventive interventions might best be assessed, such trials remain relatively uncommon in the injury and violence field. Studies comparing the incidence of injury before and after the implementation of an intervention, sometimes with reference to a control population (i.e., a group in which the intervention has not been introduced), more commonly provide evidence of

effectiveness. Of course, randomized controlled trials are clearly not needed for interventions whose benefits are obvious. In contrast, for other interventions, particularly those that may have modest but important benefits, rigorous evaluation methods are required.

Studies addressing the effectiveness of interventions in LMICs, as distinct from their effectiveness in high-income countries, are also relatively uncommon. While the proven efficacy of some interventions in high-income countries does not require replication in LMICs (e.g., the use of motorcycle helmets), strategies that may be effective in increasing the rates of helmet-wearing in high-income countries may not necessarily be appropriate in LMICs; thus, specific evidence of their effectiveness is required. For this reason, tailoring of the interventions found effective in high-income countries to ensure that they are appropriate in LMICs, with rigorous evaluation to confirm their applicability, is increasingly being endorsed. Moreover, determining how best to implement programs based on such interventions is increasingly needed and implementation research is often needed in these instances.

Interventions to Prevent Road Traffic Injuries

"Safer systems, safer roads, safer vehicles, and safer people" is the motto employed by those working to reduce road traffic injuries (Peden et al., 2004). In the public health sector, much of the emphasis in terms of interventions has—not surprisingly—focused on the last of these elements (WHO, 2015a). In addition, there is significant political support for road safety. The United Nations General Assembly, the World Health Assembly, and more recently the inclusion of a target related to road safety in the 2030 Agenda for Sustainable Development have all encouraged countries to implement good practices in regard to preventing road traffic injuries. Guidance is provided to countries through the global plan for the Decade of Action, which includes five pillars—road safety management, safer infrastructure, safer vehicles, safer road users, and post-crash response (United Nations Road Safety Collaboration, 2011)—and more recently through the Save LIVES technical package on road safety, which encourages countries to put in place 22 evidence-based interventions (WHO, 2017d) (**TABLE 9-6**).

Safer Systems

A focus on safer systems involves managing exposure to the risk of a road traffic injury, by accounting for human error through managing interactions between the environment, infrastructure, and the road user (OECD, 2008). In particular, it involves implementing strategies aimed at reducing motor vehicle traffic, encouraging use of safer modes of traffic, and minimizing exposure to high-risk scenarios. Reductions in motor vehicle traffic are possible through application of efficient fuel taxes, land-use restrictions, and safety impact assessments of transport and land-use plans; by

TABLE 9-6 Save LIVES: 6 Components and 22 Interventions

Acronym	Component	Interventions
S	**Speed management**	Establish and enforce speed limit laws nationwide, locally, and in cities
		Build or modify roads to calm traffic (e.g., roundabouts, road narrowing, speed bumps, chicanes, and rumble strips)
		Require car makers to install new technologies, such as intelligent speed adaptation, to help drivers keep to speed limits
L	**Leadership**	Create an agency to spearhead road safety
		Develop and fund a road safety strategy
		Evaluate the impact of road safety strategies
		Monitor road safety by strengthening data systems
		Raise awareness and public support through education and campaigns

I	**Infrastructure design and improvement**	Provide safe infrastructure for all road users, including sidewalks, safe crossings, refuges, overpasses, and underpasses
		Put in place bicycle and motorcycle lanes
		Make the sides of roads safer by using clear zones, collapsible structures, or barriers
		Design safer intersections
		Separate access roads from through-roads
		Prioritize people by putting in place vehicle-free zones
		Restrict traffic and speed in residential, commercial, and school zones
		Provide better, safer routes for public transport
V	**Vehicle safety**	Establish and enforce motor vehicle safety standard regulations related to: ■ Seat belts ■ Seat belt anchorages ■ Frontal impact ■ Side impact ■ Electronic stability control ■ Pedestrian protection ■ ISOFIX child restraint points
		Establish and enforce regulations on motorcycle antilock braking
E	**Enforcement of traffic laws**	Establish and enforce laws at national, local, and city levels on: ■ Drinking and driving ■ Motorcycle helmets ■ Seat belts ■ Child restraints
S	**Survival**	Develop organized and integrated prehospital and facility-based emergency care systems
		Train those who respond to crashes in basic emergency care
		Promote community first responder training

Reprinted from World Health Organization (WHO). (2017). *Save LIVES: A road safety technical package*. Geneva, Switzerland: World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/10665/255199/1/9789241511704-eng.pdf

the provision of shorter, safer routes; and through trip reduction measures, including greater emphasis on the development and use of public transport systems. Minimizing exposure to high-risk scenarios includes strategies such as restricting access to different parts of the road network, giving priority in the road network to higher-occupancy vehicles or to pedestrians and slow-moving transports, placing restrictions on speed and engine performance of motorized two-wheelers, separating different traffic modes, increasing the legal age for operating motorized two-wheelers, and introducing graduated driver licensing systems. While it seems obvious that the implementation of these strategies should lead to reductions in road traffic injuries, only limited documentation of their effectiveness in low-income contexts is readily available.

Safer Roads

Intervention strategies focusing on safer roads need to incorporate safety awareness in planning road networks, safety features in road design, and remedial action

at high-risk crash sites. Although they may not have been examined in rigorously controlled studies, many of these strategies have been adapted and adopted over many years in both high-income countries and LMICs.

Traffic-calming measures are among the strategies recommended with respect to incorporating safety features into road design (Table 9-6). Systematic reviews of the evidence confirm that traffic-calming measures reduce road traffic injuries, although evidence from randomized controlled trials is not yet available (Bunn et al., 2003; Retting, Ferguson, & McCartt, 2003; WHO, 2017d). Although most of the evidence cited in support of these measures comes from high-income countries, a "before and after" study conducted in Ghana showed clearly that speed bumps were effective in reducing traffic injuries, and especially in decreasing pedestrian injuries (Afukaar, Antwi, & Ofosu-Amah, 2003).

The introduction of speed cameras (designed to capture images of drivers violating speed limits) and other speed-enforcement devices has also been shown to be highly cost-effective in reducing road traffic injuries (WHO, 2017c, 2017d). Nevertheless, the quality of these studies remains low and relatively few have been conducted in LMICs (ICF Consulting & Imperial College Centre for Transport Studies, 2003; Wilson, Willis, Hendrikz, & Bellamy, 2006).

Other safety features that might be incorporated into road design include provisions for slow-moving traffic and for vulnerable road users, lanes for overtaking other road users, median barriers, street lighting, advisory speed limits, and systematic removal of roadside hazards, such as trees or utility poles (Beyer & Ker, 2009; Peden et al., 2004).

Safer Vehicles

Strategies focusing on safer vehicles that have been suggested include a core set of seven United Nations vehicle regulations for motor vehicles and one for motorcycles that are expected to decrease the incidence of road traffic injuries and save lives (WHO, 2015a, 2017d). The seven vehicle regulations are standards on frontal and side impact; electronic stability control; pedestrian protection; seat belts and anchorages; and child restraints. A recent study conducted in Latin America revealed that more than 440,000 deaths and serious injuries could be averted and up to $143 billion saved if basic UN vehicle regulations were applied to key countries in the region by 2030 (Wallbank, McRae-McKee, Durrell, & Hynd, n.d.).

Establishing and enforcing regulations on motorcycle antilock braking and daytime running lights are important evidence-based interventions for powered two-wheelers (WHO, 2017a, 2017d). The positive effects of daytime running lights on motorcycles have been shown in LMICs (Radin Umar, Mackay, & Hills, 1996; Yuan, 2000), while meta-analyses of the effects of (automatic) daytime running lights on cars, particularly in countries in the northern hemisphere, have shown that they decrease the frequency of road crashes, including pedestrian and cycle crashes (Elvik & Vaa, 2004).

The installation and maintenance of seat belts in cars is probably the most well-known and most effective crash vehicle design strategy. Nevertheless, data from surveys conducted in 2009 showed that more than one-fourth of all countries that manufacture or assemble cars do not require that seat belts be fitted (WHO, 2009). By 2013, although 161 countries had national seat belt laws, only 105 of them had laws that met best practice standards by requiring all occupants to be belted (WHO, 2015a). Among other intelligent vehicle devices that might prove useful in LMICs are speed limiters and ignition interlock devices that prevent alcohol users from starting a vehicle (McGinty et al., 2017).

Safer People

In recent decades, effective intervention strategies aimed at improving road user behavior have increasingly focused on the introduction and enforcement of relevant legislation, combined with enforcement and education, rather than advocating educational efforts alone (Poli de Figueiredo, Rasslan, Bruscagin, Cruz, & Rocha e Silva, 2001; Staton et al., 2016; WHO, 2013). Many of these legislative intervention strategies have addressed speeding, alcohol use and driving, motorcycle helmet use, seat belts, and child restraints (WHO, 2015a). Despite this focus, few countries have national and subnational laws relating to all five of these risk factors. The situation is most advanced in regard to seat belt laws, as 105 countries now have laws that match the best practice (**FIGURE 9-5**). Even when legislation has been passed, enforcement of the laws is often weak, especially in LMICs.

A large body of research—albeit little of it conducted in LMICs—shows that the setting and enforcement of speed limits can lead to reductions of as much as 34% in road traffic injuries (Peden et al., 2004; WHO, 2017c). Similarly, numerous studies demonstrate that establishment and enforcement of legal blood alcohol limits, minimum drinking-age laws, and use of alcohol checkpoints lead to important reductions in road traffic injuries, albeit of varying magnitude (Peden et al., 2004).

Both bicycle helmets and, in particular, motorcycle helmets have been shown to have a significant impact in reducing head injuries among riders (Liu et al., 2008;

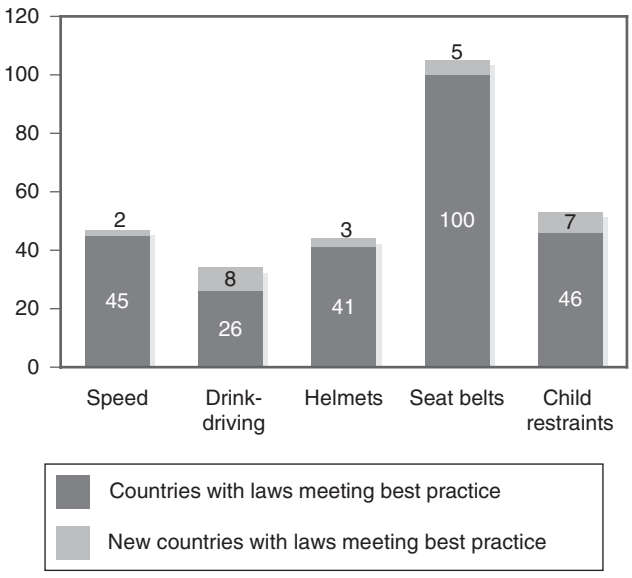


FIGURE 9-5 Countries with road safety laws meeting best practice.

Reprinted from World Health Organization (WHO). (2015). *Global status report on road safety, 2015*. Geneva, Switzerland: World Health Organization.

Thompson et al., 2003; WHO, 2017a). The introduction of mandatory seat belt laws and mandatory child restraint laws has been shown to have a major impact in reducing occupant deaths and injuries—decreases of as much as 25% have been noted in some areas (FIA Foundation for the Automobile and Society, 2009; WHO, 2007). Additionally, systematic reviews have shown the greater effectiveness of enforcement strategies that allow enforcement officers to specifically stop and check seat belt use (primary enforcement) compared with strategies that allow seat belt use to be checked only when other enforcement strategies are the focus of the traffic stop (secondary prevention) (Dinh-Zarr et al., 2001; Rivara, Thompson, Beahler, & MacKenzie, 1999). Moreover, many studies have shown that enforcement needs to be selective, highly visible, and well publicized; conducted over a sufficiently long period; and repeated several times a year (European Transport Safety Council, 2016; Solomon, Ulmer, & Preusser, 2002).

Interventions to Prevent Poisonings

The prevention of unintentional poisonings includes consideration of both occupational and non-occupational poisonings, including household poisonings. Efforts to prevent occupational exposures include the promotion, legislation, and enforcement of nonchemical methods of pest control and the promotion of the safe use of pesticides when nonchemical methods are not feasible. Interventions such as these largely fall within the domain of specialists working in occupational and environmental health and are not considered further in this chapter.

Suggested interventions to reduce exposure to non-occupational poisonings include better storage of poisonous materials, in terms of both the storage position and the nature of the storage vessels (Nixon, Spinks, Turner, & McClure, 2004; Peden et al., 2008). With respect to the former, the suggested interventions include storing poisonous materials outside the home, and at levels beyond the reach of children (above head height). With respect to the nature of the storage containers, it has been suggested that efforts need to be directed toward reducing the use of *secondhand* household containers (e.g., cola bottles), including the introduction and enforcement of legislation to prohibit sales of poisons in such containers (Nhachi & Kasilo, 1994). While these interventions clearly have merit, evidence demonstrating their effectiveness is lacking.

The efficacy of child-resistant containers (CRCs) in preventing access to poisons has been demonstrated, however. According to data from a controlled "before and after" study undertaken in South Africa, the free distribution of CRCs appears to be a highly effective means of preventing poisoning in children (Krug, Ellis, Hay, Mokgabudi, & Robertson, 1994).

Other interventions that have been suggested, but not rigorously examined, include the use of warning labels on poison packaging, appropriate first aid education, and the introduction of poison control centers that are charged with monitoring the incidence of poisonings and providing appropriate preventive advice (Nixon et al., 2004; Peden et al., 2008). Likewise, home safety education and the provision of safety education have been proposed as potentially effective interventions; while the available evidence appears promising, the impact of such interventions on poisoning rates is unclear (Kendrick et al., 2008).

Interventions to Prevent Fall-Related Injuries

Fall-Related Injuries in Older People

In community-dwelling individuals, exercise interventions show the greatest promise in reducing the risk and rate of falls. These include multiple-component group exercises such as Tai Chi, which involve balance and strengthening exercises, as well as individually prescribed multiple-component home-based exercise (Gillespie et al., 2012).

Assessment and multifactorial interventions have been shown to reduce the rate of falls but not the risk of falling, as has gradual withdrawal of psychotropic medication. Overall, vitamin D supplementation does

not reduce the risk of falls, although this measure may well do so in persons with lower vitamin D levels. Similarly, home safety interventions do not reduce falls overall, but are effective in preventing falls among persons with visual impairment and among others at higher risk of falling. Other effective interventions are directed toward individuals who reside in certain living environments and those with specific medical conditions (Gillespie et al., 2012).

Some evidence indicates that multifactorial interventions can reduce both the number of falls and the risk of falling in hospitals, and may do so in nursing care facilities as well. Vitamin D supplementation is also effective in reducing the rate of falls in nursing care facilities. By comparison, exercise in subacute hospital settings appears effective in decreasing falls, but its effectiveness in nursing care facilities remains uncertain (Cameron et al., 2012).

Hip protectors initially appeared to be a promising intervention to reduce the impact of a fall in older people. More recently, a growing number of studies have questioned the effectiveness of these devices, especially given the relatively poor compliance rates that are achieved in real-life settings (Cameron et al., 2010; Parker, Gillespie, & Gillespie, 2006).

Fall-Related Injuries in Younger People

The *World Report on Child Injury Prevention* document (Peden et al., 2008) highlights a number of areas where interventions to prevent falls among children have been shown to be effective or appear promising. These strategies include engineering measures focused on the redesign of equipment, such as nursery equipment, and environmental measures that focus on the redesign of playgrounds or buildings. Previous successes with the Children Can't Fly program in low-income areas in the United States have been replicated in low-income countries, where the use of window guards have been shown to be a promising approach to the prevention of falls from buildings. By comparison, while educational strategies for the prevention of falls in children might seem to be appealing at first glance, little evidence supports the effectiveness of such approaches in reducing the rates of falls (Kendrick et al., 2008; Kendrick et al., 2013).

Interventions to Prevent Burn-Related Injuries

Installed, working smoke detectors have been found to reduce the risk of death from fires by up to 50%; the challenge, however, is making sure that all homes have working smoke alarms on all levels of the residence (DiGuiseppi et al., 2010). Most studies in high-income countries have focused on the effectiveness of smoke detector giveaway programs, community- and school-based educational campaigns, and community-based burns prevention measures. Unfortunately, these education-based approaches have little evidence supporting their effectiveness in terms of reduced incidence of fires and fire-related injuries. Notably, combining building codes and legislation on smoke alarms with education seems to have a greater effect in achieving this goal (DiGuiseppi et al., 2010).

Interventions that have been proposed but not yet shown to be effective largely rely on reducing exposure to fires and flames (Peden et al., 2008). These measures include separating cooking areas from living areas (including efforts to reduce the use of indoor fires for cooking), ensuring cooking surfaces are at appropriate heights, reducing the storage of flammable substances in households, and providing for greater supervision of younger children. Other proven or promising interventions include the introduction of safer lamps and stoves, child-resistant lighters, banning the sale of fireworks to children, and treatment of patients at dedicated burn centers (Peden et al., 2008).

Evidence for the effectiveness of interventions to prevent water-related burn injuries is minimal, but promising. In an increasing number of high-income countries, much effort has been directed toward measures that not only educate communities about the dangers of high tap water temperatures, but also legislate and enforce efforts to regulate the temperature of water flowing from household taps (DiGuiseppi et al., 2010; Macarthur, 2003). In contrast, in low-income countries, scalds due to hot water are more likely to be associated with cooking and boiling of water, rather than water from taps. Consequently, interventions related to the separation of the cooking areas from living and play areas have been emphasized, as well as suggestions that cooking vessels holding water might be better designed to minimize the chances of spillage.

Finally, interventions directed at increasing awareness of burns prevention (whether fire or water related) have been proposed, especially given the success of *safe community* interventions involving a multitude of strategies (Ytterstad & Sogaard, 1995). To date, limited evidence shows that such interventions can be successful (Peden et al., 2008; Turner, Spinks, McClure, & Nixon, 2004).

Interventions to Prevent Drowning

Evidence for the effectiveness of interventions to prevent drowning has been recently summarized in a new global report on drowning (WHO, 2014a). In this

report, ten actions to prevent drowning were identified. These included community-based actions, such as installing barriers (e.g., pool fencing, doorway barriers and playpens, well covers); providing safe places for children away from water; teaching school-age children basic swimming skills; and training bystanders. In addition, the report recommended more effective policies or laws, such as setting and enforcing safe boating standards; building resilience and managing floor risk; coordinating drowning efforts; and developing a national water safer plan. Finally, the report authors called for well-designed studies to address priority research questions in this area (WHO, 2014a).

The effectiveness of learn-to-swim programs, while a common component of prevention programs in high-income countries, have only recently been studied in LMICs. Programs in Australia, Bangladesh, China, Thailand, the United States, and Vietnam have shown that teaching children (older than 5 years) basic swimming, water safety, and safe rescue skills reduces the risk of drowning; in one study in Bangladesh, such a program was shown to be cost-effective (Rahman et al., 2012).

Education regarding the burden and risk factors for drowning, especially the risks posed by prior consumption of alcohol, has also been postulated as a potential intervention strategy (Quan, Bennett, & Branche, 2007). Increased supervision of children around bodies of water and the provision of lifeguards at popular swimming areas have been proposed as other measures that might reduce drowning (Hyder et al., 2003; WHO, 2017e).

Recent work largely done in Bangladesh has been exploring the use of barriers such as playpens and door barriers, as well as keeping at-risk children away from the hazard in community day care centers or crèches (Hyder et al., 2014a, 2014b). The crèches appeared very promising on earlier studies (Rahman et al., 2012); in a large study conducted more recently, they were found to protect children age 1–4 years from drowning deaths. The crèches also involve community mobilization, the development of village injury prevention committees, and educational programs for both children and parents (Hyder et al., 2014a, 2014b).

Interventions to prevent water-related transport drowning include equipping boats with flotation devices, ensuring that boats and flotation devices are well maintained, introducing legislation and enforcing regulations relating to the maximum number of individuals who may be carried on specific types of boats, and providing fully trained and responsive coast guard services (WHO, 2014a). However, as with the preventive measures proposed to prevent non-transport-related drowning, evidence of these interventions' effectiveness is lacking (Peden et al., 2008).

Interventions to Prevent Self-Directed Violence

Knowledge about risk factors for suicide and effective prevention strategies has increased dramatically over the past decade, although only 28 countries today have national suicide prevention strategies (WHO, 2014b). Effective strategies for preventing suicides and suicide attempts include restricting access to the most common means, including pesticides, firearms, and certain medications (WHO, 2014b). Attempts to limit the availability of suicide means have included reducing access to sedatives, pesticides, carbon monoxide in domestic gas and in car exhausts, and handguns in the home. Safer storage, bans, and replacement of pesticides with less toxic alternatives could prevent many of the estimated 370,000 suicides caused by ingestion of these chemicals every year. Although international conventions have long attempted to manage hazardous substances, many highly toxic pesticides continue to be widely used. Research suggests, however, that bans must be accompanied by evaluations of agricultural needs and replacement with low-risk alternatives for pest control.

Early identification and effective management of mental disorders and harmful alcohol use can contribute to reducing suicides and should be systematically integrated in the health sector response. Community-based prevention efforts include encouraging attendance at suicide prevention centers and self-help groups and school-based interventions that involve training of school staff, community members, and healthcare providers to identify those persons at risk and refer them to treatment. By comparison, societal approaches include both restricting access to the means of suicide and managing media reports of suicides.

Interventions to Prevent Interpersonal Violence

This section is adapted from *INSPIRE—Seven Strategies for Ending Violence Against Children*, which is a technical package of prevention strategies that have shown success in reducing interpersonal violence, in particular violence against children (WHO, 2016b). The strategies were selected based on convergence in the research-based guidance already published by the 10 international agencies that endorsed INSPIRE. While the INSPIRE technical package is specifically aiming to reduce violence against children age 0–18 years, most of these strategies have been shown to have preventive effects on violence in adulthood as well, and benefits in areas

such as mental health, education, and crime reduction (WHO, 2016b). This section summarizes the seven strategies and spotlights the effectiveness of those interventions within each strategy that have the strongest evidence for preventing violence. **TABLE 9-7** presents an overview of the seven strategies, specific approaches within each strategy, and the sectors that are usually responsible for its implementation.

Implementation and Enforcement of Laws

The aim of this strategy is to ensure that laws to prevent violent behaviors, reduce excessive alcohol use,

TABLE 9-7 INSPIRE Strategies, Approaches, and Sectors for Preventing and Responding to Violence Against Children

Strategy	Approaches	Sectors	Cross-cutting Activities
Implementation and enforcement of laws	■ Laws banning violent punishment of children by parents, teachers, or other caregivers ■ Laws criminalizing sexual abuse and exploitation of children ■ Laws that prevent alcohol misuse ■ Laws limiting youth access to firearms and other weapons	Justice	Multisectoral actions and coordination
Norms and values	■ Changing adherence to restrictive and harmful gender and social norms ■ Community mobilization programs ■ Bystander interventions	Health, education, social welfare	Multisectoral actions and coordination
Safe environments	■ Reducing violence by addressing "hotspots" ■ Interrupting the spread of violence ■ Improving the built environment	Interior, planning	Multisectoral actions and coordination
Parent and caregiver support	■ Delivered through home visits ■ Delivered in groups in community settings ■ Delivered through comprehensive programs	Social welfare, health	Multisectoral actions and coordination
Income and economic strengthening	■ Cash transfers ■ Group saving and loans combined with gender equity training ■ Microfinance combined with gender norm training	Finance, labor	Monitoring and evaluation
Response and support services	■ Counseling and therapeutic approaches ■ Screening combined with interventions ■ Treatment programs for juvenile offenders in the criminal justice system ■ Foster care interventions involving social welfare services	Health, justice, social welfare	Monitoring and evaluation
Education and life skills	■ Increase enrollment in preschool, primary schools, and secondary schools ■ Establish a safe and enabling school environment ■ Improve children's knowledge about sexual abuse and how to protect themselves against it ■ Life and social skills training ■ Adolescent intimate-partner violence prevention programs	Education	Monitoring and evaluation

Reprinted from World Health Organization (WHO). (2016). *INSPIRE: Seven strategies for ending violence against children*. Retrieved from http://apps.who.int/iris/bitstream/10665/207717/1/9789241565356-eng.pdf

and limit youths' access to firearms and other weapons are both implemented and enforced (WHO, 2016b, p. 31). Implementation and enforcement of such laws has the potential to reduce physical violence, sexual abuse and exploitation, and firearms-related deaths, and nonfatal injuries among children (DeSimone, Markowitz, & Xu, 2013; Matzopoulos, Thompson, & Myers, 2014; Modig, 2009; Santaella-Tenorio, Cerdá, Villaveces, & Galea, 2016). Laws prohibiting violent behaviors including sexual abuse or violent punishment of children signal that society does not consider these behaviors acceptable (Modig, 2009; United Nations, 2014). Such laws provide a way to hold perpetrators accountable for their actions. Laws and policies can also reduce key risk factors for violence against children, such as excessive alcohol consumption and binge drinking (Fitterer, Nelson, & Stockwell, 2015; Wagenaar, Toomey, & Erickson, 2005; Wechsler & Nelson, 2010).

Norms and Values

This strategy aims to strengthen norms and values that support nonviolent, respectful, nurturing, positive, and gender-equitable relationships for all children and adolescents (WHO, 2016b, p. 37). Achieving this often requires modifying deeply ingrained social and cultural norms and behaviors—in particular, the idea that some forms of violence are not only normal, but sometimes justifiable. It involves approaches such as community mobilization programs, bystander interventions, and small-group programs that challenge harmful gender and social norms (WHO, 2016b, p. 37).

Successful implementation of this strategy should result in greater recognition of what constitutes abusive behavior toward intimate partners and children, a reduced acceptance of this violence, and, ultimately, a decrease in the occurrence of such violence (Abramsky et al., 2014; Jewkes et al., 2008; Kyegombe et al., 2015; Paine et al., 2002; Raising Voices, 2016; Usdin, Scheepers, Goldstein, & Japhet, 2005; Verma et al., 2008). In addition, it is intended to produce more favorable attitudes and beliefs toward nonviolent approaches to parental discipline, gender equity, and gender-equitable division of labor (Dworkin, Hatcher, Colvin, & Peacock, 2013; Lundgren, Beckman, Chaurasiya, Subhedi, & Kerner, 2013).

Safe Environments

This strategy aims to create and sustain safe streets and other environments where children and youth gather and spend time, and in so doing reduce assault-related injuries (WHO, 2016b, p. 43). It focuses on modifying communities' social and physical environment (rather than the individuals within it) to foster positive—and deter harmful—behaviors. It involves approaches like problem-oriented policing directed toward "hotspots" for violence, interrupting violent conflicts by stopping retaliatory violence, and changing the built environment through, for example, lighting, creation of green places in vacant lots, and renovating abandoned buildings (Braga, Papachristos, & Hureau, 2012; Cassidy, Inglis, Wiysonge, & Matzopoulos, 2014; Cerdá et al., 2012; Florence, Shepherd, Brennan, & Simon, 2011, 2014; Henry, Knoblauch, & Sigurvinsdottir, 2014; Picard-Fritsche & Cerniglia, 2013; Skogan, Harnett, Bump, & DuBois, 2009; Webster, Whitehill, Vernick, & Parker, 2012).

Parent and Caregiver Support

This strategy aims to reduce harsh parenting practices and create positive parent–child relationships by helping parents and caregivers understand the importance of positive, nonviolent discipline and of close, effective communication. Parent and caregiver support can be provided through parent training programs delivered via home visits or in groups. In these approaches, parents are educated about their children's development and trained to adopt positive parenting practices, such as nonviolent discipline, and to engage in effective parent–child communication on sensitive topics (WHO, 2016b, p. 49).

The parent and caregiver support strategy has been shown to reduce child maltreatment; bullying and being bullied; physical, emotional, and sexual violence victimization by partners or peers; and aggression and delinquency during adolescence; and to increase parental monitoring of child and youth safety (Beets et al., 2009; Bilukha et al., 2005; "Building Happy Families," 2014; Kärnä et al., 2011; Knerr, Gardner, & Cluver, 2013; Knox & Burkhart, 2014; Nurse Family Partnership, 2011; Olds et al., 1998; "Parents Make the Difference," 2014; Salmivalli & Poskiparta, 2012; Washburn et al., 2011).

Income and Economic Strengthening

This strategy aims to improve families' economic security and stability, thereby reducing intimate-partner violence and child maltreatment. It involves interventions such as making cash transfers to families in combination with parent training and/or on condition that they ensure their children attend school, or providing microfinance in combination with education for men and women on gender norms, domestic violence, and sexuality (WHO, 2016b, p. 55).

The income and economic strengthening strategy can be expected to reduce physical violence

toward children by parents or other caregivers; early and forced marriage of young girls; intimate-partner violence; children witnessing intimate-partner violence in the home; and social norms and attitudes that approve of intimate-partner violence (Cancian, Yang, & Slack, 2013; Gupta et al., 2013; Huston et al., 2003; Jan et al., 2011; Kim et al., 2007; Ozer, Fernald, Manley, & Gertler, 2009; Pronyk et al., 2006; Pronyk, Hargreaves, & Morduch, 2007; Vyas & Watts, 2009).

Response and Support Services

This strategy aims to improve access to good-quality health, social welfare, and criminal justice support services for all children who need them—including for reporting violence—to reduce the long-term impact of violence (WHO, 2016b, p. 61). Children who have experienced violence need access to a variety of health and support services to help them heal. These services can also help break the cycle of violence in children's lives and help them better cope and recover. Basic health services, such as emergency medical care for violence-related injuries, and clinical care for victims of sexual violence (including post-exposure prophylaxis against human immunodeficiency virus [HIV] in cases of rape when indicated), are the main priority (WHO, 2016b, p. 61). Where these basic health services are in place and are child focused, mechanisms for children to seek help, protection, support, and care—and to report violent incidents—can be put in place. These mechanisms can include counseling and therapeutic approaches, screening combined with interventions, treatment programs for juvenile offenders in the criminal justice system, and foster care interventions involving social welfare services (WHO, 2016b, p. 61).

Successful implementation of the strategy should reduce the recurrence of the same type of violence in the short term; victimization or perpetration of violence in the short and long terms; sexually transmitted infections and negative reproductive health outcomes; and trauma symptoms (e.g., post-traumatic stress disorder, depression, anxiety) (Dubowitz, Feigelman, Lane, & Kim, 2009; Koehler, Lösel, Akoensi, & Humphreys, 2013; MacMillan & Wathen, 2014; Murray et al., 2015; Wethington et al., 2008).

Education and Life Skills

This strategy aims to increase children's access to education, social–emotional learning, and life-skills training, and to ensure that school environments are safe and enabling (WHO, 2016b, p. 67). Gains in education protect against both victimization and perpetration of violence. This strategy involves approaches such as increasing enrollment of children in preschool, primary schools, and secondary schools; establishing a safe and enabling school environment; improving children's knowledge about sexual abuse and how to protect themselves against it; adolescent intimate-partner violence prevention programs; and life and social skills training programs. Programs that strengthen children's social and emotional learning enhance their communication and relationship skills and help them learn to solve problems, deal with emotions, empathize, and safely manage conflict—all life skills that can prevent violence (WHO, 2016b).

The education and life skills strategy can be expected to result in higher levels of school attendance and academic success and increased awareness of, and improved attitudes toward, intimate-partner violence. It should also lead to reductions in bullying, aggressive, and violent behaviors; physical and sexual intimate-partner violence victimization and perpetration; and drug use and excessive alcohol use (Chaux, 2007; Devries et al., 2015; Foshee et al., 2005; Hahn et al., 2007; Institute of Medicine, 2011; Reynolds, Temple, Ou, Arteaga, & White, 2011; Shek & Ma, 2012; Wallace, 2006; Wilson & Lipsey, 2007; WHO, 2015a).

Effects of Multiple Strategies

Research specifically on the combined effects of multiple INSPIRE strategies has yet to become available, but some examples of such studies can be gleaned from the literature on HIV prevention. For instance, a study in rural South Africa examined the effects on mental health, substance abuse, violence, and risk behavior for HIV/AIDS of a program geared toward social protection (cash) and psychosocial support (care). In boys, violence perpetration incidence was 20% without cash and care provision, 14% with cash provision, 13% with care alone, and 9% with cash plus care. In girls, sexual exploitation incidence was 14%, without cash or care, 11% with cash provision, 11% with care, and 8% with cash plus care; and HIV risk behavior incidence was 15% without cash and care, 10% with cash provision, 10% with care, and 7% with cash plus care (Cluver, Orkin, Boyes, & Sherr, 2014). The findings from these studies suggest that the additive effects of multiple INSPIRE strategies could be large, and that a priority for future research should be to evaluate the appropriate effects of multiple interventions.

Interventions to Prevent Collective Violence

The prevention of collective violence involves reducing the potential for violent conflicts and providing appropriate responses to violent conflicts when they

occur. Not surprisingly, scientific evidence about the effectiveness of such interventions is lacking (Zwi, Garfield, & Loretti, 2002). A systematic review focusing on the evidence for approaches to reduced sexual violence in the context of armed conflict and humanitarian crises found only one study that specifically addressed the disaster setting. Most studies focused on opportunistic forms of sexual violence committed in post-conflict settings (Spangaro, Adogu, Ranmuthugala, Davies, Steinacker, & Zwi, 2013). The evidence is inconclusive, but it appears that multicomponent interventions that engaged the affected communities were among the more promising approaches. Nevertheless, policies that facilitate reductions in poverty, that make decision making a more accountable process, that reduce inequalities between groups, and that reduce access to biological, chemical, nuclear, and other weapons have been recommended. When planning responses to violent conflicts, recommended approaches include assessing at an early stage which populations are most vulnerable and what their needs are, coordinating the activities of the various players, and working toward global, national, and local capabilities so as to deliver effective health services during the various stages of an emergency (Zwi et al., 2002).

The Role of Health Services in Preventing Death and Disability from Unintentional Injuries and Violence

While the focus of this section has been on identifying strategies to prevent the occurrence of injuries, it is important to recognize that access to health services plays a major role in preventing death and disability from injuries and violence. Differential access to health and medical services has been identified in other fields of health as having an important bearing on long-term outcomes, and the importance of such services is now being increasingly recognized in the injury field (Razzak & Kellermann, 2002).

For example, although prehospital transport systems exist in major cities in LMICs, the prehospital care that is delivered to trauma victims in the field is typically minimal or nonexistent in rural areas. Not surprisingly, this lack of quality emergency medical services care and timely transport to the hospital translates into higher mortality rates. Significant reductions in mortality have been reported as a result of improvements in trauma care (Mock, Lormand, Joshipura, Goosen, & Peden, 2004). WHO has recently suggested that even in environments where fewer resources are available, injury mortality and disability can be reduced by reorganizing systems, upgrading the skills of the health staff, and ensuring minimum physical and human resources are available (Mock, Quansah, Krishnan, Arreola-Risa, & Rivara, 2004).

In most LMICs, access to violence-related health services is limited. In the Violence Against Children Surveys, fewer than 10% of children who experienced sexual violence reported receiving services in Cambodia, Haiti, Kenya, Malawi, Swaziland, Tanzania, and Zimbabwe (Sumner et al., 2015).

Basic health services to mitigate the negative consequences of violence include emergency medical care for violence-related injuries and clinical care for victims of sexual violence, including post-exposure prophylaxis against HIV in cases of rape, when indicated. These services should be available at the primary healthcare level everywhere. More comprehensive services include counseling and social services for victims and perpetrators of violence; notably, such services can help victims better cope with violence (WHO, 2016b).

Economic Analyses of Interventions to Prevent Unintentional Injuries and Violence

Currently, there is a dearth of literature on economic evaluations of interventions for injuries. This scarcity of research in large part reflects the infancy of this field and the limited, but growing evidence base identifying effective preventive strategies. Unfortunately, even for interventions that are routinely used in high-income countries, proof of their cost-effectiveness is not always easy to find. The situation is even more critical in LMICs, where evaluations of these interventions are rarely available.

Exemplifying this paucity as well as the potential importance of such data are two reviews that documented economic analyses of interventions directed at reducing road traffic injuries (Hyder, Waters, Philipps, & Rehwinkel, 2004; Waters, Hyder, & Phillips, 2004). These reviews, which relied on mostly data from high-income countries, suggest that interventions such as mandatory helmet laws, laws mandating motor vehicle inspections, installation of automatic daytime running lights, and seat belt laws are likely to provide widespread benefits with high benefit–cost ratios, even in LMICs. A recently published article by Sheehan et al. (2017) indicated that investing $5.20 per capita across 75 LMICs in five areas, including road safety, would result in economic and social benefits 10 times their costs and save up to 12.5 million lives. This study identified seven road safety interventions—namely, behavioral measures (helmet use, speed compliance, alcohol testing, seat belt use, and graduated licensing), motor vehicle safety

measures, and infrastructure expenditures (safer roads)—as most effective at reducing deaths, injuries, and disabilities among adolescents.

Costs of interventions for injuries were analyzed as part of the Disease Control Priorities project (www.dcp-3.org/dcp2/; Norton, Hyder, Bishai, & Peden, 2006). Specifically, Bishai and Hyder (2006) estimated the cost-effectiveness of five interventions that could reduce injuries in LMICs: enforcing traffic laws more strictly, erecting speed bumps, promoting helmets for bicycle riders, promoting helmets for motorcycle riders, and storing kerosene in childproof containers. They estimated what each intervention would cost in six world regions over a 10-year period from both governmental and societal perspectives, with costs being measured in U.S. dollars for 2001. Some data were available on the effectiveness of each intervention and were used to form models of DALYs averted by each strategy for various regions. The interventions had cost-effectiveness ratios ranging from $5 to $556 per DALY averted, depending on the region. Enhanced speeding control, for example, was found to be highly cost-effective at $93 per DALY. Similarly, at $13.98 per DALY, placing speed bumps at 25% of the most dangerous junctions turned out to be highly cost-effective, although this measure requires the identification of such intersections. Bicycle helmet legislation is not universally implemented in the high-income world, yet its cost-effectiveness ($170 per DALY) makes it an attractive option. Interestingly, motorcycle helmet legislation in East Asia was found to have a higher cost and lower benefit at $556 per DALY. At $96 per DALY, child-resistant containers are a highly cost-effective intervention that warrants serious consideration by a large part of the low- and middle-income world where paraffin (kerosene) is used for cooking and heating purposes, including sub-Saharan Africa, South Asia, East Asia, and parts of the Middle East.

Empirical work conducted in LMICs remains rare. One notable example of such work is a study that examined the costs and potential effectiveness of increasing traffic enforcement in Uganda, as assessed on the four major roads to the capital Kampala (Bishai, Asiimwe, Abbas, Hyder, & Bazeyo, 2008). By using monthly data on traffic citations and casualties for the years 2001 to 2005 and employing time series regression, the researchers computed costs from the perspective of the police department. The average cost-effectiveness of better road safety enforcement in Uganda was determined to be $603 per death averted or $27 per life-year saved discounted at 3% (amounting to 1.5% of Uganda's $1,800 GDP per capita).

Economic impact studies on other types of unintentional injuries are scarce, especially from LMICs. A recent study reported that the average out-of-pocket payments varied significantly (range: $8–$830) by injury type and outcome (fatal versus nonfatal), while the total injury out-of-pocket expenditures were $355,795 and $5,000 for nonfatal and fatal injuries, respectively, per 100,000 people in Bangladesh (Alfonso et al., 2017). In the same study, the majority of household heads with injuries reported financial distress following that injury.

Although relatively few economic evaluations of interventions targeting interpersonal violence have been published, the available studies suggest that behavioral, legal, and regulatory interventions cost less money than they save, in some cases by several orders of magnitude (Waters et al., 2004). An evaluation of the Cardiff model, which strengthened data collection regarding the place, the time, the mechanism, and the perpetrators of violence-related injuries in hospital emergency departments, and anonymized and shared this information with the police, showed that this model reduced hospital admissions by 42% and police-recorded injuries by 32%. These reductions represented savings of £6.9 million (~$9.6 million) per year and a cost–benefit ratio of £82 million (~$115 million) in benefits for each £1 million ($1.4 million) spent on the program (Florence, Shepherd, Brennan, & Simon, 2014). A review of the costs and benefits of early intervention programs to prevent child maltreatment concluded that some home-visiting programs targeting high-risk/low-income mothers returned between $2 and $3 for each $1 spent (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004). In a further review of nine early childhood programs, seven were found to be cost-effective, yielding between $2 and $17 in benefits for every dollar invested (Kilburn & Karoly, 2008). Despite this benefit, both reviews concluded that not all childhood interventions are cost-effective, with some being ineffective and expensive.

Along with outcome evaluation studies, investigations into the cost-effectiveness of preventive strategies for interpersonal violence should be a research priority.

Advancing the Injury and Violence Prevention Agenda: Opportunities and Challenges

Over the past decades, our knowledge about the burden of injuries and violence, as well as risk factors and effective interventions addressing these issues, has increased exponentially. In many high-income countries, the benefits of this increased knowledge

have resulted in significant declines in injury- and violence-related mortality and morbidity rates. In contrast, for many LMICs, increases in rates of injury and violence are predicted over the next 20 years, as a result of both changing socio-demographic patterns and these countries' success in addressing the burden of communicable disease and maternal and child ill health. Consequently, there is an imperative to rapidly advance an injury and violence prevention agenda, while recognizing that many of the challenges faced by LMICs will be different from those that have been faced in the previous decades.

Advocacy for Injury and Violence Prevention

It remains critical that the global health community fully understand the health, economic, and social impacts of unintentional injuries and violence. Likewise, the health sector must recognize that clearly defined risk factors for injuries and violence exist and that effective interventions can prevent these problems and reduce the burden of death and disability to which they give rise. One of the challenges is that while the health sector often has to deal with the consequences of injuries and violence, many of the approaches to prevent injuries and violence come under the auspices of other sectors. However, despite the recent efforts by WHO and many injury and violence prevention professionals throughout the world, many governments and the wider community often fail to recognize that injuries are no "accident" and that violence does not "just happen"—and that we can work to prevent increases in the incidence of these conditions. For this reason, targeted and evidence-based advocacy is essential at local, national, and international levels.

The key messages to be transmitted in advocacy efforts are as follows: Injuries and violence cause considerable death and disability; they are predictable events with clearly identified causes; and they can be effectively prevented. Moreover, there is a great need, especially in LMICs, to initiate national conversations around key injury issues and stimulate an intersectoral approach to address them. The WHO reports on violence and health, on road traffic injury prevention, and on child injury prevention, for example, propose such an approach to national stakeholders and stress the need for national ownership and local action (Krug et al., 2002; Peden et al., 2004; Peden et al., 2008; WHO, 2014b). Some early successes can be seen at the global level. For example, two ministerial summits on global road safety were held in Moscow in 2009 and in Brasilia in 2015, led by the United Nations Road Safety Collaboration; national leaders have pledged their support for the "Decade of Action for Global Road Safety 2010–2020"; many countries have sponsored United Nations General Assembly resolutions on "improving global road safety" as well as two World Health Assembly resolutions (2004 and 2016); and most recently two road safety targets were included in the 2030 Agenda for Sustainable Development. These high-level advocacy events call on all members of the global community to pay attention to the increasing burden of road traffic injuries and encourage them to invest in evidence-based efforts to improve road safety.

The emergence of violence prevention as a key component of the global health agenda can in part be ascribed to the role played by the World Health Assembly and its partner organizations in highlighting violence prevention and injury prevention in general. Since the publication of *World Report on Violence and Health* in 2002, two World Health Assembly resolutions have called on countries to invest in violence prevention; by 2010, three out of six WHO regional committees (Africa, the Americas, and Europe) had adopted similar resolutions. The inclusion of violence prevention on the agenda of other multilateral agencies is another indication of this topic's emergence as a global priority. The World Health Assembly particularly recognized the need to strengthen national health systems in addressing interpersonal violence, and adopted a resolution in during the 67th assembly in 2014. WHO's member states adopted the "global plan of action to strengthen the role of the health system within a national multisectoral response to address interpersonal violence, in particular against women and girls and against children" at the 69th World Health Assembly in May 2016.

For the first time in history, violence and injury prevention have taken a prominent place on the global development agenda. The Sustainable Development Agenda 2030 has several targets that directly focus on the reduction of homicide, violence against children, violence against women, and road traffic accidents. It further addresses a number of important risk factors for violence and injury prevention, such as harmful alcohol use.

The United Nations General Assembly has also reviewed special reports on violence against children, violence against women, and armed violence, which have resulted in resolutions calling for greater investment in multisectoral efforts to address these forms of violence.

Research and Development Needs

To date, global investments in injury and violence research and interventions, compared with

investments made in other health areas, have not matched the burden of disease. Historically, road traffic injuries have been identified as a highly neglected area for investments compared with the burden of disease they represent, measured in dollars per DALY (approximately $0.40 per DALY), especially when considered in relation to other major health problems (Ad Hoc Committee on Health Research Relating to Future Intervention Options, 1996). And in the past century, it was not surprising that national analyses of safety investments demonstrated low investment rates in low-income countries such as Pakistan and Uganda—approximately $0.07 per person in Pakistan (Bishai, Hyder, Ghaffar, Morrow, & Kobusingye, 2003).

Some notable exceptions to this pattern do exist, such as the multi-million-dollar investments made by Bloomberg Philanthropies to establish model intervention areas in 10 countries and 10 cities since 2009 to demonstrate the impact of road traffic injury prevention and control measures (Hyder & Bishai, 2012; Hyder, Allen, Peters, Chandran, & Bishai, 2013). Even so, while we already know a great deal about the burden, causes, and effective interventions for injuries and violence, significant gains in our understanding of these factors could still be made if equitable resources were directed to the field.

Given that 90% of the world's population lives in LMICs, and the burden of injuries and violence is predicted to increase in these countries, it is imperative that future research and development activities focus especially on the needs of these countries. While it would be inappropriate to outline a detailed agenda for future research and developments that is relevant to all countries, a generic framework that encompasses the types of research required can be recommended.

Epidemiologic research describing the existing burden, causes, and distribution of unintentional injuries and violence is still needed in LMICs. Assessing the loss of health and life from unintentional injuries and violence, identifying the populations affected by these conditions, and determining the specific circumstances in which they occur all deserve a place on the continuing research agenda for LMICs. Problems of underreporting and other biases in the available data need to be addressed, along with the specific impacts of these conditions on the very poor. Organizations such as the global Road Traffic Injuries Research Network[9] are attempting to promote key research on such issues and produce important findings that benefit prevention and control (Road Traffic Injuries Research Network Multicenter Study Collaborators, 2013).

Of course, it is critical not only to identify the determinants of unintentional injuries and violence, but also to address them. The lack of *intervention research* in LMICs has left a huge gap in health research globally. Scientific trials of injury and violence prevention interventions have generally not been conducted in these countries, and there is a great need to modify, adapt, and test both existing and proposed interventions in these settings. While some might argue that such research should be a priority in most LMICs, unless the basic research on the burden and determinants of unintentional injuries and violence has been conducted, the political and financial support for such follow-on research will not be forthcoming.

The lack of empirical information on the cost-effectiveness of injury and violence prevention interventions remains a major policy issue, especially given that interventions used in high-income countries might potentially be transferred to LMICs without regard to their appropriateness or relevance in the latter regions. Perhaps equally important is the concern that without such research, highly cost-effective interventions will not be implemented at all. *Policy-oriented research* that identifies the barriers to implementation and translation of research findings is, therefore, an essential component in advancing the injury and violence prevention agenda.

The Significance of a Trained Workforce

Research and programs in the field of unintentional injuries and violence require trained individuals with specific skills and tools. Injury and violence prevention is a science replete with conceptual frameworks, epidemiologic approaches, intervention testing methods, and analytical techniques. Well-developed training programs in high-income countries have emerged that can produce a workforce who can carry the injury and violence prevention agenda forward. At present, the general lack of such trained human resources in the low- and middle-income world means that advocates must necessarily focus on developing the appropriate human resources capacity in these countries.

[9] https://rtirn.net/

Issues of both quantity and quality arise with regard to the development of a trained workforce. Clearly, there needs to be a critical mass of trained injury and violence prevention professionals in a country if authorities there are to understand, develop, and implement interventions. At the same time, the field covers a broad range of topics, so a diverse group of skills are needed for prevention. Epidemiology, statistics, health information systems, health policy, economics, sociology, and criminology—these are merely some of the fields that have a role in reducing the burden of disease from injuries and violence. In this context, the role of the public health sector is both to assume leadership in developing and organizing a response to injuries and violence and to facilitate the inputs provided by other fields and disciplines. For this reason, management and leadership skills are important assets for a well-trained workforce in injury and violence prevention.

Efforts to address these capacity development needs are being implemented both in high-income countries and in LMICs. For example, WHO has created a teaching curriculum (TEACH-VIP) that may be freely used in training programs and courses and that covers all basic aspects of injury prevention and control.[10] It focuses on enhancing the knowledge base of participants and their technical foundations for work in the injury field. A number of electronic learning courses are also available. The most recent course on legislation and road safety, developed by WHO in collaboration with Johns Hopkins Bloomberg School of Public Health's International Injury Research Unit, was launched in 2017.

WHO also has a free mentoring program (MENTOR-VIP) that allows people from LMICs to enroll in a one-year mentoring relationship with more experienced colleagues from around the world.[11] The mentoring program focuses on hands-on learning and the development of specific skills for practicing injury prevention and control.

Other examples of capacity building include the work by partner organizations such as the Road Traffic Injuries Research Network, which facilitates the development of both junior and senior researchers from LMICs who are seeking to develop their research skills in road traffic injuries research. In addition, WHO Collaborating Centers such as the Johns Hopkins International Injury Research Unit have a free, online road safety training program that leads to the award of a certificate.[12]

The Roles of National and International Organizations

The roles of national agencies (e.g., ministries of health and medical research councils), international organizations (e.g., the United Nations Development Program [UNDP], WHO, UNICEF, and the World Bank), and international societies (e.g., International Federation of the Red Cross and Red Crescent, Handicap International) need to be emphasized in advancing the injury and violence prevention agenda. The roles of national organizations include the following:

- *Advocacy*: Accepting ownership of the problem and promoting engagement with other national institutions.
- *Implementation*: Developing policy and ensuring programs are available on the ground in countries for addressing injuries and violence.
- *Evaluation*: A continuous process of assessing interventions and programs, in addition to supporting nationally relevant research.

The role of international organizations and societies can be summarized as follows:

- *Strategic*: First internally recognizing the toll on societies of injuries and violence, and then convincing governments externally to appreciate the same. This has already been done in the case of road traffic injuries by the joint WHO-World Bank report in 2004 and for child injuries by the joint WHO–UNICEF report in 2008.
- *Facilitation*: Providing technical assistance, appropriate advice, and relevant tools for national policies and action.
- *Resource mobilization*: Demonstrating true commitment to supporting the implementation of effective interventions in countries and promoting research in the field.

The international movements currently under way for violence prevention and road traffic injuries prevention are examples of how joint global–national partnerships are needed for making change.

[10] www.who.int/violence_injury_prevention/media/news/2010/13_01_2010/en/index.html
[11] www.who.int/violence_injury_prevention/capacitybuilding/mentor_vip/en/
[12] https://www.jhsph.edu/research/centers-and-institutes/johns-hopkins-international-injury-research-unit/

Conclusion

Unintentional injuries and violence represent a major global health problem, and one that will increase in magnitude unless systematic, scientifically based approaches to prevention are implemented. While the public health community has recognized the significance of the global injury epidemic only relatively recently, much new knowledge has been generated in a short time. Much of this new knowledge has focused on the burden of injury in high-income countries, however. The challenge for the future, therefore, is to extend our knowledge to the growing injury and violence burden faced by LMICs.

The international public health community has an important role to play in addressing this challenge, by facilitating the description of the problem, the development of solutions, the implementation of programs, and the analysis of effects. Moreover, the public health community can play a leadership role in galvanizing a multisectoral response to injuries and violence, advocating for investments at national and international levels, and catalyzing the sharing of experiences around the world. The large and often devastating health impacts of injuries and violence make it imperative that the international public health community not only take an interest in these issues, but also engage in proactive efforts to address the burden of disease from injuries and violence at local, national, and international levels.

In summary, injuries are no "accident," and violence does not "just happen"; rather, these are highly predictable and preventable events. The international public health community has already played an important role in taking this message forward. The challenge is to continue to do so, thereby ensuring that the predicted epidemics of unintentional injuries and violence in LMICs can be prevented.

Discussion Questions

1. Why should public health programs in LMICs address injuries and violence?
2. List three indicators you would expect to change as a result of a road traffic injury program in a country. Which data would you need to measure these changes? How would you obtain those data?
3. Choose a specific type of injury. Which steps can be taken to reduce the burden of this injury in LMICs?
4. Suppose you were funding research on violence. Describe five topics or studies that would be your priority for research in LMICs.
5. Compare and contrast how the prevention of violence and unintentional injuries differ in respect to their use of passive versus active prevention strategies.

Acknowledgments

We thank Dr. Kacem Iaych, statistician at WHO, for checking the Global Health Estimate data.

Disclaimer

Drs. Butchart and Ms. Kieselbach are staff members of the World Health Organization. They alone are responsible for the views expressed in this publication, and their views do not necessarily represent the decisions or policies of the World Health Organization. Dr. Peden is a Senior Research Fellow and Head of the UK Injury Program at The George Institute for Global Health, Oxford University. Dr. Hyder is Director of the International Injury Research Unit at the Johns Hopkins Bloomberg School of Public Health.

References

Abramsky, T., Devries, K., Kiss, L., Nakuti, J., Kyegombe, N., Starmann, E., & Watts, C. (2014). Findings from the SASA! Study: A cluster randomized controlled trial to assess the impact of a community mobilization intervention to prevent violence against women and reduce HIV risk in Kampala, Uganda. *BMC Medicine, 12*, 122. https://doi.org/10.1186/s12916-014-0122-5

Ad Hoc Committee on Health Research Relating to Future Intervention Options. (1996). *Investing in health research and development* (Document TDK/ Gen/96.1). Geneva, Switzerland: World Health Organization.

Aeron-Thomas, A., Jacobs, G. D., Sexton, B., Gururaj, G., & Rahmann, F. (2004). *The involvement and impact of road crashes on the poor: Bangladesh and India case studies* (Published Project Report No. PRP 010). Crowthorne, UK: Transport Research Laboratory.

Afukaar, F. K., Antwi, P., & Ofosu-Amah, S. (2003). Pattern of road traffic injuries in Ghana: Implications for control. *Injury Control and Safety Promotion, 10*(1–2), 69–76.

Alfonso, Y. N., Alonge, O., Hoque, D. M. E., Baset, K. U., Hyder, A. A., & Bishai, D. (2017). Care-seeking patterns and direct economic burden of injuries in Bangladesh. *International Journal of Environmental Research and Public Health, 14*(5), 472.

Ambrose, A. F., Paul, G., & Hausdorff, J. M. (2013). Risk factors for falls among older adults: A review of the literature. *Maturitas, 75*, 51–61.

Aos, S., Lieb, J., Mayfield, J., Miller, M., & Pennucci, A. (2004). *Benefits and costs of prevention and early intervention programs for youth*. Olympia, WA: Washington State Institute for Public Policy.

Azizi, B. H., Zulkifli, H. I., & Kasim, M. S. (1993). Risk factors for accidental poisoning in urban Malaysian children. *Annals of Tropical Paediatrics, 13*(2), 183–188.

Bachani, A. M., Zia, N., Hung, Y. W., Adetunji, R., Cuong, P. V., Faried, A., . . . Hyder, A. A. (2017). Speeding in urban South East Asia: Results from a multi-site observational study. *Journal of the Australian College of Road Safety, 28*(2), 27.

Beets, M. W., Flay, B. R., Vuchinich, S., Snyder, F. J., Acock, A., Li, K.-K., . . . Durlak, J. (2009). Use of a social and character development program to prevent substance use, violent behaviors, and sexual activity among elementary-school students in Hawaii. *American Journal of Public Health, 99*(8), 1438–1445. https://doi.org/10.2105/AJPH.2008.142919

Beyer, F. R., & Ker, K. (2009). Street lighting for preventing road traffic injuries. *Cochrane Database of Systematic Reviews, 1*, CD004728.

Bierens, J. J. L. M. (ed). 2006. *Handbook on Drowning. Prevention, rescue, treatment*. Berlin, Germany, Springer-Verlag.

Bilukha, O., Hahn, R. A., Crosby, A., Fullilove, M. T., Liberman, A., Moscicki, E., . . . Task Force on Community Preventive Services. (2005). The effectiveness of early childhood home visitation in preventing violence: A systematic review. *American Journal of Preventive Medicine, 28*, 11–39.

Bishai, D., Asiimwe, B., Abbas, S., Hyder, A. A., & Bazeyo, W. (2008). Cost-effectiveness of traffic enforcement: Case study from Uganda. *Injury Prevention, 14*, 223–227.

Bishai, D., & Hyder, A. (2006). Modeling the cost effectiveness of injury interventions in lower and middle income countries: Opportunities and challenges. *Cost Effectiveness and Resource Allocation, 4*(1), 2.

Bishai, D., Hyder, A. A., Ghaffar, A., Morrow, R. H., & Kobusingye, O. (2003). Rates of public investment for road safety in developing countries: Case studies of Uganda and Pakistan. *Health Policy and Planning, 18*(2), 232–235.

Braga, A. A., Papachristos, A. V., & Hureau, D. M. (2014). The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly, 31*(4):633–663.

Brenner, R. A. (2003). Prevention of drowning in infants, children and adolescents. *Pediatrics, 112*(2), 440–445.

Building happy families: IRC research brief. (2014). *Impact evaluation of a parenting and family skills intervention for migrant and displaced Burmese families in Thailand*. New York, NY: International Rescue Committee. Retrieved from https://www.rescue.org/report/building-happy-families-irc-research-brief

Bunn, F., Collier, T., Frost, C., Ker, K., Roberts, I., & Wentz, R. (2003). Traffic calming for the prevention of road traffic injuries: Systematic review and meta-analysis. *Injury Prevention, 9*(3), 200–204.

Butchart, A., Brown., D., Khanh-Huynh, A., Corso, P., Florquin, N., & Muggah, R. (2008). *Manual for estimating the economic costs of injuries due to interpersonal and self-directed violence*. Geneva, Switzerland: World Health Organization.

Butchart, A., & Mikton, C. (2014). *Global status report on violence prevention*. Geneva, Switzerland: World Health Organization.

Butchart, A., Phinney, A., Check, P., & Villaveces, A. (2004). *Preventing violence: A guide to implementing the recommendations of the World Report on Violence and Health*. Geneva, Switzerland: Department of Injuries and Violence Prevention, World Health Organization.

Cameron, I. D., Gillespie, L. D., Robertson, M. C., Murray, G. R., Hill, K. D., Cumming, R. G., & Kerse, N. (2012). Interventions for preventing falls in older people in care facilities and hospitals. *Cochrane Library*. doi:10.1002/14651858.CD005465.pub3

Cameron, I. D., Robinovitch, S., Birge, S., Kannus, P., Khan, K., Lauritzen, J., . . . Kiel, D. P. (2010). Hip protectors: Recommendations for conducting clinical trials: An international consensus statement (part II). *Osteoporosis International, 21*(1), 1–10.

Cancian, M., Yang, M.-Y., & Slack, K. S. (2013). The effect of additional child support income on the risk of child maltreatment. *Social Service Review, 87*(3), 417–437. https://doi.org/10.1086/671929

Carlini-Cotrim, B., & da Matta Chasin, A.A. (2000). Blood alcohol content and death from fatal injury: A study in metropolitan area of Sao Paulo, Brazil. *Journal of Psychoactive Drugs, 32*(3), 269–275.

Carnegie Commission on Preventing Deadly Conflict. Preventing deadly conflict: final report. New York, NY: Carnegie Corporation, 1997.

Cassidy, T., Inglis, G., Wiysonge, C., & Matzopoulos, R. (2014). A systematic review of the effects of poverty deconcentration and urban upgrading on youth violence. *Health & Place, 26*, 78–87. https://doi.org/10.1016/j.healthplace.2013.12.009

Cerdá, M., Morenoff, J. D., Hansen, B. B., Tessari Hicks, K. J., Duque, L. F., Restrepo, A., & Diez-Roux, A. V. (2012). Reducing violence by transforming neighborhoods: A natural experiment in Medellín, Colombia. *American Journal of Epidemiology, 175*(10), 1045–1053. https://doi.org/10.1093/aje/kwr428

Chaux, E. (2007). Aulas en Paz: A multicomponent program for the promotion of peaceful relationships and citizenship competencies. *Conflict Resolution Quarterly, 25*(1), 79–86. https://doi.org/10.1002/crq.193

Chew, F. L., Yong, C. K., Mas Ayu, S., & Tajunisah, I. (2010). The association between various visual function tests and low fragility hip fractures among the elderly: A Malaysian experience. *Age and Ageing, 39*(2), 239–245.

Cluver, L. D., Orkin, F. M., Boyes, M. E., & Sherr, L. (2014). Cash plus care: social protection cumulatively mitigates HIV-risk behaviour among adolescents in South Africa. *Aids, 28*, S389–S397.

Coutinho, E. S., Bloch, K. G., & Rodrigues, L. C. (2009). Characteristics and circumstances of falls leading to severe fractures in elderly people in Rio de Janeiro, Brazil. *Cadernos de Saúde Pública, 25*(2), 455–459.

Craig, W., Harel-Fisch, Y., Fogel-Grinvald, H., Dostaler, S., Hetland, J., Simons-Morton, B., . . . HBSC Bullying Writing Group. (2009). A cross-national profile of bullying and victimization among adolescents in 40 countries. *International Journal of Public Health, 54*(2 suppl), 216–224.

Dahdah, S., & McMahon, K. (2008). *The true cost of road crashes: valuing life and the cost of a serious injury*. Washington, DC: International Road Assessment Programme, World Bank Global Road Safety Facility.

Dandona, R., Kumar, G. A., Ivers, R., Joshi, R., Neal, B., & Dandona, L. (2010). Characteristics of non-fatal fall injuries in rural India. *Injury Prevention, 16*, 166–171.

Danseco, E. R., Miller, T. R., & Spicer, R. S. (2000). Incidence and costs of 1987–1994 childhood injuries: Demographic breakdowns. *Pediatrics, 105*(2), E27.

DeSimone, J., Markowitz, S., & Xu, J. (2013). Child access prevention laws and nonfatal gun injuries. *Southern Economic Journal, 80*(1), 5–25. https://doi.org/10.4284/0038-4038-2011.333

Devries, K. M., Knight, L., Child, J. C., Mirembe, A., Nakuti, J., Jones, R., . . . Naker, D. (2015). The Good School Toolkit for reducing physical violence from school staff to primary school students: A cluster-randomised controlled trial in Uganda. *The Lancet: Global Health, 3*(7), e378–e386. https://doi.org/10.1016/S2214-109X(15)00060-1

DiGuiseppi, C., Jacobs, D. E., Phelan, K. J., Mickalide, A., & Ormandy, D. (2010). Housing interventions and control of injury-related structural deficiencies: A review of the evidence. *Journal of Public Health Management & Practice, 16*(3 suppl), S35–S43.

Dinh-Zarr, T. B., Sleet, D. A., Shults, R. A., Zaza, S., Elder, R. W., Nichols, J. L., . . . Task Force on Community Preventive Services. (2001). Reviews of evidence regarding interventions to increase the use of safety belts. *American Journal of Preventive Medicine, 21*(4 suppl), 48–65.

Donroe, J., Tincopa, M., Gilman, R. H., Brugge, D., & Moore, D. A. (2008). Pedestrian road traffic injuries in urban Peruvian children and adolescents: Case control analyses of personal and environmental risk factors. *PLoS One, 3*(9), e3166.

Driscoll, T. R., Harrison, J. A., & Steenkamp, M. (2004). Review of the role of alcohol in drowning associated with recreational aquatic activity. *Injury Prevention, 10*(2), 107–113.

Dubowitz, H., Feigelman, S., Lane, W., & Kim, J. (2009). Pediatric primary care to help prevent child maltreatment: The Safe Environment for Every Kid (SEEK) Model. *Pediatrics, 123*(3), 858–864. https://doi.org/10.1542/peds.2008-1376

Dworkin, S. L., Hatcher, A. M., Colvin, C., & Peacock, D. (2013). Impact of a gender-transformative HIV and antiviolence program on gender ideologies and masculinities in two rural, South African communities. *Men and Masculinities, 16*(2). https://doi.org/10.1177/1097184X12469878

Elvik, R., & Vaa, T. (2004). *Handbook of road safety measures*. Amsterdam, Netherlands: Elsevier.

European Transport Safety Council. (2016). How traffic law enforcement can contribute to safer roads. PIN Flash Report 31. Retrieved from http://etsc.eu/wp-content/uploads/PIN_FLASH31_Final-1.pdf

Euser, S., Alink, L. R., Stoeltenborgh, M., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2015). A gloomy picture: A meta-analysis of randomized controlled trials reveals disappointing effectiveness of programs aiming at preventing child maltreatment. *BMC Public Health, 15*, 1068. https://bmcpublichealth.biomedcentral.com/track/pdf/10.1186/s12889-015-2387-9.

Fang, X., Brown, D. S., Florence, C. S., & Mercy, J. A. (2012). The economic burden of child maltreatment in the United States and implications for prevention. *Child abuse & neglect, 36*(2), 156–165.

Fang, X., Fry, D. A., Brown, D. S., Mercy, J. A., Dunne, M. P., Butchart, A. R., . . . & McCoy, A. (2015). The burden of child maltreatment in the East Asia and Pacific region. *Child abuse & neglect, 42*, 146-162.

FIA Foundation for the Automobile and Society. (2009). *Seat-belts and child restraints: A road safety manual for decision-makers and practitioners*. London, UK: Author.

Fitterer, J. L., Nelson, T. A., & Stockwell, T. (2015). A review of existing studies reporting the negative effects of alcohol access and positive effects of alcohol control policies on interpersonal violence. *Frontiers in Public Health, 3*, 253. https://doi.org/10.3389/fpubh.2015.00253

Florence, C., Shepherd, J., Brennan, I., & Simon, T. (2011). Effectiveness of anonymised information sharing and use in health service, police, and local government partnership for preventing violence related injury: Experimental study and time series analysis. *BMJ (Clinical Research Edition), 342*, d3313.

Florence, C., Shepherd, J., Brennan, I., & Simon, T. R. (2014). An economic evaluation of anonymised information sharing in a partnership between health services, police and local government for preventing violence-related injury. *Injury Prevention, 20*(2), 108–114. https://doi.org/10.1136/injuryprev-2012-040622

Foshee, V. A., Bauman, K. E., Ennett, S. T., Suchindran, C., Benefield, T., & Linder, G. F. (2005). Assessing the effects of the dating violence prevention program "Safe Dates" using random co-efficient regression modelling. *Prevention Science, 6*, 245–257.

García-Moreno, C., Jansen, H., Ellsberg, M., Heise, L., & Watts, C. (2005). *WHO multi-country study on women's health and domestic violence against women*. Geneva, Switzerland: World Health Organization.

García-Moreno, C., Pallitto, C., Devries, K., Stöckl, H., Watts, C., & Abrahams, N. (2013). Global and regional estimates of violence against women: Prevalence and health effects of intimate partner violence and non-partner sexual violence: World Health Organization.

GBD 2015 DALYs and HALE Collaborators. (2016). Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388*, 1603–1658.

Gillespie, L. D., Robertson, M. C., Gillespie, W. J., Sherrington, C., Gates, S., Clemson, L. M., & Lamb, S. E. (2012). Interventions for preventing falls in older people living in the community. *Cochrane Database of Systematic Reviews, 9*, CD007146. doi:10.1002/14651858. CD007146.pub3

Global Road Safety Partnership. (2007). Drinking and driving: An international good practice manual. Retrieved from http://www.who.int/roadsafety/projects/manuals/alcohol/en/

Global Road Safety Partnership. (2008). Speed management: A road safety manual for decision-makers and practitioners. Retrieved from http://www.who.int/roadsafety/projects/manuals/speed_manual/en/

Gupta, J., Falb, K. L., Lehmann, H., Kpebo, D., Xuan, Z., Hossain, M., . . . Annan, J. (2013). Gender norms and economic empowerment intervention to reduce intimate partner violence against women in rural Côte d'Ivoire: A randomized controlled pilot study. *BMC International Health and Human Rights, 13*, 46. https://doi.org/10.1186/1472-698X-13-46

Haddon, W., Jr. (1968). The changing approach to the epidemiology, prevention, and amelioration of trauma: The transition to approaches etiologically rather than descriptively based. *American Journal of Public Health and the Nation's Health, 58*(8), 1431–1438.

Hahn, R., Fuqua-Whitley, D., Wethington, H., Lowy, J., Liberman, A., Crosby, A., . . . Task Force on Community Preventive Services. (2007). The effectiveness of universal school-based programs for the prevention of violent and aggressive behavior: A report on recommendations of the Task Force on Community Preventive Services. *Morbidity and Mortality Weekly Report: Recommendations and Reports, 56*, 1–12.

Haslam, R. A., Hidea, S. A., Gibb, A. G. F., Gyi, D. E., Pavitt, T., Atkinson, S., & Duff, A. R. (2005). Contributing factors in construction accidents. *Applied Ergonomics, 36*, 401–415.

Heise, L., & García-Moreno, C. (2002). Violence by intimate partners. In E. Krug, L. L. Dahlberg, J. A. Mercy, A. B. Zwi, & R. Lozano (Eds.), *World report on violence and health* (pp. 87–121). Geneva, Switzerland: World Health Organization.

Henry, D., Knoblauch, S., & Sigurvinsdottir, R. (2014). *The effect of intensive ceasefire intervention on crime in four Chicago police beats: Quantitative assessment*. Chicago, IL: Robert R. McCormick Foundation.

Hippisley-Cox, J., & Coupland, C. (2009). Predicting risk of osteoporotic fracture in men and women in England and Wales: Prospective derivation and validation of QFractureScores. *British Medical Journal, 339*, b4229.

Huston, A., Miller, C., Richburg-Hayes, L., Duncan, G., Eldred, C., Weisner, T., . . . Redcross, C. (2003). *New hope for families and children: Five year results of a program to reduce poverty and reform welfare*. New York, NY: Manpower Demonstration Research Corporation.

Hyder, A. A., Allen, K. A., Peters, D. H., Chandran, A., & Bishai, D. (2013). Large-scale road safety programmes in low- and middle-income countries: An opportunity to generate evidence. *Global Public Health, 8*(5), 504–518.

Hyder, A. A., Alonge, O., He, S., Wadhwaniya, S., Rahman, F., Rahman, A., & Arifeen, S. E. (2014a). Saving of children's lives from drowning project in Bangladesh. *American Journal of Preventive Medicine*, *47*(6), 842–845.

Hyder, A. A., Alonge, O., He, S., Wadhwaniya, S., Rahman, F., Rahman, A., & Arifeen, S. E. (2014b). A framework for addressing implementation gap in global drowning prevention interventions: Experiences from Bangladesh. *Journal of Health, Population, and Nutrition, 32*, 564–576.

Hyder, A. A., Arifeen, S., Begum, N., Fishman, S., Wali, S., & Baqui, A. H. (2003). Death from drowning: Defining a new challenge for child survival in Bangladesh. *Injury Control and Safety Promotion, 10*(4), 205–210.

Hyder, A. A., Waters, H., Philipps, T., & Rehwinkel, J. (2004). *Exploring the economics of motorcycle helmet laws: Implications for low and middle-income countries*. Baltimore, MD: Johns Hopkins University.

ICF Consulting & Imperial College Centre for Transport Studies. (2003). *Cost–benefit analysis of road safety improvements: Final report*. London, UK: ICF Consulting.

Institute of Medicine. (2011). *Preventing violence against women and children: Workshop summary*. Washington, DC: National Academies Press.

International Road Assessment Programme. (2015). *Vaccines for roads (third edition)*. Retrieved from http://www.irap.org/phocadownload/Vaccines_for_Roads_3.pdf

Jacobs, G., Aeron-Thomas, A., & Astrop, A. (2000). *Estimating global road fatalities* (TRL Report 445). Crowthorne, UK: Transport Research Laboratory.

Jan, S., Ferrari, G., Watts, C. H., Hargreaves, J. R., Kim, J. C., Phetla, G., . . . Pronyk, P. M. (2011). Economic evaluation of a combined microfinance and gender training intervention for the prevention of intimate partner violence in rural South Africa. *Health Policy and Planning*, *26*(5), 366–372. https://doi.org/10.1093/heapol/czq071

Jaye, C., Simpson, J. C., & Langley, J. D. (2001). Barriers to safe hot tap water: Results from a national study of New Zealand plumbers. *Injury Prevention, 7*, 302–306.

Jewkes, R., Nduna, M., Levin, J., Jama, N., Dunkle, K., Puren, A., & Duvvury, N. (2008). Impact of stepping stones on incidence of HIV and HSV-2 and sexual behaviour in rural South Africa: Cluster randomised controlled trial. *British Medical Journal, 337*, a506. doi:10.1136/bmj.a506

Jiang, X., Zhang, Y., Wang, Y., Wang, B., Xu, Y., & Shang, L. (2010). An analysis of 6215 hospitalized unintentional injuries among children aged 0–14 in northwest China. *Accident Analysis & Prevention, 42*(1), 320–326.

Jitapunkul, S., Yuktananandana, P., & Parkpian, V. (2001). Risk factors of hip fracture among Thai female patients. *Journal of the Medical Association of Thailand, 84*(11), 1576–81.

Kärnä, A., Voeten, M., Little, T. D., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2011). A large-scale evaluation of the KiVa antibullying program: Grades 4–6. *Child Development, 82*(1), 311–330. https://doi.org/10.1111/j.1467-8624.2010.01557.x

Kendrick, D., Smith, S., Sutton, A., Watson, M., Coupland, C., Mulvaney, C., & Mason-Jones, A. (2008). Effect of education and safety equipment on poisoning-prevention practices and poisoning: Systematic review, meta-analysis and metaregression. *Archives of Disease in Childhood, 93*(7), 599–608.

Kendrick, D., Watson, M. C., Mulvaney, C. A., Smith, S. J., Sutton, A. J., Coupland, C. A., & Mason-Jones, A. J. (2008). Preventing childhood falls at home: Meta-analysis and meta regression. *American Journal of Preventive Medicine, 35*(4), 370–379.

Kendrick, D., Young, B., Mason-Jones, A. J., Ilyas, N., Achana, F. A., Cooper, N. J., . . . Coupland, C. (2013). Home safety education and provision of safety equipment for injury prevention [Review]. *Evidence-Based Child Health: A Cochrane Review Journal, 8*, 761–939.

Khan, M. M., & Reza, H. (1998). Gender differences in non-fatal suicidal behavior in Pakistan: Significance of sociocultural factors. *Suicide and Life-Threatening Behavior, 28,* 62–68.

Kilburn, R., & Karoly, L. A. (2008). *The economics of early childhood policy: What the dismal science has to say about investing in children* (OP-227-CFP). Santa Monica, CA: RAND Corporation. Retrieved from http://www.rand.org/pubs/occasional_papers/ OP227.html

Kim, J. C., Watts, C. H., Hargreaves, J. R., Ndhlovu, L. X., Phetla, G., Morison, L. A., . . . Pronyk, P. (2007). Understanding the impact of a microfinance-based intervention on women's empowerment and the reduction of intimate partner violence in South Africa. *American Journal of Public Health, 97*(10), 1794–1802.

Kloeden, C. N., McLean, A. J., Baldock, M. R. J., & Cockington, A. J. T. (1998). *Severe and fatal car crashes due to roadside hazards: A report to the Motor Accident Commission*. Adelaide, Australia: National Health and Medical Research Council Road Accident Research Unit, University of Adelaide.

Knerr, W., Gardner, F., & Cluver, L. (2013). Improving positive parenting skills and reducing harsh and abusive parenting in low- and middle-income countries: A systematic review. *Prevention Science, 14*(4), 352–363. https://doi.org/10.1007/s11121-012-0314-1

Knox, M., & Burkhart, K. (2014). A multi-site study of the ACT Raising Safe Kids program: Predictors of outcomes and attrition. *Children and Youth Services Review, 39*, 20–24. https://doi.org/10.1016/j.childyouth.2014.01.006

Kobusingye, O., Guwatudde, D., & Lett, R. (2001). Injury patterns in rural and urban Uganda. *Injury Prevention, 7*(1), 46–50.

Koehler, J. A., Lösel, F., Akoensi, T. D., & Humphreys, D. K. (2013). A systematic review and meta-analysis on the effects of young offender treatment programs in Europe. *Journal of Experimental Criminology, 9*(1), 19–43. https://doi.org/10.1007/s11292-012-9159-7

Kopits, E., & Cropper, M. (2003). *Traffic fatalities and economic growth* (Policy Research Working Paper 3035). Washington, DC: World Bank.

Krug, A., Ellis, J. B., Hay, I. T., Mokgabudi, N. F., & Robertson, J. (1994). The impact of child-resistant containers on the incidence of paraffin (kerosene) ingestion in children. *South African Medical Journal, 84*(11), 730–734.

Krug, E., Dahlberg, L. L., Mercy, J. A., Zwi, A. B., & Lozano R. (Eds.). (2002). *World report on violence and health*. Geneva, Switzerland: World Health Organization.

Kyegombe, N., Abramsky, T., Devries, K. M., Michau, L., Nakuti, J., Starmann, E., . . . Watts, C. (2015). What is the potential for interventions designed to prevent violence against women to reduce children's exposure to violence? Findings from the SASA! study, Kampala, Uganda. *Child Abuse & Neglect, 50*, 128–140. https://doi.org/10.1016/j.chiabu.2015.10.003

Liu, B. C., Ivers, R., Norton, R., Boufous, S., Blows, S., & Lo, S. K. (2008). Helmets for preventing injury in motorcycle riders. *Cochrane Database of Systematic Reviews, 1*, CD004333.

Liu, G. F., Han, S., Liang, D. H., Wang, F. Z., Shi, X. Z., Yu, J., & Wu, Z. L. (2003). Driver sleepiness and risk of car crashes in Shenyang, a Chinese northeastern city: Population-based case-control study. *Biomedical and Environmental Sciences, 16*(3), 219–226.

Lord, S. R., Sherrington, C., Menz, H. B., & Close, J. C. T. (2007). *Falls in older people: Risk factors and strategies for prevention* (2nd ed.). Cambridge, UK: Cambridge University Press.

Lundgren, R., Beckman, M., Chaurasiya, S. P., Subhedi, B., & Kerner, B. (2013). Whose turn to do the dishes? Transforming gender attitudes and behaviours among very young adolescents in Nepal. *Gender & Development, 21*(1), 127–145. https://doi.org/10.1080/13552074.2013.767520

Macarthur, C. (2003). Evaluation of Safe Kids Week 2001: Prevention of scald and burn injuries in young children. *Injury Prevention, 9*(2), 112–116.

MacMillan, H., & Wathen, C. (2014). *Research brief: Interventions to prevent child maltreatment*. London, Ontario, Canada: Preventing Violence Across the Lifespan Research Network.

Matzopoulos, R. G., Thompson, M. L., & Myers, J. E. (2014). Firearm and nonfirearm homicide in 5 South African cities: A retrospective population-based study. *American Journal of Public Health, 104*(3), 455–460. https://doi.org/10.2105/AJPH.2013.310650

McGinty, E. E., Tung, G., Shulman-Laniel, J., Hardy, R., Rutkow, L., Frattaroli, S., & Vernick, J. S. (2017). Ignition interlock laws: Effects on fatal motor vehicle crashes, 1982–2013. *American Journal of Preventive Medicine, 52*(4), 417–423.

Mercy, J. A., Butchart, A., Farrington, D., & Cerda, M. (2002). Youth violence. In E. Krug, L. L. Dahlberg, J. A. Mercy, A. B. Zwi, & R. Lozano (Eds.), *World report on violence and health* (pp. 23–56). Geneva, Switzerland: World Health Organization.

Mercy, J. A., Butchart, A., Rosenberg, M. L., Dahlberg, L., & Harvey, A. (2008). Preventing violence in developing countries: A framework for action. *International Journal of Injury Control and Safety Promotion, 15*(4), 197–208.

Mock, C. J., Jurkovich, G. J., nii-Amon-Kotei, D., Arreola-Risa, C., & Maier, R. V. (1998). Trauma mortality patterns in three nations at different economic levels: Implications for global trauma system development. *Journal of Trauma, 44*(5), 804–812.

Mock, C., Lormand, J. P., Joshipura, M., Goosen, J., & Peden, M. (2004). *Guidelines for essential trauma care*. Geneva, Switzerland: World Health Organization.

Mock, C., Quansah, R., Krishnan, R., Arreola-Risa, C., & Rivara, F. (2004). Strengthening the prevention and care of injuries worldwide. *The Lancet, 363*(9427), 2172–2179.

Modig C. (2009). Never violence: 30 years on from Sweden's abolition of corporal punishment. Government Offices of Sweden & Save the Children Sweden. Retrieved from http://www.government.se/contentassets/6bfb214c582448b6ace4d32978361577/never-violence---thirty-years-on-from-swedens-abolition-of-corporal-punishment

Mohan, D. (2002). Road safety in less-motorized environments: Future concerns. *International Journal of Epidemiology, 31*(3), 527–532.

Murray, L. K., Skavenski, S., Kane, J. C., Mayeya, J., Dorsey, S., Cohen, J. A., . . . Bolton, P. A. (2015). Effectiveness of trauma-focused cognitive behavioral therapy among trauma-affected children in Lusaka, Zambia: A randomized clinical trial. *JAMA Pediatrics, 169*(8), 761–769. https://doi.org/10.1001/jamapediatrics.2015.0580

Nafukho, F. M., & Khayesi, M. (2002). Livelihood, conditions of work, regulation, and road safety in the small-scale public transport sector: A case of the matatu mode of transport in Kenya. In X. Godard & I. Fatouzoun (Eds.), *Urban mobility for all: Proceedings of the Tenth International CODATU Conference, Lome, Togo, 12–15 November 2002* (pp. 241–425). Lisse, Netherlands: AA Balkema.

Nhachi, C. F., & Kasilo, O. M. (1994). Household chemicals poisoning admissions in Zimbabwe's main urban centres. *Human and Experimental Toxicology, 13*(2), 69–72.

Nixon, J., Spinks, A., Turner, C., & McClure, R. (2004). Community based programs to prevent poisoning in children 0–15 years. *Injury Prevention, 10*(1), 43–46.

Norton, R., Hyder, A., Bishai, D., & Peden, M. (2006). Unintentional injuries. In D. Jamison, J. Breman, A. Measham, G. Alleyne, M. Claeson, & D. Evans (Eds.), *Disease control priorities in developing countries* (pp. 737–753). New York, NY: Oxford University Press.

Nurse Family Partnership. (2011). Published research: Nurse Family Partnership. Retrieved from http://www.nursefamilypartnership.org/Proven-Results/Published-research

Odero, W. & Zwi, A. B. (1997). Drinking and driving in an urban setting in Kenya. *East Afr Med J., 74*(11):675-679.

Odero, W., Garner, P., & Zwi, A. B. (1997). Road traffic injuries in developing countries: A comprehensive review of epidemiological studies. *Tropical Medicine and International Health, 2*(5), 445–460.

Olds, D., Henderson, C. R., Jr., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., . . . Powers, J. (1998). Long-term effects of nurse home visitation on children's criminal and antisocial behavior: 15-year follow-up of a randomized controlled trial. *Journal of the American Medical Association, 280*(14), 1238–1244.

Organisation for Economic Co-operation and Development (OECD). (2008). Towards zero: Ambitious road safety targets and the safe system approach.

Organisation for Economic Co-operation and Development (OECD). (2015). Improving safety for motorcycle, scooter and moped riders. Retrieved from http://www.oecd-ilibrary.org/transport/improving-safety-for-motorcycle-scooter-and-moped-riders_9789282107942-en

Organization of Motor Vehicle Manufacturers. (2016). World motor vehicle production by country and type. OICA correspondents survey, 2016. Retrieved from http://www.oica.net/category/production-statistics/

Ozer, E. J., Fernald, L. C. H., Manley, J. G., & Gertler, P. J. (2009). Effects of a conditional cash transfer program on children's behavior problems. *Pediatrics, 123*(4), e630–e637. https://doi.org/10.1542/peds.2008-2882

Paine, K., Hart, G., Jawo, M., Ceesay, S., Jallow, M., Morison, L., . . . Shaw, M. (2002). "Before we were sleeping, now we are awake": Preliminary evaluation of the Stepping Stones sexual health programme in the Gambia. *African Journal of AIDS Research, 1*, 41–52.

Parents make the difference. (2014). *Findings from a randomized impact evaluation of a parenting program in rural Liberia*. New York, NY: International Rescue Committee.

Parker, M. J., Gillespie, W. J., & Gillespie, L. D. (2006). Effectiveness of hip protectors for preventing hip fractures in elderly people: Systematic review. *British Medical Journal, 332*(7541), 571–574.

Peden, M., Oyebite, K., Ozanne-Smith, J., Hyder, A. A., Branche, C., Rahman, F. A. K. M., . . . Bartolomeos, K. (Eds.). (2008). *World report on child injury prevention*. Geneva, Switzerland: World Health Organization.

Peden, M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A. A., Jarawan, E., & Mathers, C. (Eds.). (2004). *World report on road traffic injury prevention*. Geneva, Switzerland: World Health Organization.

Picard-Fritsche, S., & Cerniglia, L. (2013). *Testing a public health approach to gun violence*. New York, NY: Center for Court Innovation. Retrieved from http://cureviolence.org/post/resource/testing-a-public-health-approach-to-gun-violence/

Poli de Figueiredo, L. F., Rasslan, S., Bruscagin, V., Cruz, R., & Rocha e Silva, M. (2001). Increases in fines and driver licence withdrawal have effectively reduced immediate deaths from trauma on Brazilian roads: First-year report on the new traffic code. *Injury, 32*(2), 91–94.

Prinja, S., Jagnoor, J., Chauhan, A. S., Aggarwal, S., & Ivers, R. (2015). Estimation of the economic burden of injury in north India: A prospective cohort study. High private out-of-pocket expenditure for treatment of injury poses major economic burden on families. *The Lancet, 385*(2), S27–S57.

Pronyk, P. M., Hargreaves, J. R., Kim, J. C., Morison, L. A., Phetla, G., Watts, C., . . . Porter, J. D. (2006). Effect of a structural intervention for the prevention of intimate-partner violence and HIV in rural South Africa: A cluster randomised trial. *The Lancet, 368*(9551), 1973–1983.

Pronyk, P. M., Hargreaves, J. R., & Morduch, J. (2007). Microfinance programs and better health: Prospects for sub-Saharan Africa. *Journal of the American Medical Association, 298*(16), 1925–1927. https://doi.org/10.1001/jama.298.16.1925

Quan, L., Bennett, E. E., & Branche, C. M. (2007). Interventions to prevent drowning. In L. S. Doll, S. E. Bonzo, D. A. Sleet, & J. A. Mercy (Eds.), *Handbook of injury and violence prevention*. Atlanta, GA: Centers for Disease Control and Prevention, pg 81-96.

Radin Umar, R. S., Mackay, G. M., & Hills, B. L. (1996). Modelling of conspicuity-related motor cycle accidents in Seremban and Shah Alam, Malaysia. *Accident Analysis and Prevention, 28*(3), 325–332.

Rahman, F., Bose, S., Linnan, M., Rahman, A., Mashreky, S.R., Haaland, B., & Finkelstein, E. (2012). Cost-effectiveness of an injury and drowning prevention program in Bangladesh. *Pediatrics, 130*(6), e1621–e1628.

Raising Voices. (2016). Raising Voices: Preventing violence against women and children. Retrieved from http://raisingvoices.org/

Razzak, J. A., & Kellermann, A. L. (2002). Emergency medical care in developing countries: Is it worthwhile? *Bulletin of the World Health Organization, 80*(11), 900–905.

Retting, R. A., Ferguson, S. A., & McCartt, A. T. (2003). A review of evidence-based traffic engineering measures designed to reduce pedestrian–motor vehicle crashes. *American Journal of Public Health, 93*(9), 1456–1463.

Reynolds, A. J., Temple, J. A., Ou, S. R., Arteaga, I. A., & White, B. A. (2011). School-based early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. *Science, 333*(6040), 360–364.

Rivara, F. P., Thompson, D. C., Beahler, C., & MacKenzie, E. J. (1999). Systematic reviews of strategies to prevent motor vehicle injuries. *American Journal of Preventive Medicine, 16*(1 suppl), 1–5.

Road Traffic Injuries Research Network Multicenter Study Collaborators. (2013). The use of non-standard motorcycle helmets in low- and middle-income countries: a multicenter study. *Injury Prevention, 19*(3), 158–163.

Robb, G., Sultana, S., Ameratunga, S., & Jackson, R. (2008). A systematic review of epidemiological studies investigating risk factors for work-related road traffic crashes and injuries. *Injury Prevention, 14*(1), 51–58.

Robbins, J., Aragaki, A. K., Kooperberg, C., Watts, N., Wastawski-Wende, J., Jaskson, R. D., . . . Cauley, J. (2007). Factors associated with 5-year risk of hip fracture in postmenopausal women. *Journal of the American Medical Association, 298*(20), 2389–2398.

Rosenberg, M., Butchart, A., Mercy, J., Narasimhan, V., & Waters, H. (2006). Interpersonal violence. In D. T. Jamison, J. G. Breman, A. R. Measham, G. Alleyne, M. Claeson, D. B. Evans, . . . Musgrove, P. (Eds.), *Disease control priorities in developing countries* (2nd ed., pp. 755–770). Washington, DC: Oxford University Press & World Bank.

Salmivalli, C., & Poskiparta, E. (2012). KiVa antibullying program: Overview of evaluation studies based on a randomized controlled trial and national rollout in Finland. *International Journal of Conflict and Violence, 6*(2), 294–301.

Santaella-Tenorio, J., Cerdá, M., Villaveces, A., & Galea, S. (2016). What do we know about the association between firearm legislation and firearm-related injuries? *Epidemiologic Reviews, 38*(1), 140–157. https://doi.org/10.1093/epirev/mxv012

Shannon, M. (2000). Ingestion of toxic substances by children. *New England Journal of Medicine, 342*(3), 186–191.

Sheehan, P., Sweeny, K., Rasmussen, B., Wils, A., Friedman, H. S., Mahon, J., . . . Laski, L. (2017). Building the foundations for sustainable development: A case for global investment in the capabilities of adolescents. *The Lancet.* [Epub April 19, 2017]. http://dx.doi.org/10.1016/S0140-6736(17)30872-3

Shek, D. T. L., & Ma, C. M. S. (2012). Impact of Project P.A.T.H.S. on adolescent developmental outcomes in Hong Kong: Findings based on seven waves of data. *International Journal of Adolescent Medicine and Health, 24*(3), 231–244. https://doi.org/10.1515/ijamh.2012.034

Skogan, W., Harnett, S., Bump, N., & DuBois, J. (2009). *An evaluation of CeaseFire-Chicago*. Chicago, IL: Northwestern University Institute for Policy Research.

Sobngwi-Tambekou, J., Bhatti, J., Kounga, G., Salmi, L. R., & Lagarde, E. (2010). Road traffic crashes on the Yaoundé-Douala road section, Cameroon. *Accident Analysis & Prevention, 42*(2), 422–426.

Solomon, M. G., Ulmer, R. G., & Preusser, D. F. (2002). *Evaluation of click it or ticket model programs* (DOT HS-809-498). Washington, DC: National Highway Traffic Safety Administration.

Soori, H. (2001). Developmental risk factors for unintentional childhood poisoning. *Saudi Medical Journal, 22*(3), 227–230.

Spangaro, J., Adogu, C., Ranmuthugala, G., Davies, G. P., Steinacker, L., & Zwi, A. (2013). What evidence exists for initiatives to reduce risk and incidence of sexual violence in armed conflict and other humanitarian crises? A systematic review. *PLoS One, 8*(5), e62600.

Staton, C., Vissoci, J., Gong, E., Toomey, N., Wafula, R., . . . Hocker, M.D. (2016). Road traffic injury prevention initiatives: A systematic review and metasummary of effectiveness in low and middle income countries. *PLOS One*. doi:10.1371/journal.pone.0144971

Stoltenborgh, M., van Ijzendoorn, M. H., Euser, E. M., & Bakermans-Kranenburg, M. J. (2011). A Global Perspective

on Child Sexual Abuse: Meta-Analysis of Prevalence Around the World. *Child Maltreatment, 16*(2), 79–101. doi:10.1177/1077559511403920.

Sumner, S. A., Mercy, A. A., Saul, J., Motsa-Nzuza, N., Kwesigabo, G., Buluma, R., . . . & Kilbane, T. (2015). Prevalence of sexual violence against children and use of social services-seven countries, 2007-2013. *MMWR Morbidity and Mortality Weekly Report, 64*(21), 565–569.

Thanh, N. X., Hang, H. M., Chuc, N. T., & Lindholm, L. (2003). The economic burden of unintentional injuries: A community-based cost analysis in Bavi, Vietnam. *Scandinavian Journal of Public Health Supplement, 62*, 45–51.

Thompson, D. C., Rivara, F. P., & Thompson, R. (2003). Helmets for preventing head and facial injuries in bicyclists. *Annals of Emergency Medicine, 41*(5), 738–740.

Transport for London. (2016). Casualties in Greater London during 2015. Retrieved from http://content.tfl.gov.uk/casualties-in-greater-london-2015.pdf

Turner, C., Spinks, A., McClure, R., & Nixon, J. (2004) Community-based interventions for the prevention of burns and scalds in children. *Cochrane Database Systematic Reviews, 3*, CD004335.

United Nations. (2014). UN Resolution A/RES/69/194 Part one (1) Ensuring the prohibition by law of all forms of violence against children.

United Nations Road Safety Collaboration. (2011). *Global plan for the Decade of Action for Road Safety 2011–2020*. Geneva, Switzerland: World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/10665/255196/1/9789241511933-eng.pdf?ua=1

Usdin, S., Scheepers, E., Goldstein, S., & Japhet, G. (2005). Achieving social change on gender-based violence: A report on the impact evaluation of Soul City's fourth series. *Social Science and Medicine, 61*, 2434–2445.

Verma, R., Pulerwitz, J., Mahendra, V., Khandekar, S., Singh, A., Das, S., . . . Barker, G. (2008). *Promoting gender equity as a strategy to reduce HIV risk and gender-based violence among young men in India. Horizons Final Report*. Washington, DC: Population Council.

Vyas, S., & Watts, C. (2009). How does economic empowerment affect women's risk of intimate partner violence in low and middle income countries? A systematic review of published evidence. *Journal of International Development, 21*(5), 577–602. https://doi.org/10.1002/jid.1500

Wagenaar, A. C., Toomey, T. L., & Erickson, D. J. (2005). Complying with the minimum drinking age: Effects of enforcement and training interventions. *Alcoholism, Clinical and Experimental Research, 29*(2), 255–262.

Wallace, T. (2006). *Evaluating Stepping Stones: A review of existing evaluations and ideas for future M&E work*. Johannesburg, South Africa: ActionAid International. Retrieved from http://www.steppingstonesfeedback.org/resources/5/CountriesfromSTEPPINGSTONESREVIEWOFREVIEWS2006Wallace.pdf

Wallbank, C., McRae-McKee, K., Durrell, L., & Hynd, D. (n.d.). *The potential for vehicle safety standards to prevention deaths and injuries in Latin America: An assessment of the societal and economic impact of inaction*. Wokingham, UK: Transport Research Laboratory. Retrieved from https://trl.co.uk/reports/potential-vehicle-safety-standards-prevent-deaths-and-injuries-latin-america

Warda, L., Tenenbein, M., & Moffatt, M. E. K. (1999). House fire injury prevention update. Part I: A review of risk factors for fatal and non-fatal house fire injury. *Injury Prevention, 5*(3), 145–150.

Washburn, I. J., Acock, A., Vuchinich, S., Snyder, F., Li, K.-K., Ji, P., . . . Flay, B. R. (2011). Effects of a social–emotional and character development program on the trajectory of behaviors associated with social–emotional and character development: Findings from three randomized trials. *Prevention Science, 12*(3), 314–323. https://doi.org/10.1007/s11121-011-0230-9

Waters, H. R., Hyder, A. A., & Phillips, T. L. (2004). Economic evaluation of interventions for reducing road traffic injuries: A review of literature with applications to low and middle-income countries. *Asia Pacific Journal of Public Health, 16*(1), 23–31.

Waters, H., Hyder, A., Rajkotia, Y., Basu, S., Rehwinkel, J. A., & Butchart, A. (2004). *The economic dimensions of interpersonal violence*. Geneva, Switzerland: World Health Organization.

Webster, D., Whitehill, J., Vernick, J., & Parker, E. (2012). *Evaluation of Baltimore's Safe Streets Program: Effects on attitudes, participants' experiences, and gun violence*. Baltimore, MD: Johns Hopkins Center for the Prevention of Youth Violence.

Wechsler, H., & Nelson, T. F. (2010). Will increasing alcohol availability by lowering the minimum legal drinking age decrease drinking and related consequences among youths? *American Journal of Public Health, 100*(6), 986–992. https://doi.org/10.2105/AJPH.2009.178004

Wethington, H. R., Hahn, R. A., Fuqua-Whitley, D. S., Sipe, T. A., Crosby, A. E., Johnson, R. L., . . . Task Force on Community Preventive Services. (2008). The effectiveness of interventions to reduce psychological harm from traumatic events among children and adolescents: A systematic review. *American Journal of Preventive Medicine, 35*(3), 287–313. https://doi.org/10.1016/j.amepre.2008.06.024

Wilson, C., Willis, C., Hendrikz, J. K., & Bellamy, N. (2006). Speed enforcement detection devices for preventing road traffic injuries. *Cochrane Database of Systematic Reviews*, *2*, CD004607.

Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine, 33*, S130–S143.

Wolf, R., Daichman, L., & Bennett, G. (2002). Abuse of the elderly. In E. Krug, L. L. Dahlberg, J. A. Mercy, A. B. Zwi, & R. Lozano (Eds.), *World report on violence and health* (pp. 125–145). Geneva, Switzerland: World Health Organization.

Woratanarat, P., Ingsathit, A., Suriyawongpaisal, P., Rattanasiri, S., Chatchaipun, P., Wattayakorn, K., & Anukarahanonta, T. (2009). Alcohol, illicit and non-illicit psychoactive drug use and road traffic injury in Thailand: A case-control study. *Accident Analysis & Prevention, 41*(3), 651–657.

World Health Organization (WHO). (2007). Youth and road safety. Retrieved from http://apps.who.int/iris/bitstream/10665/43607/1/9241595116_eng.pdf

World Health Organization (WHO). (2009). Global status report on road safety: Time for action. Retrieved from http://apps.who.int/iris/bitstream/10665/44122/1/9789241563840_eng.pdf

World Health Organization (WHO). (2011). Mobile phone use: A growing problem of distraction. Retrieved from http://who.int/violence_injury_prevention/publications/road_traffic/distracted_driving_en.pdf?ua=1&ua=1

World Health Organization (WHO). (2013). Strengthening road safety legislation: A practice and resource manual for countries. Retrieved from http://apps.who.int/iris/bitstream/10665/85396/1/9789241505109_eng.pdf?ua=1

World Health Organization (WHO). (2014a). Global report on drowning: Preventing a leading killer. Retrieved from http://apps.who.int/iris/bitstream/10665/143893/1/9789241564786_eng.pdf?ua=1&ua=1

World Health Organization (WHO). (2014b). *Preventing suicide: A global imperative*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2015a). Global status report on road safety 2015. Retrieved from http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/

World Health Organization (WHO). (2015b). *Preventing youth violence: An overview of the evidence*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2016a). Drug use and road safety: A policy brief. Retrieved from http://apps.who.int/iris/bitstream/10665/249533/1/WHO-MSD-NVI-2016.01-eng.pdf?ua=1

World Health Organization (WHO). (2016b). *INSPIRE: Seven strategies for ending violence against children*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2017a). Powered two- and three-wheeler safety: A road safety manual for decision-makers and practitioners. Retrieved from http://apps.who.int/iris/bitstream/10665/254759/1/9789241511926-eng.pdf?ua=1

World Health Organization (WHO). (2017b). Global Accelerated Action for the Health of Adolescents (AA-HA!): Guidance to support country implementation. Retrieved from http://apps.who.int/iris/bitstream/10665/255415/1/9789241512343-eng.pdf?ua=1

World Health Organization (WHO). (2017c). Managing speed. Retrieved from http://apps.who.int/iris/bitstream/10665/254760/1/WHO-NMH-NVI-17.7-eng.pdf?ua=1

World Health Organization (WHO). (2017d). Save LIVES: A road safety technical package. Retrieved from http://apps.who.int/iris/bitstream/10665/255199/1/9789241511704-eng.pdf?ua=1

World Health Organization (WHO). (2017e). Preventing drowning: An implementation guide. Retrieved from http://apps.who.int/iris/bitstream/10665/255196/1/9789241511933-eng.pdf?ua=1

World Health Organization (WHO) & International Society for Prevention of Child Abuse and Neglect. (2006). *Preventing child maltreatment: A guide to taking action and generating evidence*. Geneva, Switzerland: WHO.

World Health Organization (WHO) & London School of Hygiene & Tropical Medicine. (2010). *Preventing intimate partner and sexual violence: Taking action and generating evidence*. Geneva, Switzerland: WHO.

Yip, P. S. F. (1998). Age, sex, marital status and suicide: An empirical study of East and West. *Psychological Reports, 82*, 311–322.

Yon, Y., Mikton, C. R., Gassoumis, Z. D., & Wilber, K. H. (2017). Elder abuse prevalence in community settings: A systematic review and meta-analysis. *The Lancet: Global Health, 5*(2), e147–e156.

Ytterstad, B., & Sogaard, A. J. (1995). The Harstad injury prevention study: Prevention of burns in small children by a community-based intervention. *Burns, 21*(4), 259–266.

Yuan, W. (2000). The effectiveness of the "Ride Bright" legislation for motorcycles in Singapore. *Accident Analysis and Prevention, 32*(4), 559–563.

Zaloshnja, E., Miller, T. R., Lawrence, B. A., & Romano, E. (2005). The costs of unintentional home injuries. *American Journal of Preventive Medicine, 28*(1), 88–94.

Zhao, X. G., He, X. D., Wu, J. S., Zhao, G. F., Ma, Y. F., Zhang, M., . . . Jiang, G. Y. (2009). Risk factors for urban road traffic injuries in Hangzhou, China. *Archives of Orthopaedic and Trauma Surgery, 129*(4), 507–513.

Zwi, A. B., Garfield, R., & Loretti, A. (2002). Collective violence. In E. Krug, L. L. Dahlberg, J. A. Mercy, A. B. Zwi, & R. Lozano (Eds.), *World report on violence and health* (pp. 213–239). Geneva, Switzerland: World Health Organization.

CHAPTER 10

Global Mental Health

Vikram Patel, Alex Cohen, Brandon Kohrt, Harvey Whiteford, and Crick Lund

Introduction

In 2015, all nations united around a shared mission of achieving the Sustainable Development Goals (SDGs). This set of goals represented an exponential advance from the Millennium Development Goals (MDGs), which the SDGs replaced, both in its aspiration to encompass a substantially broader agenda and through its explicit recognition that these were global concerns, affecting all nations, and requiring global actions. One notable example of this transformative vision was the recognition that health burdens went beyond the MDG focus on a selection of infectious diseases and maternal and child health. Noncommunicable diseases, mental health, and substance abuse received recognition, and several targets related to these concerns were specified (**TABLE 10-1**). At last, it seemed that the decades of science and advocacy had finally allowed mental health to take its rightful place on the global development agenda.

The discipline of global mental health has played a key role in the inclusion of mental health in the SDGs. The field of global mental health was the product of decades of interdisciplinary research and practice in diverse transnational contexts. Governmental and nongovernmental organizations in the United Kingdom such as the Department for International Development (DFID) and the Wellcome Trust funded much of the early work in mental health research and services in low- and middle-income countries (LMICs) in the later part of the 20th century. A series of publications drawing upon this large body of evidence generated by epidemiologists, clinicians, social scientists, and human rights advocates led to a "call to action" in 2007 to "scale up services for people affected by mental disorders built on the twin foundations of cost-effective interventions and respect for human rights" in all countries of the world, and in particular in LMICs, where the attainment of these rights was most seriously compromised (*The Lancet* Global Mental Health Group, 2007).

For the purposes of this chapter, we consider the definition of "mental health" to include all conditions that affect the nervous system that are leading causes of disease burden. Conditions with a vascular or infectious etiology, such as human immunodeficiency virus (HIV) infection of the brain or cerebrovascular diseases, are excluded here, as they are addressed in other chapters of this text. Mental disorders include intellectual disability, epilepsy, anxiety and mood disorders, psychoses (schizophrenia and bipolar disorders), substance use disorders (alcohol and drug use disorders), and dementia. Thus, this list of health conditions includes disorders that clinicians may categorize as psychiatric, neurological, or substance use disorders. This chapter uses the acronym MNS disorders (meaning "mental, neurological, and substance use disorders"), which is common nomenclature for the World Health Organization (WHO) and other health and development institutions.

TABLE 10-1 Mental Health in the Sustainable Development Goals

	United Nations' Sustainable Development Goals
SDG 3	Ensure healthy lives and well-being for all at all ages
SDG Target 3.4	Requests that countries: "By 2030, reduce by one third premature mortality from non-communicable diseases through prevention and treatment and promote mental health and well-being"
SDG Target 3.5	Requests that countries: "Strengthen the prevention and treatment of substance abuse, including narcotic drug abuse and harmful use of alcohol"
SDG Target 3.8	Requests that countries: "Achieve universal health coverage, including financial risk protection, access to quality essential health-care services and access to safe, effective, quality and affordable essential medicines and vaccines for all"

Modified from United Nations Sustainable Development goals: http://www.un.org/sustainabledevelopment/sustainable-development-goals/

This rich interdisciplinary treasure of knowledge laid the foundation for the landmark 2007 *The Lancet* series on global mental health. The authors of this series of articles arrived at the conclusion that the high burden and unmet needs for care constituted a global health crisis. They deliberated on what might be the most urgent, clear, and specific "call to action" for the global health community and, in the end, chose to focus on the needs of those individuals affected by a mental disorder, calling for actions to reduce the treatment gap by scaling up the coverage of services for mental disorders in all countries, but especially in LMICs (*The Lancet* Global Mental Health Group, 2007). The years following the publication of *The Lancet* series witnessed a tangible increase in the attention to the treatment gap in LMICs, as evidenced by the increase in development assistance for mental health, which more than doubled in absolute dollars in the years immediately after 2007 (Gilbert, Patel, Farmer, & Lu, 2015). In 2011, the Grand Challenges in Global Mental Health initiative, led by the U.S. National Institute for Mental Health (NIMH), emphasized implementation research questions as the priorities to reduce the treatment gap for mental disorders (Collins et al., 2011) (**TABLE 10-2**).

The publication of these priorities led to a slew of new research initiatives by the NIMH to support research and training in global mental health as well as a set of international "hubs" for research on task sharing and scaling up mental health interventions, while Grand Challenges Canada supported dozens of projects addressing some of these priorities in a number of LMICs. Simultaneously, a number of global institutions and coalitions began to advocate for mental health. WHO launched its flagship mental health Gap Action Programme (mhGAP) to scale up care for MNS disorders in 2008 and developed a series of seminal publications that provide guidance to health practitioners in nonspecialist settings regarding treatments for MNS disorders (WHO, 2010); track the status of mental health systems at the country level (WHO, 2015); and establish standards of care for mental health facilities (see www.who.int/mental_health for a comprehensive listing of WHO resources for mental health). This culminated in the Comprehensive Mental Health Action Plan 2013–2020, which was agreed by all nations of the world in 2013, and the establishment of a roadmap for achievement of a broad range of mental health–related targets (Saxena, Funk, & Chisholm, 2013). The Disease Control Priorities Network published its recommendations, targeting governments and development agencies, for which interventions should be scaled up through diverse platforms from the community to specialist care, ultimately forming the mental and neurological health component of the package of interventions for universal health care (Patel et al., 2016). Notably, both the Comprehensive Mental Health Action Plan and the Disease Control Priorities Network recommendations took a much broader view of mental health, emphasizing the continuum from promotion of mental health to prevention of mental disorders, to treatment, long-term care, and inclusion of persons with mental disorders.

Preceding these developments were a number of reform initiatives in specific countries—for example, in Brazil and Italy—that sought to influence and promote a public health approach to mental health care. These efforts aimed to provide community-oriented

TABLE 10-2 Grand Challenges in Global Mental Health, 2011

Rank	Five Leading Challenges for Global Mental Health
1	Integrate core packages of mental health services into routine primary health care
2	Reduce the cost and improve the supply of effective psychotropic drugs for mental, neurological, and substance use disorders
3	Train health professionals in LMICs to provide evidence-based care for children with mental, neurological, and substance use disorders
4	Provide adequate community-based care and rehabilitation for people with chronic mental illness
5	Strengthen the mental health component in the training of all healthcare personnel to create an equitable distribution of mental health providers

Data from Collins, P. Y., Patel, V., Joestl, S., March, D., Insel, T. R., Daar, A. S. (2011). Grand challenges in global mental health. *Nature, 475*, 27–30.

care along with both medical and psychosocial interventions, while strongly emphasizing the priorities of those affected by mental disorders. The ultimate goal was to demonstrate that such approaches could improve access to quality care.

Reforms in Italy began in the 1960s, and reforms in Brazil can be traced to the 1990s, with the psychosocial community center program known as *Centros de Atenção Psicossocial* (CAPS) beginning approximately four to five years prior to the mhGAP initiative. Civil society partnership with mental health professionals to promote a shared vision continued to grow during the twenty-first century. The most notable example was the Movement for Global Mental Health (www.globalmentalhealth.org), which was launched in 2008 as a virtual global alliance. By 2015, the Movement included 170 member institutions representing diverse stakeholders, from academics to civil society representatives. Since 2013, the Movement has been led by persons affected by mental disorders. In several countries, prominent individuals have "come out" with their personal accounts of mental disorders, indicating the growing acceptance of this form of human suffering. The field of global mental health has become a respected discipline in its own right, complete with academic programs and centers in universities around the world, specialist journals and books on the subject, and an annual calendar of scientific events. Not surprisingly, the discipline has been described as having "come of age" (Patel & Prince, 2010).

This chapter is organized in three parts. First, it presents a brief history of global mental health prior to the *The Lancet* series. Next, it discusses four foundations of the discipline: the influence of culture and social determinants on mental health; the burden and impact of mental disorders (including substance use disorders and neurological disorders); effective prevention and treatment strategies; and the ways that these strategies can be delivered in low-resource contexts. Finally, it considers the limitations of the field as it is currently conceptualized and the strategies for addressing global mental health issues in the future.

Historical Development of Global Mental Health

The earliest developments of public mental health care can be traced at least as far back as the early Islamic world of the Middle East, North Africa, and Spain. Although accounts differ, it seems that the first hospitals that cared for persons with mental disorder were established by Islamic physicians during the ninth century CE in Baghdad and Egypt (Dols, 1987). Within a few hundred years, institutional care had spread 3,000 miles to the west in Marrakech (twelfth century) and Fez (thirteenth century), Morocco (Moussaoui & Glick, 2015). Influenced by the practices in Morocco, in the fourteenth century a hospital in Granada, Spain, began to accept persons with mental disorders. In the fifteenth and sixteenth centuries, hospitals for persons with mental disorders were established in at least five cities in Spain. In Northern Europe, there is evidence of the establishment of institutional care that coincides with or predates the founding of the hospitals in Spain (Mora, 2008; Pierloot, 1975). The most famous examples from Northern Europe are Bethlem

Hospital in London, which traces its establishment to the thirteenth century, and the Colony of Geel in Belgium, whose origins date from approximately the same time (Mora, 2008).

The beginning of modern public mental health can be traced to the late eighteenth century, when there was a decided shift in beliefs about the nature of mental disorder. Before this time, "madness" was associated with a loss of rationality, which meant that persons with mental disorders were considered as less than human and, in an effort to restore them to reason, were treated as brutes (Scull, 1989). "Moral treatment," which was developed simultaneously and independently in France (Weiner, 1992), England (Digby, 1985), and elsewhere (Scull, 2015), rejected the notion that mentally ill people lacked reason and suggested, instead, that tolerance and confinement in a well-ordered and pleasant environment could restore a person to rationality and mental health (Grob, 1994). The example of the York Retreat in England, which was established in the late eighteenth century as one of the sites in which moral treatment was developed, gave rise to "a wave of enthusiasm and optimism" for the curability of madness (Scull, 2015). This new perspective on mental disorders and their treatment brought about a powerful movement to abolish the abuses and to establish public systems of institutions that would offer beneficent care and the prospect of recovery to persons with mental disorders. One of the best examples of this advocacy was the work of Jean-Étienne Esquirol, who, after visiting a number of psychiatric institutions in France, wrote a report in 1819 in which he advocated for a state-run system of asylums (Goldstein, 2001). In 1838, France followed Esquirol's advice; other countries established similar systems in first half of the nineteenth century (Scull, 2015). The importance of this form of care for persons with mental disorders can be seen readily in the thousands of articles about asylums that were published in the nineteenth century in the leading English-language medical journals (Cohen & Minas, 2016).

As soon as the public asylums opened, they were filled beyond capacity. Throughout the second half of the nineteenth century, the notion of small curative institutions was abandoned due to increasing demands for services and a reluctance on the part of governments to allocate more funds for the care of mentally ill indigent persons (Grob, 1994; Scull, 1989). In addition, as conditions in the asylums grew worse, the effectiveness of asylum care—an assumption that had underpinned the widespread establishment of asylums—began to be questioned. These developments can be seen in publications in *The Lancet* over a period of 50 years. In 1827, the journal published a letter decrying the conditions in private asylums and called for "the establishment of public hospitals for insane paupers," and called to task the British Colleges of Physicians and Surgeons for not devoting more attention to these issues (Humanitas, 1827). Thirty years later, *The Lancet* (1857) published an editorial, "The Crime of Lunacy and How We Punish It," that questioned the efficacy of asylums: "They are . . . mere houses of detention." In 1875, the journal established the Commission on Lunatic Asylums to investigate "the treatment, in public and special institutions, of patients laboring under the various forms of mental disease" (*The Lancet*, 1875). The conclusions of the Commission were not positive: "It has been found impossible 'to formulate the system of treatment' either for cure or relief. Practically, there is no general 'system' beyond that which may be described as control" (*The Lancet*, 1877). Thus, by the late nineteenth century, public mental health efforts were inextricably associated with the wretched, overcrowded conditions in asylums: "the positive images of hospitals that had prevailed in the mid-nineteenth century [had given] way to far more negative ones associated with hopelessness, abuse, and ultimately death" (Grob, 1994).

Despite these concerns, the colonial enterprise saw the global expansion of institutional care for persons with mental disorders starting in 1567 when Spain established a psychiatric institution in Mexico City. The first institution of its kind in the Western Hemisphere, it heralded what later became known as colonial psychiatry. Nearly 200 years later, and continuing throughout the nineteenth century and into the first decades of the 20th century, "lunatic" asylums were established by British authorities in India (Basu, 2016; Weiss, 1983). Asylums were also established in French (Edington, 2013; Keller, 2008), Italian (Scarfone, 2016), Dutch (Pols, 2006), Spanish (Meyer, 2010), and other British (Ng & Chee, 2006; Sadowsky, 1997; Swartz, 1999) colonies. These institutions have often been depicted as being a form of racism and a means of exerting social control over indigenous populations (Goddard, 1992; Jackson, 1999; Schmidt, 1967; Swartz, 1999). However, it is also true that asylums in the colonies merely reflected the values and treatments in the home-country asylums of the colonial powers (Weiss, 1983).

Despite deteriorating conditions in asylums (rebranded as mental hospitals and psychiatric hospitals in the United Kingdom and the United States, respectively [Cohen & Minas, 2016]), the number of psychiatric inpatients continued to increase in the wealthy nations of the West. Then, beginning in the 1950s, efforts in North America, Western Europe, and Australia were initiated to remove long-term patients

from psychiatric facilities and provide treatment and care in the community. The incentive for what came to be called *deinstitutionalization* evolved from a convergence of several forces. First, encouraged by successful treatment of soldiers traumatized by their experiences in World War II, psychiatrists became optimistic about their ability to effectively treat mental disorders outside of hospital settings (Grob, 1994). Second, there was a growing awareness that the abusive conditions found in public psychiatric hospitals, as well as the negative effects on patients of long-term institutionalization, were at least as harmful as the chronic mental disorders themselves. Third, caring for patients in large institutions that did not provide effective care was expensive. Finally, the discovery in 1954 of chlorpromazine, the first effective antipsychotic medication, offered people with chronic mental disorders the prospect of living in the community rather than as inpatients (Greenblatt, 1992).

Together, these forces brought about dramatic changes in institutionalized populations. In the United States, for example, 559,900 people were in psychiatric hospitals in 1955; 25 years later, that number had decreased to 138,000 (Goldman, 1983). In Britain, the number of psychiatric inpatients peaked at 148,100 in 1954 and then steadily decreased; by 1985, there were only 64,800 such inpatients (Thornicroft & Bebbington, 1989). In Italy, deinstitutionalization took place somewhat later, but had similar results: Early in the 1970s there were about 75,000 psychiatric inpatients, but that number had dwindled to only 38,000 in 1981 (Morosini, Repetto, De Salvia, & Cecere, 1985).

Deinstitutionalization has had mixed reviews. Although large numbers of patients were discharged from hospitals, many came to be accommodated, at least in the United States, in prisons, nursing homes, and adult homes for mentally disabled persons (Scull, 1985), where the worst aspects of the old asylums were often recreated (Human Rights Watch, 2015; Levy, 2002a, 2002b, 2002c, 2002d). In Europe, some evidence suggests that "reinstitutionalization" has taken place. For example, as the number of psychiatric beds in hospitals has decreased, the number of persons with mental disorders in prisons and forensic hospitals has increased (Fakhoury & Priebe, 2007; Human Rights Watch, 2016; Priebe et al., 2005).

Although deinstitutionalization is considered a global policy (WHO & Gulbenkian Global Mental Health Platform, 2014), it has occurred primarily in Western countries, with a few exceptions. For example, the number of psychiatric hospital beds has increased in South Korea (Kim, 2017). Not only does Japan have the highest rate of psychiatric beds in the world (28.4 beds per 10,000 people), but it also has the longest average length of psychiatric hospitalization (Imai et al., 2005). Deinstitutionalization has not taken place at all in the great majority of LMICs, such that hospital-based care remains at the center of what exists of those countries' national mental health systems (Saxena, Thornicroft, Knapp, & Whiteford, 2007). A noteworthy exception is Brazil, which, between 1995 and 2005, reduced its number of psychiatric beds by 41% while increasing community services by a factor of 9 (Andreoli, Almeida-Filho, Martin, Mateus, & Mari, 2007). Other research suggests that while psychiatric beds decreased in six South American countries, including Brazil, prison populations have increased substantially (Mundt et al., 2015). It has been noted, however, that the increase in the prison population cannot be definitively attributed to deinstitutionalization (Winkler et al., 2016). More recently, other large middle-income countries have begun reforming their mental healthcare systems: India launched its first national mental health policy in 2014 with similar principles while China has greatly expanded coverage of care for mental disorders through its 686 Project, which refers to the first 6.86 million Renminbi (USD$ 829,000 in 2004) invested by the Chinese government to free patients from seclusion and restraints throughout the country (Ma, 2012; Patel et al., 2017).

Perhaps the best way to consider the positive and negative effects of deinstitutionalization is to ask the consumers of mental health services whose lives have been most affected by this policy. When that has been done, the answer is clear. Despite the difficulties of life outside the hospital, the relative lack of supportive services, and the effects of stigma and discrimination, former long-term inpatients generally agree they prefer living in the community (Davidson, Hoge, Godleski, Rakfeldt, & Griffith, 1996).

The development of community mental health services was both an impetus for and a product of deinstitutionalization. Although the potential for treating people with mental disorders outside of hospital settings was recognized prior to the discovery of effective antipsychotic medications, the evolution of community-based services came about with the recognition that treatment and care required a range of social and rehabilitation services and involved more than just dispensing medication. Thus, while the definition of community care may have once simply meant care outside hospitals, it now encompasses, at least ideally, professional services in community settings, social reintegration, and support services, such as housing, employment, medical care, and welfare (Tansella & Thornicroft, 2001). The provision of community services to individuals with mental

disorders is now a central principle of global mental health (Alem, 2002; Alem, Jacobsson, & Hanlon, 2008; Hanlon, Wondimagegn, & Alem, 2010; Semrau, Barley, Law, & Thornicroft, 2011; Thara & Padmavati, 2013; Thornicroft, Alem, Dos Santos, et al., 2010; Thornicroft, Tansella, & Law, 2008), though this remains an unfulfilled goal for the vast majority of persons affected by mental disorders globally.

The scientific and policy foundations of global mental health can be traced, at least in part, to activities at WHO that were begun soon after the establishment, in 1949, of a mental health division and an Expert Committee on Mental Health (Lovell, 2014). Two reports by the Expert Committee were of particular importance to the field that would become known as global mental health. The first of these reports, *Epidemiology of Mental Disorders* (WHO, 1960), sought to lay the groundwork for an international system for the classification of the psychiatric disorders and, more generally, to advance the ultimate goal of creating a scientifically rigorous approach to psychiatric epidemiology. Direct outcomes of this work included WHO's international studies of schizophrenia (Jablensky et al., 1992; WHO, 1973), as well as studies of schizophrenia in Mauritius (Murphy & Raman, 1971) and Sri Lanka (Waxler, 1979). In fact, there was a burgeoning of psychiatric research in developing countries during the 1960s (e.g., Lambo, 1960; Leighton et al., 1963; Lin, Rin, Yeh, Hsu, & Chu, 1969), 1970s (e.g., Harding, 1973; Jilek & Jilek-Aall, 1970; Kulhara & Wig, 1978), and 1980s (e.g., Harding et al., 1980; Kleinman, 1980; Kulhara & Chandiramani, 1988). The second of the Expert Committee's reports, *Organization of Mental Health Services in Developing Countries* (WHO, 1975), followed a *British Journal of Psychiatry* series on psychiatric problems in the developing world (Carstairs, 1973; German, 1972; Leon, 1972; Neki, 1973) and recommended many actions to address the burden of mental disorders in developing countries. One of the most farsighted statements in the report was the following: "A number of innovations have been recommended, notably the sharing of mental tasks by a wide range of health workers and by other community agencies."

WHO followed up these recommendations with a series of projects that attempted to integrate mental health services into primary care settings in seven low-income countries (Sartorius & Harding, 1983). Now, more than 40 years later, this strategy remains central to the development of mental health services in low-resourced settings.

As important as all of these projects were, perhaps the "founding" event of global mental health was the publication of the results of the first Global Burden of Disease (GBD) study, which introduced the concept of disability-adjusted life-years, a single measure that represented the burden imposed by both mortality and morbidity (the most recent findings of this initiative are described in more detail later in this chapter) (World Bank, 1993). Prior to this report, infectious diseases were prioritized by international public health efforts. However, the GBD study revealed that noncommunicable diseases—neuropsychiatric disorders, specifically—accounted for a significant proportion of the GBD. These findings provided the most compelling evidence that mental disorders were priority public health disorders in low-, middle-, and high-income countries alike. Thus, the GBD study opened the way to the field that became known as global mental health.

Culture and Mental Disorders

It is crucial to understand the role of culture in the experience, diagnosis, and treatment of mental disorder (Kirmayer & Swartz, 2014). Common elements in the definition of culture are "values, beliefs, knowledge, norms, and practices and the notion that that these are shared among a specific set of people" (Hruschka & Hadley, 2008, p. 947). Beliefs refer to conscious psychological processes. Norms are behaviors maintained by social sanctioning and affective responses. Values are valences placed on beliefs, knowledge, and norms that lead to engagement in or avoidance of behaviors. These beliefs, norms, and values shape both lay and professional understandings of mental disorder. This section addresses the following topics: (1) the role of culture in biomedical classification systems; (2) alternative approaches to mental disorder categorization across diverse cultural contexts; and (3) cross-cultural methods and approaches to improve mental health research and services.

Cultural Considerations for Biomedical Psychiatric Classification

Understanding the origins and assumptions of different diagnostic systems is vital to implement best practices in research and clinical care for diverse global populations. The classification of diseases leads, in theory, to more accurate diagnoses and effective treatments. Valid and reliable systems of classification make it possible to determine accurate prevalence and incidence rates and, therefore, should guide decisions about the development of services. The classification of mental disorders, however, presents some unique

challenges. Psychiatric diagnoses do not "carve nature at the joint" and the boundaries between different conditions may not be distinct (Blacker & Tsuang, 1992; Kendler & Gardner, 1998; Tsuang, Stone, & Faraone, 2000). Unlike other diseases, there are no specific and replicable pathophysiological pathways to distinguish most mental disorders in a clinical setting. Moreover, clinical classification in psychiatry is based on the symptom profiles, rather than on the disorders' etiology (as in the case of infectious diseases) or their pathology (as in the case of vascular disease).

Two main biomedical systems of psychiatric classification are used today: the *International Classification of Diseases* [ICD] (WHO, 1990) and the *Diagnostic and Statistical Manual of Mental Disorders* [*DSM*] (American Psychiatric Association [APA], 2013). These approaches to psychiatric categorization took shape in the late 1800s with the work of Emile Kraepelin and others who developed diagnoses based on standardized factors related to symptom presentation, prognosis, and hereditary risks (Jilek, 1995). Kraepelin and his followers also influenced modern frameworks regarding the incorporation of culture into psychiatric classification. Kraepelin traveled to Java, Sri Lanka, India, Singapore, and the Americas to evaluate his diagnostic criteria with non-European populations. He and his followers developed a model with a biological core for mental disorders (i.e., pathogenic), with cultural influences shaping the presentation of that biological core (i.e., pathoplastic) (Jilek, 1995). For example, hallucinations were considered a universal biologically-determined feature of psychosis whereas culture shaped the content of those hallucinations. Most current classifications (e.g., DSM and ICD) continue this approach with cultural considerations for different conditions while assuming conditions are grounded in a core of presumed universal psychiatric pathology.

These classification systems—DSM and ICD—and their assumption of biological universality are often considered *etic* perspectives in cross-cultural psychiatry. "Etic" refers to a universal classification systems and is contrasted with "emic," which refers to classifications specific to a culture, which may or may not overlap with etic categorizations (Hahn, 1995). The use of the DSM and ICD psychiatric categorizations as overarching etic frameworks applicable across all cultures has been critiqued by some psychiatrists and social scientists (Kleinman, 1988; Kleinman & Good, 1985; Littlewood, 2002). The DSM and ICD do not incorporate all emic divisions from other cultures and medical nosologies, such as Chinese, Ayurvedic, or Q'uranic classifications of mental disorder. It is argued that classifications of psychiatric disorders largely reflect American and European emic concepts of psychopathology based on implicit cultural concepts of normality and deviance. In addition, the classification systems have not been comparably applied across racial/ethnic groups within the same clinical context. For example, in the United Kingdom and United States, African- and Caribbean-descent populations are more likely to be diagnosed with schizophrenia compared with white/Western European-descent populations with the same symptom profile; the latter are diagnosed with bipolar disorder more frequently than with schizophrenia (Bell, Williamson, & Chien, 2008; Louden, 1995). For a challenge to these claims, see the work of Selten and Hoek (2008) and Singh, Greenwood, White, and Churchill (2007).

Even when comparable symptoms are observed, the social and personal significance of the behavior or experience may not have the same meaning and their explanatory models may differ. Researchers have cautioned that there is a risk of confounding culturally distinctive behavior with psychopathology on the basis of superficial similarities of behavior patterns or phenomena, an assumption referred to as a "category fallacy" (Kleinman, 1987). Within European and North American cultural contexts, there also have been debates about the scope of depression and how it is—or is not—distinguished from normal grief (Wakefield & Demazeux, 2015).

Due to these shortcomings, cross-cultural studies in psychiatric epidemiology historically have suffered from several problems. For example, case identification techniques varied from site to site, and methods were not standardized (Gorenstein, 1992). These inconsistencies led to a movement to standardize the process of psychiatric measurement and diagnosis so that diagnoses would be more replicable among clinicians and across patient populations. In the United States, one expectation since the third edition of the DSM (APA, 1980) was that this kind of standardization would help to reduce bias and increase the reliability of diagnoses across populations and among clinicians. Similarly, the tenth revision of the ICD (ICD-10) was developed with the explicit purpose of being an international standard (WHO, 1990). Thus, efforts were made to ensure that the drafters of the ICD-10 were drawn from as many countries as was feasible, and the revised system of classification was field-tested by more than 700 clinicians in 39 countries from all continents. The vast majority of ICD-10 conditions have reasonable reliability (Sartorius et al., 1993).

In preparation for the ICD-11, which was due to be published in 2018, one approach to explore cultural bias has been to evaluate how clinicians across cultures categorize disorders based on their clinical experiences. In a study of 517 clinicians in eight countries ranging from Brazil to India to Japan, clinicians had strong inter-rater agreement regardless of their national or cultural origin (Reed et al., 2013). The same study found that the cross-cultural categorizations were generally consistent with DSM and ICD, but the clinician-generated categories were more consistent with the proposed changes for ICD-11 compared to the DSM-IV or ICD-10, in particular with regard to personality disorders.

Critiques of cultural biases in diagnoses are part of a broader movement in psychology that has detected biases when the vast majority of research is based on a single population demographic. This problem has been referred to as the bias of psychology research being dominated by WEIRD populations (Western, Educated, Industrialized, Rich, and Democratic), which account for 90% of psychology publications (Henrich, Heine, & Norenzayan, 2010b). When psychological studies have been conducted with non-WEIRD populations, researchers have discovered that presumed universal processes such as visual perception, spatial reasoning, and behavioral motivation related to fairness and cooperation have cultural variations. In addition, it is important to consider that diagnostic criteria of syndromes can and do change over time, as is well demonstrated by the regular revisions of international psychiatric classifications, and that these changes are considerably influenced by attitudinal, political, and historical factors. Anthropologists have investigated how diagnoses are developed out of specific social, economic, and political environments; for example, there are in-depth historical accounts with anthropological critiques of disorders including posttraumatic stress disorder (PTSD) (Young, 1995), bipolar disorder (Martin, 2007), and dissociative disorders (Hacking, 1998).

Fifty years of research suggests that concepts of psychosis are remarkably similar across diverse settings (Cohen et al., 2016). Cross-cultural reviews of depression, PTSD, and other conditions have found that some symptoms may vary across cultures, though they have also found many universals. In a systematic review covering 178 social groups (varied by location, ethnicity, culture, and/or geography), of which 115 were non-Western populations, DSM hallmark symptoms of depressed mood/sadness, fatigue/loss of energy, problems with sleep, appetite/weight problems, and suicidal thoughts were common across all groups (Haroz et al., 2017). At the same time, the review revealed that some symptoms with near universality are not captured in current biomedical classifications. For example, symptoms such as social isolation/loneliness, crying a lot, somatic complaints, and thinking too much were among the most common symptoms but are not included in DSM criteria. Even among Western populations, the symptoms of social isolation and crying a lot were among the top 10 most associated with depression. In contrast, DSM diagnostic symptoms such as psychomotor agitation/retardation, feelings of worthlessness/guilt, and poor concentration were not among the top 10 associated symptoms in Western and non-Western cultural groups.

A cross-cultural review of PTSD concluded that while this diagnosis is generally "valid"—that is, it reflects the reality of human experience across populations—there is considerable cultural variability (Hinton & Lewis-Fernandez, 2011). The PTSD review highlighted cultural differences in the meaning and interpretation of trauma symptoms, the prevalence of the diagnosis across groups, the prevalence of specific symptoms such as those related to avoidance and numbing, the likelihood that a person of one culture will develop PTSD compared to a person of another culture given the same trauma exposure, and the association with nontraumatic stressors.

These types of cross-cultural systematic reviews are helpful to identify potential cultural biases in psychiatric diagnoses and to point us toward approaches that may be more appropriate across populations.

Alternative Approaches to Categorize and Conceptualize Mental Disorders

Historically, classifications of mental disorders prior to the DSM and ICD have existed in different cultures. All cultures have some form of distinguishing normal from abnormal, or socially acceptable versus unacceptable behavior, and this is often the province of religious practitioners and traditional healers (Clifford, 1990; El-Islam, 1982; Kleinman, 1980; Kohrt, Hruschka, Kohrt, Panebianco, & Tsagaankhuu, 2004; Weiss et al., 1988).

As an example, the first Chinese Classification of Mental Disorders (CCMD) appeared in 1979; since then, the system has undergone several revisions. Its third and most recent version is heavily influenced by the ICD-10 and DSM-IV systems, but still retains local features. The main differences between the ICD-10 and the CCMD-3 are the CCMD-3's retention of the term "neurosis" and categories of neurotic disorders such as neurasthenia (Lee, 2001). Personality disorders are

less often diagnosed in Chinese populations; thus, two categories of personality disorders—borderline personality disorder and avoidant personality disorder—are excluded from the Chinese scheme. The CCMD also includes its own section of culture-related mental disorders such as *qigong*-induced mental disorder. *Qigong* is a trance-based form of a traditional Chinese healing system. The disorder is similar to a dissociative state, featuring identity disturbance, irritability, hallucinations, and aggressive and bizarre behaviors. These often acute, brief episodes are linked to excessive practice of *qigong* meditation by physically or psychologically ill subjects.

Emic conditions (i.e., locally relevant categorizations) that are not typically observed in Western culture have been referred to as culture-bound syndromes (Simons & Hughes, 1985). Examples include *latah* (a startle-based form of distress in Southeast Asia), *koro* (a fear of genital retraction also noted in Southeast Asia), and *ode ori* (a West African form of distress characterized by thinking too much and attributed to a worm crawling in the skull) (Makanjuola, 1987; Simons & Hughes, 1985). Over time, the "bounded" aspect of culture-bound syndromes has been challenged due to findings of similar patterns of distress in disparate cultural settings, a lack of cohesive symptom presentation characterizing a syndrome, and the diversity in etiological attributions and vulnerable groups (Hahn, 1995; Kirmayer & Minas, 2000). Dissatisfaction with the term "culture-bound syndrome" has led researchers to propose other labels, such as "idiom of distress," "popular category of distress," "cultural syndrome," and "explanatory model" (Nichter, 1981; Weiss, 1997). In DSM-5, the term "cultural concept of distress" (CCD) attempts to aggregate these concepts without implying cultural exclusivity: "*Cultural concept of distress* refers to ways that cultural groups experience, understand, and communicate suffering, behavioral problems, or troubling thoughts and emotions" (APA, 2013, p. 758).

There is now a large literature indicating commonalities in emic categories of psychological distress. A systematic review (Kaiser et al., 2015) identified 138 publications mentioning "thinking too much," with examples from every populated continent including terms such as *kufungisisa* (Zimbabwe), *reflechi twòp* (Haiti), *pensando mucho* (Nicaragua), and *kut careen* (Cambodia), and, in more recent studies, *ucingakakhulu* (South Africa) (Davies, Schneider, Nyatsanza, & Lund, 2016). Across cultures, "thinking too much" is characterized by ruminative, intrusive, and anxious thoughts that, if prolonged, result in a range of physical and psychological complaints. Common associated symptoms include low mood, anhedonia, poor concentration, social withdrawal, sleep disruptions, and somatic complaints across most populations. "Thinking too much" is associated with similar socioeconomic and traumatic stressors across populations. Moreover, coping mechanisms share commonalities across cultures, such as controlling or suppressing thoughts, distraction, and engaging in social activities—all of which overlap with evidence-based psychological treatments. This study is one of the few examples of how starting with a CCD rather than with a psychiatric category can help generate information on commonalities in psychological distress across populations.

A systematic review of CCDs and their overlap with DSM and ICD psychiatric categories demonstrates that the presence of a CCD increases the likelihood of persons meeting criteria for a psychiatric disorder (Kohrt et al., 2014). However, there is not a one-to-one match; that is, no CCD predicts a psychiatric diagnosis in 100% of individuals. On average, presence of a CCD does strongly predict PTSD, and to a lesser degree CCDs are predictive of depression and general psychological distress (**FIGURE 10-1**). Despite somatic complaints (i.e., physical symptoms such as headaches, digestive problems, and numbness and tingling) frequently being an aspect of CCDs, most CCDs are more strongly associated with common mental disorders than with somatoform disorders. Guidance for conducting rigorous epidemiologic studies of CCDs is now available with the tool known as Systematic Assessment of Quality in Observational Research for Cultural Psychiatry Epidemiology (SAQOR-CPE; Kohrt et al., 2014).

Cultural concepts of distress are increasingly recognized as an important component of culturally adapting mental health treatments, especially psychological treatments. Studies with the strongest effect sizes (i.e., those with the biggest improvements in mental health from pre- to post-treatment) do not utilize psychiatric categories when framing distress for patients and clients. For example, in India, depression and common mental disorders (CMDs) are often framed as "stress" or "tension" (Patel et al., 2010; Petersen et al., 2016), which are cultural idioms of distress that have many features of CMDs but are nonstigmatizing and facilitate treatment engagement (Weaver, 2017). In Zimbabwe, a lay health worker–delivered psychological treatment, "The Friendship Bench," uses the CCD *kufungisisa* (Shona for "thinking too much"; Patel, Simunyu, & Gwanzura, 1995) as a culturally acceptable way to discuss CMDs (Chibanda et al., 2016). Similarly, treatment of refugees and immigrants with culturally adapted cognitive-behavioral therapy has framed the interventions with CCDs (e.g., *nervios*-related conditions for

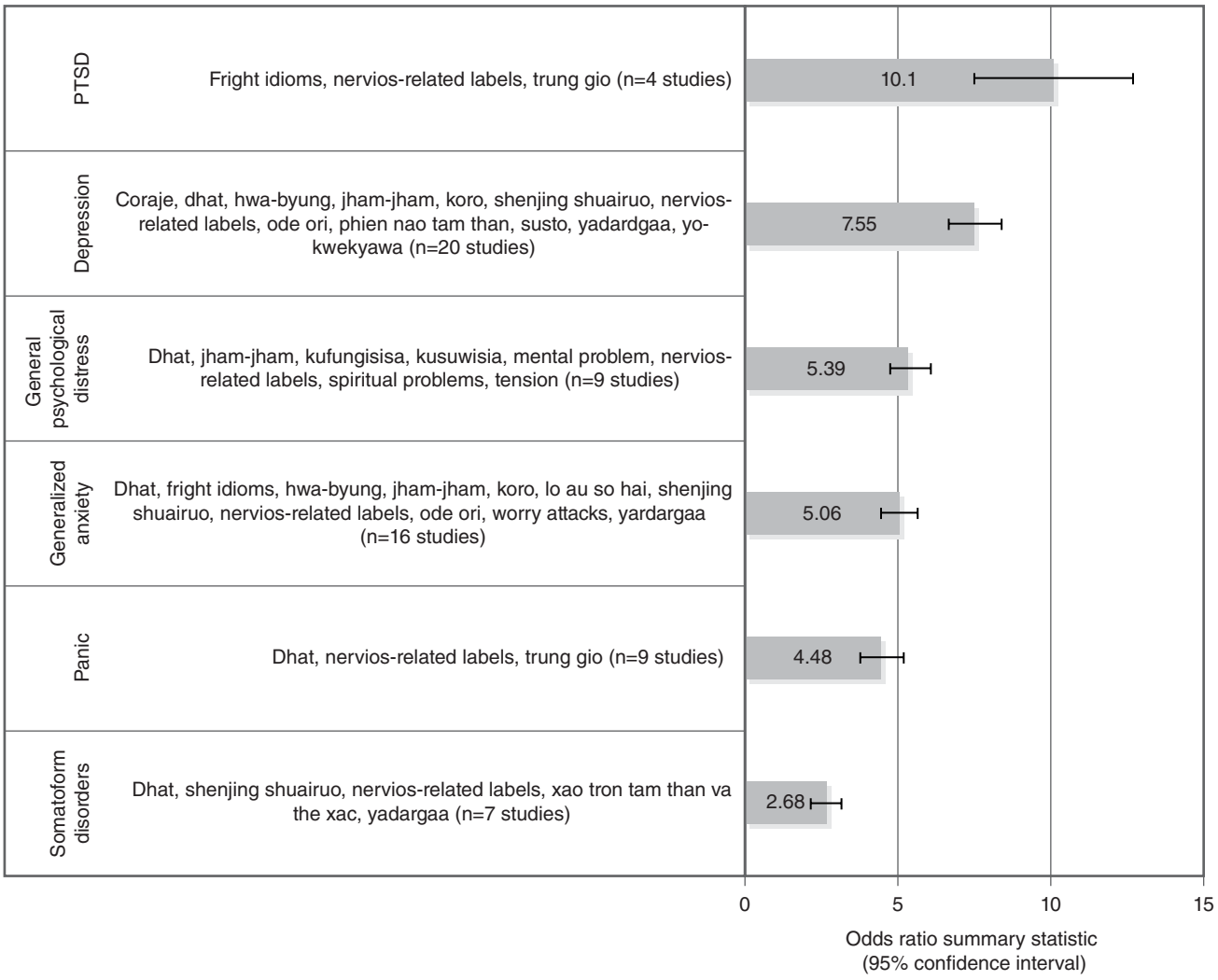


FIGURE 10-1 Psychiatric disorders and cultural concepts of distress.

Modified from Kohrt, B. A., Rasmussen, A., Kaiser, B. N., Haroz, E. E., Maharjan, S. M., Mutamba, B. B., de Jong, J. T., & Hinton, D. E. (2014). Cultural concepts of distress and psychiatric disorders: literature review and research recommendations for global mental health epidemiology. *International Journal of Epidemiology 43*(2): 365–406.

Latinos and *khyaal attacks* for Cambodians), and studies have shown significant improvements in both the CCDs and in reduction of depression and PTSD (Hinton, Hofmann, Pollack, & Otto, 2009; Hinton, Rivera, Hofmann, Barlow, & Otto, 2012).

Within high-resource settings in European and North American cultures, there are also concerns about the generalizability and conceptual assumptions in psychiatric nosologies such as the DSM. In the United States, NIMH has proposed the Research Domain Criteria (RDoC) as a way to look at processes from genes to functional neuroscience to environment in an effort to understand etiology and presentation and ultimately guide treatment (Insel et al., 2010). This more dimensional system provides an alternative to the categorical divisions between "mentally ill" and "normal," which have been contested by cross-cultural researchers (Bilder, 2015). Moreover, this approach can incorporate social genomics and social neuroscience methods that are increasingly demonstrating that culture and biology interact (Cole, 2014; Rule, Freeman, & Ambady, 2013). Individual and collective experiences including acculturation, especially during child development, determine much of the biological and psychological processes observed in adulthood.

Ultimately, much of RDoC focuses on the individual level. Nevertheless, there are increasing calls for an analogous approach to societal-level causal mechanisms that would be useful from a public and global mental health perspective (Stein, Lund, & Nesse, 2013). Going forward, there is potential for social neuroscience and cross-cultural psychiatry to find areas of convergence to develop more accurate diagnostic approaches to apply across diverse populations (Kirmayer & Gold, 2012b).

Cross-Cultural Methods and Approaches for Mental Health Research and Services

Ultimately, diagnostic categories are a moving target, as they are continually being transformed by improved biocultural research that seeks to better understand mechanisms and improve the fit between types of interventions and persons with psychological distress. As diagnoses continue to change, it is crucial to have a set of rigorous methodologies that can integrate culture into global mental health research and interventions. Examples of these rigorous methods are transcultural translation and validation procedures for assessment and the Cultural Formation Interview for clinical care.

Cultural equivalence is needed for psychiatric assessment tools used for research, screening, and treatment monitoring. Lack of cultural equivalence risks category fallacies and under-, over-, and misdiagnosis. A tool is considered to have cultural equivalence (i.e., measurement invariance) if it measures the same construct across cultures. Content, semantic, technical, criterion, and conceptual equivalence are needed to achieve measurement invariance (Flaherty et al., 1988). These are defined as follows:

- *Content equivalence:* Does the phenomenon in question occur and is it locally recognized as distressing within the target culture?
- *Semantic equivalence:* Does the meaning of each item remain the same after translation?
- *Technical equivalence:* How does the method used in data collection affect results differentially between cultures? Technical methods could encompass response options on scales, item structure, and administration format (e.g., pen and paper, computerized assessment, interview administered). Challenges have been identified with the use of Likert scales, statements versus questions, positive versus negative wording, and pictorial response scales (Kohrt et al., 2011; Weobong et al., 2009).
- *Criterion equivalence:* What is an instrument's relationship with previously established and independent criteria for the same phenomenon? According to criterion equivalence, there should be comparable psychometrics (e.g., sensitivity, specificity, positive predictive value, negative predictive value) with a known marker, oftentimes a structured clinical interview.

There are now best practices for transcultural translation and validation that incorporate qualitative methods followed by a clinical validation study to establish psychometric properties, which have been used for both adult and child populations (Kohrt et al., 2011; Van Ommeren et al., 1999). This transcultural translation approach involves qualitative interviews including focus group discussions with target groups and cognitive interviewing, accompanied by a series of translations and back-translations, followed by a clinical validation study.

Regarding clinical services in global mental health, the cultural formulation interview (CFI; **EXHIBIT 10-1**) in DSM-5 is a new tool that can be applied to any patient or population and assures that cultural factors are integrated in diagnosis, treatment planning, and delivering care (APA, 2013). The CFI was developed to assess five components: (1) cultural identity of the individual, (2) cultural explanations for an illness, (3) cultural factors contributing to psychosocial environment and functioning, and (4) cultural factors influencing the clinician–patient/client relationship. Value must be placed on both folk beliefs about mental disorder and the biomedical system of psychiatry. It is important to investigate patients' "explanatory models"—that is, how patients understand their problems, including their nature, origins, consequences, and remedies, as these

EXHIBIT 10-1 Cultural Formulation Interview in DSM-5

This CFI is a tool for clinicians and treatment teams to improve mental health services by assuring that cultural factors are integrated into diagnoses, treatment planning, and delivery of care. The CFI includes four components:

- Cultural definition of the problem: Explanatory models including prominent idioms of distress, reasons for treatment seeking, and impact on functioning.
- Cultural perceptions of the cause, context, and support: Cultural models of causation, impact on and influence of one's social network, culturally relevant interpretations of social stressors, and cultural identity of the individual.
- Cultural factors affecting self-coping and past help-seeking: Self-coping, past help seeking, and prior barriers to care and recovery.
- Cultural factors affecting current help seeking: Patient preferences related to social networks and religion, and clinician–patient relationship factors. The provider must identify differences and similarities in cultural and social status that might influence diagnosis and treatment.

Modified from American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders: DSM-5. Washington, DC: American Psychiatric Publishers, Incorporated.

aspects of understanding can radically alter patient–provider negotiations over appropriate treatment (Kleinman, 1988). The CFI can be used at individual, family, and community levels to inform clinical services and public health efforts in global mental health.

Ultimately, findings across diverse disciplinary approaches increasingly point toward commonalities for understanding mental disorder across cultural groups. As noted earlier, common manifestations and pathways for healing have been identified. Moreover, studies that have shown differences across groups have often shed light on the limitations of psychiatric practices that need to be changed to improve care in high-resource, Western cultural settings as well as in LMIC contexts. This point not only highlights the importance of cross-cultural work for global mental health in low-resource, non-Western settings, but also suggests how global mental health research can better inform mental health categorizations in high-resource, high-researched settings.

The Determinants of Mental Disorders

The etiology of mental disorders comprises a complex interplay among biological factors, most notably genetic predisposition, developmental factors, and psychosocial factors. Until the 1950s, the dominant notions about the etiology of mental disorders were hereditary and environmental in nature. For example, schizophrenia was attributed to abnormal parenting, and obsessive–compulsive disorder to "anal aggression." Beginning in the 1960s, however, a more balanced view emerged. Consensus was reached that both environmental and genetic influences contribute to the development of mental disorders. Psychotic conditions, which historically have been thought to arise predominantly from biological risk factors, have increasingly been shown to be strongly influenced by social determinants (Kirkbride et al., 2006; McGrath et al., 2004). The observation that many disorders have their onset in childhood and youth have placed focus on the developmental origins of these disorders (e.g., related to brain development during childhood and adolescence).

Although risk factors for the majority of mental disorders present themselves in childhood and adolescence, the major burden of mental disorders is found in early adulthood (Murray et al., 2012). Indeed, a range of childhood adversities, such as neglect and parental mental illness, are strongly associated with numerous mental disorders in later life, presumably due to the impact of "toxic stress" on the developing brain. These factors can be divided into risk factors and protective factors. Risk factors make it more likely that an individual will develop mental health difficulties, whereas protective factors mediate and reduce the effects of risk exposure. Risk and protective factors can exist in the biological, psychological, and social domains (**TABLE 10-3**). Many risk factors for mental disorder also predispose persons to a host of physical health and social problems. Adverse childhood experiences (ACEs), such as child abuse and domestic violence, loss of caregivers, nutritional deprivation, and severe childhood illness, are exposures that increase the risk of physical diseases (e.g., cardiovascular and metabolic disease), mental disorders (e.g., suicide, substance abuse, and common mental disorders), and early mortality (Anda et al., 2006; Van Niel, Pachter, Wade, Felitti, & Stein, 2014).

The question was therefore changed from *which* factors were relevant for a specific disorder, to *how much* each contributed to the condition. Ultimately, this question, too, proved to be based on an incorrect assumption—namely, that the environmental, developmental, and genetic factors exert their influences in an additive and independent manner. There is now recognition that they exert their influences in an interactive manner, which develops over the life course. Contemporary scientists are attempting to address the question of *how* they interact.

The Social Determinants of Mental Disorders

The social determinants of mental disorders may be organized into five broad domains (Lund, De Silva, & Stansfeld, 2014):

- The economic domain, which includes poverty, income, consumption, assets, income inequality, and employment
- The social domain, which includes social capital, social cohesion, and education
- The demographic domain, which includes age, gender, and ethnicity
- The neighborhood domain, which includes the area-level social and economic arrangement of people's lives such as housing, water and sanitation, and transport
- The environmental events domain, which includes natural disasters, such as floods, earthquakes, and climate change, as well as civil conflict and forced migration

Each of these domains exerts its influence on population mental health by means of both distal factors

TABLE 10-3 Selected Risk and Protective Factors for Mental Health

Domain	Risk Factors	Protective Factors
Biological	■ Exposure to toxins (e.g., tobacco and alcohol) during pregnancy ■ Genetic tendency to psychiatric disorder ■ Head trauma ■ HIV/AIDS and other physical illnesses	■ Age-appropriate physical development ■ Good physical health ■ Services provided at mother–baby clinics
Psychological	■ Maladaptive personality traits ■ Effects of emotional, physical and sexual abuse, and neglect	■ Ability to learn from experiences ■ Good self-esteem ■ High level of problem-solving ability ■ Social skills
Social		
Family	■ Divorce ■ Family conflict ■ Poor family discipline ■ Poor family management ■ No family	■ Family attachment ■ Opportunities for positive involvement in family ■ Rewards for involvement in family
School or workplace	■ Failure to perform at the expected level ■ Low degree of commitment to school or workplace ■ Inadequate/inappropriate educational provision or training opportunities ■ Experiences of bullying and victimization	■ Opportunities for involvement in school or occupational activities ■ Supportive, stimulating school environment that is tailored to children's developmental needs
Community	■ Community disorganization ■ Effects of discrimination ■ Exposure to violence ■ Social conflict and migration ■ Poverty ■ Transitions (e.g., urbanization)	■ Connectedness to community ■ Opportunities for constructive use of leisure ■ Positive cultural experiences ■ Positive role models ■ Rewards for community involvement

Modified from World Health Organization (WHO). (2015). *Child and adolescent mental health policies and plans*. Retrieved from http://www.who.int/mental_health/policy/Childado_mh_module.pdf

("upstream" social and structural factors, such as social policy, macro-economic trends, and environmental disasters) and proximal factors (the manner in which these distal factors are experienced by individuals and their families, such as employment, housing quality, and trauma).

Environmental Events Domain

According to the United Nations High Commissioner for Refugees (http://www.unhcr.org/4981c3dc2.html), as of 2007 an estimated 11.4 million refugees had fled their own countries, another 13.7 million were internally displaced, and 2.9 million were not considered citizens of any state. Many of these refugees have experienced enormous trauma in the form of violence, crime, or other humiliations; physical injury; economic dispossession; and disruption of family and community structures. Thus, the rates of mental disorders among these people would be expected to be at least as high as—and probably higher than—those for migrants in general. A study of more than 3,000 respondents

from postconflict communities in Algeria, Cambodia, Ethiopia, and Palestine found that PTSD was the most likely MNS disorder in individuals exposed to violence associated with armed conflict (de Jong, Komproe, & Ommeren, 2003). Other mental health consequences included mood and anxiety disorders.

In addition to exposure to trauma, a number of other factors may predispose refugees and immigrants to mental disorders, such as marginalization and minority status, socioeconomic disadvantage, poor physical health, the loss of social support systems, and cultural alienation in the new society. For illegal immigrants, there is also the constant fear of being found out and deported; as a consequence, access to possible sources of help is severely limited. In discussing these issues, it is relevant to note that the universal application of the concept of trauma-related mental disorders (in particular, PTSD) has been criticized because it is itself based on culturally influenced notions of how a person is supposed to react to trauma (see the earlier discussion on culture and mental disorders). While consensus exists that trauma does negatively affect a person's mental health, the question of whether this negative impact should be conceptualized in psychiatric terms (with the concomitant implications of diagnosis and treatment) or in social and cultural terms remains unresolved.

Economic Domain

There is now a substantial body of evidence demonstrating the relationship between poverty and socioeconomic inequalities with mental disorders. In the United Kingdom, for example, evidence has pointed to an association between a low standard of living and the prevalence of depression (Weich & Lewis, 1998). A systematic review located 115 studies in which the associations between poverty and anxiety and mood disorders in 36 LMICs were examined (Lund et al., 2010). Most of these studies reported positive associations between a range of poverty indicators and anxiety and mood disorders. Multivariate analyses showed that in community-based studies, 79%, 15%, and 6% reported positive, null, and negative associations, respectively. A robust association was found between anxiety and mood disorders and education, food insecurity, housing, social class, socioeconomic status, and financial stress. By comparison, the associations between anxiety and mood disorders and income, employment, and consumption were relatively inconsistent. People living in conditions of poverty are also at greater risk for physical health problems, and abundant evidence demonstrates the high degree of comorbidity between physical and mental disorders (Prince et al., 2007). Studies in developed countries have shown that mortality and morbidity rates are more affected by relative, rather than absolute, living standards. A survey in the United States, for example, showed an independent association between low income and living in income-unequal states with depression in women (Kahn, Wise, Kennedy, & Kawachi, 2000). This finding suggests that it is not just having low income, but having low income in relation to others, that increases risk for depression.

Social Domain

The association between poverty and poor mental health may be mediated both by individual psychological factors, such as low self-esteem and frustration, and by a breakdown in structural factors in the community, such as less social cohesion and poorer infrastructure. Reduced prevalence of common mental disorders has been found in populations with higher individual cognitive and ecological social capital (Ehsan & De Silva, 2015). The lack of social support and the breakdown of kinship structures may be important stressors for the millions of migrant laborers in the urban centers of Asia, Africa, and South America, as well as for the millions of dependents who are left behind in rural areas and whose primary source of income consists of the remittances that their relatives send from distant cities. In high-income countries, increased mobility of labor has reduced family ties and also led to the decline of the extended family.

The social consequences of low educational levels are obvious, especially in LMICs that are facing a growing lack of security for employees as those countries' economies are reformed. Lack of secondary education may produce a diminished opportunity for persons who are depressed—especially women—to access resources to improve their situation (Patel, Araya, de Lima, Ludermir, & Todd, 1999). Education also plays an important protective role: A consistent finding across countries is that higher educational levels are associated with reduced prevalence of common mental disorders (Lund et al., 2010), and in the long term education is protective in reducing subsequent late-life depression (Chang-Quan, Zheng-Rong, Yong-Hong, Yi-Zhou, & Qing-Xiu, 2010).

Demographic Domain

Gender is a major determinant of mental health. As described later, the excess prevalence of depression for women has been demonstrated in most community-based studies in all regions of the world.

Women are disproportionately affected by the burden of poverty, which in turn may influence their vulnerability to depression. Women are far more likely to be victims of violence in their homes; women who experience physical violence by an intimate partner are significantly more likely to suffer depression, abuse drugs, or attempt suicide. Women who are sexually abused as children are significantly more likely to suffer depression in adulthood, and experience of sexual and other forms of violence in youth is associated with depression in adolescence.

As an example of how gender influences the risk of depression, a study with women in low-income townships of Harare, Zimbabwe, revealed the high proportion of events involving humiliation and entrapment that were related to marital crises such as being deserted by husbands and left to care for children, premature death, illness of family members, and severe financial difficulties occurring in the absence of an adequate welfare safety net. As another example, studies in South Asia have shown that the culturally determined value placed on boys (as compared to girls) adversely influences maternal mental health. The risk for postnatal depression was elevated in mothers who had a girl child, especially if the desired sex was a boy or if the mother already had living girl children (Patel, Rodrigues, & DeSouza, 2002).

The excess prevalence of alcohol abuse for men has been demonstrated in every community-based study from every region of the world although the disparities are greatest in LMICs. The wide sex differences in alcohol abuse in Latin American countries and the Caribbean have been attributed to a number of gender factors (Pyne, Claeson, & Correia, 2002). Women, for example, face strict social scrutiny about many behaviors, drinking among them. Men's consumption of alcohol takes place in the public realm, whereas women's drinking more often occurs in private. Drinking among men has social meanings, such as maintaining friendships, whereas refusing a drink can imply lack of trust and denial of mutual respect. At the other extreme, intoxication of men is more socially acceptable than that of women; indeed, women often tolerate their male partners' intoxication as being a "natural" condition of manhood. Drinking and drunkenness are more often perceived to be consistent with gendered notions of masculinity; thus, men who conform closely to cultural norms are more likely to drink. Drinking may also be a coping strategy when men face serious life difficulties, such as unemployment, and are unable to live up to the traditional expectations.

The evidence that gender plays a role in eating disorders stems from two observations: (1) the enormous sex difference in prevalence (females with these disorders vastly outnumber men with the same conditions) and (2) the very low rates of these disorders in cultures that have been relatively immune to the media-driven creation of the ideal body image for women. A study from Fiji demonstrated that the introduction of television in a media-naïve non-Westernized population is associated with a rise in attitudes favoring thinner body image and self-induced vomiting in girls (Becker, Burwell, Gilman, Herzog, & Hamburg, 2002), lending credence to the theory that the emphasis on women's thinness by the media and fashion industries is now leading to a rise in disordered eating in societies that, through the forces of globalization, are being increasingly influenced by Western imagery and values.

Another demographic characteristic associated with greater risk of mental disorders is belonging to an ethnic minority population, which has been shown to be associated with an increased risk of a range of mental disorders, including psychosis, anxiety, and depression (Veling, 2013). This relationship is attributed to a range of mechanisms related to discrimination and the traumatic experiences associated with dislocation and migration. Indigenous communities represent a particularly disadvantaged group for a range of social, historical, and political reasons, and they bear a disproportionate share of the burden of mental disorders (**EXHIBIT 10-2**).

Neighborhood Domain

There is now robust evidence indicating that the characteristics of geographical areas have an important effect on population mental health, independent of individual-level characteristics such as poverty, educational level, and gender. In the current context of large-scale global urbanization, the conditions of urban slums therefore have an important influence on mental health. For example, low neighborhood socioeconomic status has been significantly associated with worse mental health, independently of individual-level socioeconomic status (Truong & Ma, 2006).

The Burden of Mental Disorders

This section describes how the burden of mental disorders (including neurological and substance use disorders) can be calculated using disability-adjusted life-years, discusses the limitations inherent in these measures, presents the burden estimates for 2015, and finally identifies important impacts of mental disorders that are not captured in burden of disease estimates.

EXHIBIT 10-2 The Mental Health of Indigenous Peoples

As many as 370 million indigenous persons may be living in approximately 5,000 distinct groups in more than 70 countries. They exhibit a wide diversity of lifestyles, cultures, social organization, histories, and political realities. Nevertheless, they share certain historical and political realities, including being subject to violence and genocide, depopulation from infectious diseases such as smallpox and measles, dislocation from their traditional lands, extreme poverty due to the destruction of their subsistence economies, and state-organized attempts to repress and eradicate their cultures. Given this history, it is not surprising that the indigenous peoples of the world are currently experiencing relatively high rates of depression, alcoholism, and suicide, as well as high rates of infectious diseases and relatively short life expectancies (Anderson et al., 2016).

The case of the indigenous communities of Australia serves to illustrate the confluence of these historical, political, social, and economic forces and contributes to our understanding of why the rates of mental disorders are higher among indigenous peoples. The indigenous peoples of Australia had a diversity of cultures dating back at least 40,000 years before the arrival of European settlers slightly more than 200 years ago. These societies had rich cultural belief systems that attributed spiritual importance to land. Social relationships were governed by codes of behavior, and local taxonomies of illness guided the treatment of health problems. The brutal history of colonization that ultimately led to the destruction and devastation of hundreds of indigenous groups, each with a distinct language and lineage, was marked by a number of severe social adversities. Notable among these were exposure to new diseases, removal from traditional lands, enslavement on white farms, imprisonment without trials, denial of basic political rights, brutal violation of human rights, sexual abuse of women, and, perhaps most tragic of all, the "stolen generations"—the children who were forcibly removed from their parents and fostered by white families in an effort to "breed" out the native population.

Among the indigenous peoples of Australia, the consequences of this history are reflected in socioeconomic, psychosocial, and health indicators of all kinds:

- High rates of unemployment, low levels of income, and poor educational status
- Age-specific mortality rates two to seven times higher, and life expectancies more than 15 years shorter, than those of the general population
- High levels of alcoholism and suicide

It is impossible to interpret the poor mental health experienced by these communities without considering the social and historical contexts of the systematic abuse of aboriginal communities. *Ways Forward*, the Australian national inquiry into indigenous mental health conducted in the early 1990s, prioritized holistic conceptions of emotional and social well-being among these groups. From these developments, greater emphasis has been given to providing access to culturally appropriate services within mainstream healthcare settings.

To understand the contributions that mental disorders make to the global burden of disease compared to other diseases and injuries, it is necessary to measure the impact of all disorders in the same way. A common metric is needed for measurement; otherwise, it is very difficult to quantify the burden imposed by cancer, for example, compared to that imposed by depression. A commonly used metric is the disability-adjusted life-year (DALY), a time-based metric that combines the years a person lives with disability (YLD) caused by a particular disorder with the years lost from life owing to a disorder (YLL). The DALY has been used by the World Bank, WHO, and in the Global Burden of Diseases, Injuries, and Risk Factors Studies to measure the gap between the health of the population at a point in time and living to a defined life expectancy in full health. Since DALYs, YLDs, and YLLs were first estimated for the year 1990, the data sources used have been greatly expanded and the methods applied to generate the estimates improved. The changes in methods have included dealing with inconsistent coding of deaths over time and between countries, introduction of disability weights from population-representative surveys, more precise severity distributions, introduction of comorbidity adjustments, and propagation of 95% uncertainty intervals around all burden estimates. As new methods are introduced, it is necessary to recalculate GBD estimates from previous years so as to make comparison with previously published data possible. A detailed description of how these burden of disease metrics are estimated for mental disorders can be found elsewhere (GBD 2015 DALYs and HALE Collaborators, 2016).

DALYs are derived by summing YLDs and YLLs for each disorder, location, age group, sex, and year, and are now provided for more than 300 diseases and injuries. In interpreting the burden of mental disorders, one must recognize that the mortality-associated (YLL) burden estimates for mental disorders in the GBD studies should not be interpreted as suggesting

that premature death in people with MNS disorders is relatively inconsequential. Premature mortality has been repeatedly shown to be significant in populations with mental disorders (Walker, McGee, & Druss, 2015), with suicide being a major contributor and large contributions coming from cardiovascular disease and cancer (Lawrence, Hancock, & Kisely, 2013). Even in high-income countries, people with mental disorders often do not receive preventive services, such as immunizations, cancer screenings, and tobacco counseling (Druss, Rosenheck, Desai, & Perlin, 2002), and they often receive a lower quality of care for their medical conditions (Björkenstam et al., 2012; Laursen, Nordentoft, & Mortensen, 2014). Excess mortality estimates can be generated based on the GBD data; while MNS disorders were directly responsible for only 840,000 deaths in 2010, individuals with these conditions were estimated to have more than 13 million excess deaths using natural history models (Charlson, Diminic, Lund, Degenhardt, & Whiteford, 2014).

Notwithstanding these caveats, the GBD estimates show mental disorders contribute substantially to the global burden of disease, with mental and substance use disorders being the leading cause of disability globally (Whiteford et al., 2013; Whiteford, Ferrari, Degenhardt, Feigin, & Vos, 2015). The burden peaks in early adulthood for mental and substance use disorders, but is more consistent across age for neurological disorders. Females accounted for more DALYs for all mental and neurological disorders, except for mental disorders occurring in childhood, schizophrenia, substance use disorders, Parkinson's disease, and epilepsy, where males accounted for more DALYs. Overall DALYs are highest in the Eastern Europe/Central Asia regions and lowest in the East Asia/Pacific regions. The relative proportion of DALYs from MNS disorders within the overall disease burden has been estimated to be 1.6 times higher in high-income regions than in lower-income regions, largely due to the relatively higher burden of other health disorders such as infectious and perinatal diseases in LMICs. Because of their larger population, however, most of the global burden from MNS disorders is found in LMICs.

The burden estimates for 2015 are provided in **TABLE 10-4** and **FIGURES 10-2** and **10-3**. Table 10-4

TABLE 10-4 Total DALYs (Numbers and Proportions) Attributable to Mental, Neurological, and Substance Use Disorders in 2015

Cause	DALYs per 100,000 Population	Proportion of All-Cause DALYs in GBD	Proportion of Mental, Neurological, and Substance Use Disorder DALYs
Neurological Disorders			
Alzheimer's disease and other dementias	238	1.0%	9.5%
Parkinson's disease	21	0.1%	0.8%
Epilepsy	124	0.5%	5.0%
Multiple sclerosis	12	0.1%	0.5%
Migraine	329	1.3%	13.2%
Tension-type headache	23	0.1%	0.9%
Medication overuse headache	92	0.4%	3.7%
Motor neuron disease	9	0.04%	0.4%
Other neurological disorders	24	0.1%	0.9%

(continues)

TABLE 10-4 Total DALYs (Numbers and Proportions) Attributable to Mental, Neurological, and Substance Use Disorders in 2015 *(continued)*

Cause	DALYs per 100,000 Population	Proportion of All-Cause DALYs in GBD	Proportion of Mental, Neurological, and Substance Use Disorder DALYs
Substance Use Disorders			
Alcohol use disorders	112	0.5%	4.5%
Opioid use disorders	121	0.5%	4.8%
Cocaine use disorders	10	0.0%	0.4%
Amphetamine use disorders	14	0.1%	0.6%
Cannabis use disorders	6	0.02%	0.2%
Other drug use disorders	19	0.1%	0.7%
Mental Disorders			
Schizophrenia	155	0.6%	6.2%
Major depressive disorder	442	1.8%	17.7%
Dysthymia	100	0.4%	4.0%
Bipolar disorder	90	0.4%	3.6%
Anxiety disorders	246	1.0%	9.9%
Anorexia nervosa	7	0.03%	0.3%
Bulimia nervosa	8	0.03%	0.3%
Autism	63	0.3%	2.5%
Asperger's syndrome and other autistic spectrum disorders	37	0.2%	1.5%
Attention-deficit/hyperactivity disorder	6	0.03%	0.2%
Conduct disorder	58	0.2%	2.3%
Idiopathic developmental intellectual disability	34	0.1%	1.4%
Other mental and substance use disorders	96	0.4%	3.9%

Data from *Global burden of disease study 2015.* Retrieved from http://ghdx.healthdata.org/gbd-2015

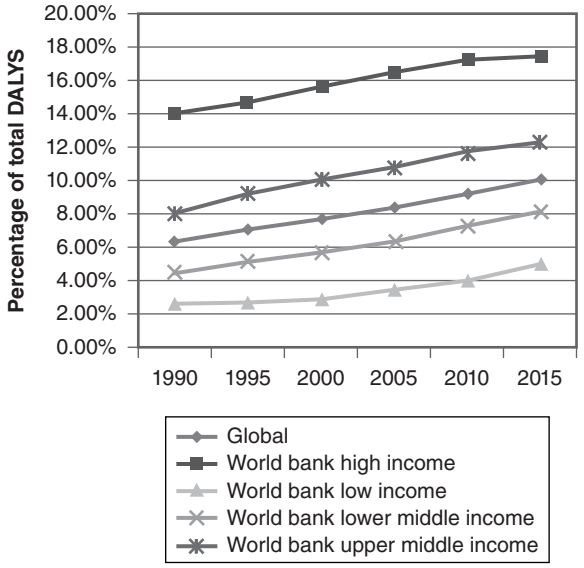


FIGURE 10-2 Mental, neurological, and substance use disorders, both sexes, all ages.

Data from *Global burden of disease study 2015*. Retrieved from http://ghdx.healthdata.org/gbd-2015

provides a global overview of the total burden from each MNS disorder included in the GBD study. The leading contributors to this burden are major depressive disorder, which accounts for almost 18% of the overall burden in the total group, with anxiety disorders and dementia both contributing close to 10%. Figure 10-2 shows that the proportionate burden from MNS disorders increased substantially in all countries, irrespective of their economic development status, between 1990 and 2015. For most disorders, the increased burden is being largely driven by population growth and aging. In areas of the world undergoing major demographic change, such as sub-Saharan Africa, where the population is expected to double by 2050, projections suggest there will be an increase of 130% in the burden of mental and substance use disorders (Charlson et al., 2014). Figure 10-3 shows that the burden is greatest in the adolescent to middle adulthood period of life, when productivity is the greatest.

To be included in the GBD data, disorders must meet the threshold for diagnosis using ICD or DSM criteria. While the application of these diagnostic criteria may bias estimates downward in regard to the burden of mental disorders in some cultures, the adoption of common case definitions is necessary to promote international comparison (Whiteford, Ferrari, & Degenhardt, 2016). Another challenge for generating estimates of mental disorder burden, and subsequently investigating trends over time and differences between countries, is the limited coverage of epidemiologic data. This is particularly problematic in LMICs and in subnational locations. As a result, estimates often have large uncertainty intervals that can mask true variations in burden between countries and over time. Furthermore, it is necessary to generate better data on the relationships between risk factors, such as economic deprivation, and the variations in burden that might be explained by these risk factors.

Depressive and Anxiety Disorders

In 2015, the global age-standardized prevalence was 3% for major depressive disorder and 1.4% for

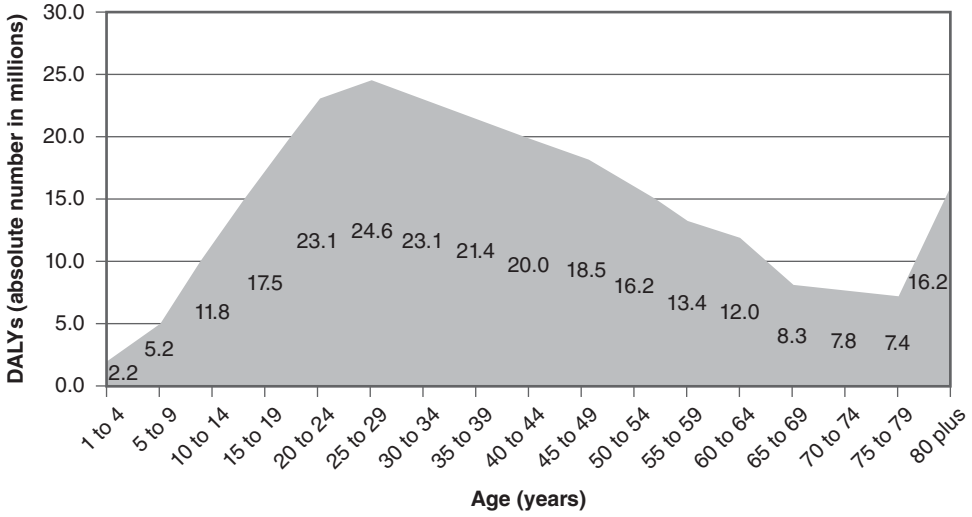


FIGURE 10-3 Global burden of mental, neurological, and substance use disorders.

Data from *Global burden of disease study 2015*. Retrieved from http://ghdx.healthdata.org/gbd-2015

dysthymia, making depressive disorders the largest contributor to the burden of disease, as measured by DALYs in the mental disorder group. The global prevalence is higher in females (5.1%) than in males (3.7%), and in countries where war and conflict are more prevalent. For example, Uganda had the highest rate of depressive disorder DALYs—although with the wide uncertainty estimates, this was not significantly different from the global mean. There is also a clear socioeconomic gradient in prevalence in many countries, with individuals having lower socioeconomic status also having increased risk for depression and anxiety disorders (Lund et al., 2010). Individual anxiety disorders are common and frequently co-occur with each other.

To prevent over-counting, GBD 2015 provides estimates for any anxiety disorder—a grouping that includes generalized anxiety disorder, panic disorder, and phobic disorders (agoraphobia, social phobia, and specific phobias) (Baxter, Vos, Scott, Ferrari, & Whiteford, 2014). It also includes obsessive–compulsive disorder, post-traumatic stress disorder, and acute stress disorders, although these are no longer classified as anxiety disorders in DSM-5. In 2015, the global age-standardized prevalence for all anxiety disorders combined was 3.6% (4.6% for females and 2.6% for males), making these disorders the second largest contributor to the burden of disease imposed by MNS disorders, as measured by DALYs. As is true for depression, there is considerable variability in the estimates, with anxiety disorder prevalence being impacted by factors such as sex, age, culture, conflict, urbanicity, and economic status (Baxter, Scott, Vos, & Whiteford, 2013). For example, a systematic review and meta-analysis of post-traumatic stress disorder alone in the general population in areas exposed to conflict found the prevalence to be 12.9%, after controlling for an extensive range of covariates (Charlson, Baxter, Cheng, Shidhaye, & Whiteford, 2016).

Schizophrenia

The age-standardized prevalence of schizophrenia was estimated in the GBD studies to be approximately 3% globally. This prevalence did not change from 1990 to 2015; however, the number of people with schizophrenia increased from 14.4 million to 23.4 million during this period, due primarily to population growth and aging. In 2015, schizophrenia contributed 15.5 million DALYs to the overall burden of disease globally, with roughly equal burdens noted in males and in females. Substantially higher prevalent cases and DALYs of schizophrenia were found lower-income regions of the world—a reflection of population size, but also coinciding with the highest treatment gaps. For example, China had the highest rate of schizophrenia DALYs, but this was closely followed by the U.S. rate, and neither country's rate was significantly higher than the global mean.

Bipolar Disorder

The global age-standardized prevalence of bipolar disorder was estimated at 0.5% across both sexes. As with the majority of mental disorders, the prevalence of bipolar disorder remained unchanged between 1990 and 2015, but due to population growth and changing age structures, the number of people with bipolar disorder increased from approximately 30 million in 1990 to 44 million in 2015. In terms of burden, bipolar disorder was responsible for 9 million DALYs, equating to 0.35% of all global DALYs in 2015. The DALY rates were consistent across the globe, with the difference between the lowest country estimate (China) and the highest country estimate (New Zealand) not differing significantly.

Substance Use Disorders

The burden of substance use disorders varies considerably across and within countries, between subgroups in a given population, and according to the substance being abused. Unlike other mental and neurological disorders, variations in substance use burden are influenced by additional factors, such as supply and availability of drugs, which can change prevalence and, in turn, burden. If dependence on tobacco (arguably the most common substance to be abused) is excluded, alcohol use disorders (AUD) are the most common substance abuse disorder in most countries. While AUD can refer to the entire range of health conditions associated with drinking alcohol above the recommended limit established by WHO, alcohol use disorders in the GBD studies refer only to alcohol dependence as defined by the ICD. Illicit drugs are those substances whose use outside medical settings has been prohibited under international and national control systems. In the GBD study, the drugs whose burden is estimated are opioids (including heroin), cannabis, amphetamines, and cocaine.

Globally, AUD were the most prevalent substance use disorders, with 63.5 million estimated cases in 2015. Cannabis dependence and opioid dependence were the most common illicit drug use disorders (19.8 million cases and 16.7 million cases, respectively). There were substantial geographic variations in the burden of disease and contrasting patterns for the association between total alcohol and illicit drug-attributable burden and per

capita income. Alcohol burden was highest in LMICs (with Russia being the highest), while the burden from illicit drugs was highest in Eastern European countries and the United States.

Childhood Disorders

The majority of mental disorders have their initial onset during childhood and adolescence, making this a crucial period for early intervention and identification. While substance use disorders tend to demonstrate a sharp increase in prevalence (and therefore burden) from late adolescence onward, certain mental and neurological disorders demonstrate significant burden between ages 5 to 14 years. Across both sexes, conduct disorder, depressive disorders, anxiety disorders, and autism spectrum disorders are responsible for almost 7.5% of all DALYs worldwide in this age group. For neurological disorders, migraine and epilepsy make up the majority of the burden of disease for children and adolescents (1.8% and 1.4%, respectively).

The implications of these estimates are significant, given that 85% of children and adolescents live in LMICs that generally have few or no child- and adolescent-specific mental health services. Furthermore, as infectious diseases continue to be successfully prevented and treated, more infants in these regions will survive into childhood and adolescence and, therefore, reach the ages where the incidence of mental disorders is higher. As such, it is important for these disorders to be recognized as significant contributors to the burden of disease in children and adolescents and for health resources to be allocated based on both current and future needs.

Dementia

The population of almost all countries is aging, and the size of the older population is growing fastest in low-income countries. Between 2015 and 2050, the number of older people living in higher-income countries is forecast to increase by 56%, compared with 138% in upper middle-income countries, 185% in lower middle-income countries, and 239% in low-income countries (https://www.alz.co.uk/research/WorldAlzheimerReport2015.pdf). Dementia is becoming a major contributor to global burden, with the total DALYs doubling for both men and women between 1990 and 2015.

Wider Health Impact of Mental Disorders

The limitations imposed by the GBD study methodology means the full impact of mental disorders is not captured in the DALY metric. Debate has arisen about where and how the burden of disease estimates related to mental disorders could be improved (Vigo, Thornicroft, & Atun, 2016; Whiteford, Ferrari, & Vos, 2016) in terms of measuring direct burden. One area where efforts to improve burden of disease estimates have been pursued is the relatively small number of deaths attributed to mental disorders. Suicide, for example, is recorded under "injuries" in the ICD cause-of-death coding guidelines. The inclusion of suicide deaths attributable to mental disorders into DALYs would have increased the burden of mental and substance use disorders in 2010 from 7.4% to 8.3% and increased the global ranking of these disorders from the fifth to the third leading cause of burden of disease (Ferrari et al., 2014).

The contribution that disorders make, as risk factors for other health outcomes, can also be dealt with through the comparative risk assessment (CRA) component of the burden of disease studies. The CRA quantifies the burden attributable to each risk factor exposure compared to an alternative (counterfactual) exposure distribution (GBD 2015 Risk Factors Collaborators, 2016). Mental disorders are considered risk factors if associated with elevated risk of mortality or disability from other diseases or injuries, though this area of the GBD studies is underdeveloped compared with the measurement of the direct burden of disease.

The burden of disease attributable to major depression as a risk factor for suicide and ischemic heart disease has been estimated (Charlson, Stapelberg, Baxter, & Whiteford, 2011; Li, Page, Martin, & Taylor, 2011). In 2010, major depression explained a further 16 million DALYs when it was considered as a risk factor for suicide. Overall, nearly half of DALYs originally allocated to suicide (included as intentional injuries in the GBD cause of death list) can be reattributed to major depression alone (Ferrari et al., 2013). Alcohol use and injecting-drug use as risk factors are now included in burden of disease estimates. In 2013, injecting-drug use was estimated to cause 39% of DALYs due to hepatitis C, 4% of DALYs due to HIV, and 1% of DALYs due to hepatitis B (Degenhardt et al., 2016). Injecting-drug use's contribution to the HIV burden of disease was highest in LMICs, and injecting-drug use's contribution to the hepatitis C burden of disease was highest in high-income countries. In 2015, DALYs attributed to alcohol use disorders were concentrated among transport injuries, cirrhosis, and cancers. Using DALYs per 100,000 population, Eastern Europe and southern sub-Saharan Africa have the highest estimated alcohol-attributable burden of disease, while Eastern Europe and North America have the highest illicit drug–attributable burden of disease.

This CRA area needs to be greatly expanded to capture the comorbidity between mental and other disorders. There is strong comorbidity between diet, diabetes, and mood disorders (Dipnall et al., 2015; Grigsby, Anderson, Freedland, Clouse, & Lustman, 2002), for example, and depression may increase the risk for diabetes (Rotella & Mannucci, 2013). Between 10% and 20% of people with schizophrenia have type 2 diabetes (Holt, Bushe, & Citrome, 2005). Depression is a risk factor for both ischemic heart disease (Charlson et al., 2013) and stroke (Pan, Sun, Okereke, Rexrode, & Hu, 2011). There is also strong evidence of the increased risk for depression *after* myocardial infarction (Strik, Lousberg, Cheriex, & Honig, 2004) and stroke (Aben et al., 2006). Infectious disease, cancer, diabetes, and cardiovascular disease are all more common with excess alcohol consumption (Rehm, 2011), with illicit drug use (Degenhardt & Hall, 2012), and in those persons with severe mental disorders (Hert et al., 2011).

In addition, anxiety and mood disorders and other common mental disorders (often called "somatoform disorders") can present in primary health care/general medical settings with physical symptoms (Escobar, Waitzkin, Silver, Gara, & Holman, 1998; Gureje, Simon, Ustun, & Goldberg, 1997). Such symptoms, which are sometimes termed "medically unexplained" because they cannot be attributed to physical disorders (Creed et al., 2012), are associated with considerable disability and high levels of help seeking and associated healthcare costs (Konnopka et al., 2012).

The coexistence of mental and physical disorders is also associated with worse outcomes of the physical disorder, through a variety of mechanisms. For example, depression after acute myocardial infarction has been associated with fatal and nonfatal cardiovascular events, adverse overall health status, and increased costs (Lichtman et al., 2014). Similarly, depression is associated with an increased risk of poor glycemic control and, therefore, increased mortality in people with diabetes (Lin et al., 2010). Similar findings have been reported for patients with HIV/AIDS, in whom depression is recognized as a predictor of worse clinical outcomes and poorer quality of life, often associated with a reduction in medication adherence (Nanni, Caruso, Mitchell, Meggiolaro, & Grassi, 2015). Clear recommendations have been made for improving the treatment of people with comorbid mental and physical disorders, especially those with severe mental disorders (Hert et al., 2011).

Another relationship that needs to be addressed is the strong association of mental disorders with a range of women's health concerns. Depression is strongly associated with dysmenorrhea, dyspareunia, and pelvic pain (Latthe, Mignini, Gray, Hills, & Khan, 2006). In some cultures in Asia, gynecologic complaints, such as abnormal vaginal discharge, are associated with depression (Patel et al., 2006). Maternal mental disorders are associated with a range of adverse outcomes in children, including low birth weight, premature birth, poor infant growth, and emotional, cognitive, and behavioral developmental delays that can persist into adolescence and beyond (Stein et al., 2014).

The Wider Social and Economic Impact of Mental Disorders

Just as one disease burden metric (such as the DALY) cannot capture the full health impact of mental disorders, it also cannot capture all of the social and economic impacts of these disorders on the individual, family, or society. Disease burden estimates do not account for the denial of basic human rights, ranging from limited opportunities for education and employment to torture and denial of freedom, sometimes within healthcare institutions. They also do not account for the time spent by family members in support and caring for the individual with mental disorder, the costs of treatment, or the productivity loss costs to the individual and his or her family.

A study undertaken for the World Economic Forum estimated that the cumulative global impact of mental disorders in terms of lost economic output may amount to $16 trillion over 20 years (Bloom et al., 2011). The Organisation for Economic Co-operation and Development (OECD, 2015) has conservatively estimated that the impact of mental disorders in high-income countries ranges from 2.3% to 4.4% of gross domestic product (GDP). Economic costs attributable to alcohol use and alcohol use disorders alone are estimated to amount to the equivalent of 1.3% to 3.3% of GDP in a range of high-income and middle-income countries, with more than two-thirds of this loss represented by productivity losses (Rehm et al., 2009). The global cost of dementia in 2015 has been estimated at $818 billion, equivalent to 1.09% of global GDP (Alzheimer's Disease International, 2015). Additionally, the social adversities associated with mental disorders, given the large and growing proportions of the global population affected by conflict or displacement due to environmental degradation and climate change, will likely increase the future burden of mental disorders.

In high- and middle-income countries, the productivity losses attributable to mental disorders have been shown to be very significant. As an example, the lost taxation revenue in Australia from people age 45 and older who retire early due to depression has been

estimated at AUD$ 278 million (USD$ 190 million), and income support for those individuals has been estimated at AUD$ 407 million (USD$ 278 million) in 2009 terms. The total impact on GDP of this group for this disorder alone was estimated at AUD$ 1.7 billion (USD$ 1.2 billion) (Schofield et al., 2011). The value of informal caring provided by family members was approximately AUD$13.2 billion (USD$10.3 billion) in 2015 (Diminic et al., 2016), much greater than the AUD$ 8 billion (USD$ 6.3 billion) that the national and state governments spent on mental health services in Australia.

Although the economic burden is large, increased spending within the health sector to increase treatment coverage for mental disorders appears to be more than offset by economic productivity gains and reduced outlays on social and income support (OECD, 2015; Wang et al., 2007). Evidence-based health and social interventions can avert the long-term cost burden and have a broad range of payoffs, both within the public sector and more widely, such as through better educational performance, improved employment/earnings, and reduced crime (Knapp, McDaid, & Parsonage, 2011). The view that treatment of mental disorders in LMICs is prohibitively expensive is a myth that needs to be dispelled. A fully scaled-up package of mental health care in sub-Saharan Africa and south Asia, based on a comparative cost-effectiveness analysis of 44 individual or combined interventions, has been estimated at USD$ 3 to USD$ 4 per person within a population (Chisholm, Naci, Hyder, Tran, & Peden, 2012). Of course, scaled-up treatment of common mental disorders globally can lead to large productivity gains. In a model consisting of 36 countries between 2016 and 2030, USD$ 230 billion was needed for scaled-up treatment of depression and USD$ 169 billion for scaled-up treatment of anxiety disorders (Chisholm, Sweeny, et al., 2016).

Interventions

This section briefly considers the role of mental health policies, human resources for mental health care, and the evidence for the prevention and treatment of mental disorders.

Mental Health Policies and Plans

A mental health policy presents the values, principles, and objectives for improving mental health and reducing the burden of mental disorders in a population. It should define a vision for the future and help to establish a model for action. A policy should be distinguished from a plan, which is a strategy for implementing actions to achieve the objectives of a policy.

In some countries, mental health policies are restricted to psychiatric services. However, a broader scope is preferable—one in which mental health services in general are addressed. These services may include primary care and specialized care, as well as all aspects of intervention—that is, promotion, prevention, treatment, and ongoing care (WHO, 2003c). Policies need to address the coordination between mental health services themselves, as well as between mental health services and other services such as housing, education, and employment. Other key issues that policies should address include financial arrangements for the private and public sectors, expenditure prioritization, and individual and organizational capacity development (WHO, 2003a). Finally, policies need to provide for continuous evaluation of mental health outcomes to ensure that those policies remain appropriate to contemporary circumstances and lead to effective services.

A country's capacity to deliver appropriate mental health services to its population is seriously hampered by the absence of a mental health policy. Thus, it is cause for concern that only 68% of countries (77% of those that responded to the WHO survey) have mental health policies (WHO, 2014). LMICs are less likely to have these kinds of policies: For example, only 71% of countries in Africa have established such policies. This is, however, a substantial improvement in recent years: In 2001, only 50% of African countries in Africa had established mental health policies (WHO, 2001).

Partly in response to this shortcoming, WHO developed the Mental Health Policy and Service Guidance Package during 2000–2005. This package consists of a series of interrelated, user-friendly modules designed to assist with policy development and service planning. One module provides a series of steps that can be taken to develop mental health policies: (1) assess the population's needs; (2) gather evidence for an effective policy; (3) consult and negotiate; (4) exchange ideas with other countries; (5) set out the vision, values, principles, and objectives of the policy; (6) determine areas for action; (7) identify the roles and responsibilities of different sectors; and (8) conduct pilot studies (WHO, 2005b).

Human Resources for Mental Health Care

The implementation of mental health policies and plans depends on both the quantity and the quality of the personnel available to implement interventions. There are vast differences among regions of the world

in terms of the availability of mental health professionals (**TABLE 10-5**). In almost all countries, there is a gap between the supply of personnel and the demand for their services. The deleterious consequences of the low numbers of mental health professionals are magnified when one considers that the distribution of mental health professionals is frequently uneven between countries in each region, and within countries. The number of mental health workers per 100,000 population is considerably higher in urban areas, for example. Also, the available personnel are often not used efficiently, and staff may be demoralized and demotivated.

Mental health programs require a cadre of well-trained mental health specialists, such as psychiatrists, psychologists, social workers, mental health occupational therapists, and psychiatric nurses. They are responsible for functions such as the management of patients with complex conditions, supervision and training of other specialists and generic health workers, research, planning, management, and consultation–liaison. It is vital that specialists stay abreast of modern international developments that are relevant for the functions they perform, including honing their skills of evaluation, capacity building, and supervision (Patel, Simon, Chowdhary, Kaaya, & Araya, 2009). At the same time, the application of such developments should be informed by local research and experience. Training efforts for mental health specialists should occur in parallel with training for generic health workers such as doctors, nurses, and community health workers, who provide the majority of care in LMICs. In keeping with the objectives of WHO's mhGAP program, a major initiative in recent years has been the training and supervision of nonspecialist health providers in the provision of basic mental health services—an approach known as "task sharing" or "task-shifting" (Kakuma et al., 2011). In keeping with the objectives of the Grand Challenges in Global Mental Health initiative (Collins et al., 2011), substantial research funding has been devoted to research in this area in LMICs, and the evidence in support of this approach is presented later in this chapter.

One sector that is particularly important in LMICs is the traditional health sector. In many LMICs, the majority of people seek care from traditional healers before seeing allopathic healers. There are several ways in which such providers can be engaged to provide mental health care. They can work side by side with allopathic mental health services, perhaps even operating from the same premises; they can be trained to recognize mental disorders and refer people suffering from them to allopathic services; and they can be recruited and trained to function as allopathic mental health workers. Whatever arrangements are made at an organizational level, individual mental health services providers should attempt to establish whether their patients are being subjected to any traditional

TABLE 10-5 Median Number of Mental Health Professionals per 100,000 Population in Each WHO Region and in the World

WHO Region	Psychiatrists	Psychiatric Nurses	Psychologists Working in Mental Health	Social Workers Working in Mental Health
Africa	0.1	0.6	0.1	0.1
Americas	1.1	5.3	1.4	0.6
Eastern Mediterranean	0.9	3.1	0.4	0.3
Europe	7.0	24.1	2.7	1.7
Southeast Asia	0.4	2.6	0.1	0.1
Western Pacific	0.9	5.7	0.9	1.5
World	0.9	5.1	0.7	0.4

Data from World Health Organization (WHO). (2014). *Mental health atlas 2014*. Geneva, Switzerland: World Health Organization.

interventions that are harmful. If they are, they should receive education and counseling that aim to reduce exposure to such negative interventions. Conversely, traditional practices that are helpful can be incorporated into allopathic care (Institute of Medicine, 2001).

Prevention and Treatment of Mental, Neurological, and Substance Use Disorders

Interventions for mental disorders have progressed enormously in the past few decades. The *Disease Control Priorities*, third edition (DCP-3), released by the World Bank, provides a synthesis of the evidence base and recommendations for the packages for prevention and treatment of MNS disorders (Patel et al., 2016). These recommendations identify the specific interventions for each group of disorders (**TABLE 10-6**) and the platforms through which they can be scaled up (**TABLE 10-7**).

Population-Level Recommendations

Evidence-based population-level recommendations include legislative measures to restrict access to means of self-harm and suicide (e.g., limiting access to guns and pesticides, such as through the use of lockboxes) and limiting availability of and demand for alcohol (e.g., through taxation and increased prices). Effective interventions, strategies, and policies to prevent and reduce substance use disorders can be categorized as regulatory, community based (including education), and health services based.

Prohibition has been attempted for alcohol products and is currently in place in some countries; this policy is also used with classes of substances including opioids, cannabinoids, and cocaine in most countries. Although prohibition can dramatically reduce substance use disorders in the short term, its costs in terms of civil disobedience and crime are enormous—so much so, that, in general, prohibition is not regarded as an acceptable policy option, with the exception of specific circumstances, such as drinking alcohol and driving (Wolf & Midanik, 2013).

Regulatory interventions include taxation, restrictions on availability, and total bans on all forms of direct and indirect advertising. Increases in alcohol taxes, for example, have been shown to reduce both the prevalence and the consumption of alcohol products. For young people, laws that raise the minimum legal drinking age reduce alcohol sales and problems among young drinkers. Reductions in the hours and days of sale, numbers of alcohol outlets, and restrictions on access to alcohol are all associated with reductions in both alcohol use and alcohol-related problems. A meta-analysis of studies in LMICs found a strong inverse effect between alcohol consumption and alcohol price and/or taxation (Sornpaisarn, Shield, Cohen, Schwartz, & Rehm, 2013), which is comparable to the impact of increasing price/taxes and decreasing consumption in high-income countries.

At the population level, there is also good practice evidence for interventions that seek to raise mental health literacy and reduce stigma and discrimination (Thornicroft et al., 2016). Social contact interventions, in which people engage with persons with mental disorders in recovery, changes attitudes and has the potential to instill behavior change (Corrigan, Morris, Michaels, Rafacz, & Rusch, 2012). Interventions within the criminal justice system are effective means to reduce alcohol and drug use, and to treat behavior disorders in adolescents and psychosis among adults. For example, mental health input in the criminal justice system can prevent the inappropriate imprisonment of people with mental disorders, make treatment for mental disorders available in prisons, and reduce the mental health sequelae of imprisonment for prisoners and their families. In high-income countries, mental health training of police through crisis intervention teams (CIT) reduces incarceration, increases mental health treatment, and reduces violence experienced both by police and by persons with mental disorders (Compton, Broussard, Munetz, Oliva, & Watson, 2011). There is increased attention to CIT and similar models in LMICs. For example, in Liberia, CIT training of police both improved mental health outcomes and reduced use of physical force against persons in acute psychological distress during the 2013–2016 Ebola outbreak (Kohrt, Blasingame, et al., 2015).

Community-Level Recommendations

Community-based mental health care is a major objective of the World Psychiatric Association and other mental health advocates and practitioners (Thornicroft et al., 2010). Life-skills training in schools to build social and emotional competencies is an example of a best practice to reduce the burden of MNS disorders. Systematic reviews of mental health promotion for children and adolescents have shown that school-based programs improve self-esteem, motivation, and self-efficacy (Barry, Clarke, Jenkins, & Patel, 2013; Fazel, Patel, Thomas, & Tol, 2014). Community mobilization and education have been used to prevent substance abuse in many countries. A crucial setting for prevention is in schools, where the goal of most alcohol education programs is to change adolescents' drinking beliefs, attitudes, and behaviors, or to modify factors such as general social skills and self-esteem

TABLE 10-6 Effective Interventions for the Prevention, Treatment, and Care of MNS Disorders

	Type of Disorder	Preventive Interventions	Drug and Physical Interventions	Psychosocial Interventions
Mental Disorders in Adulthood				
Schizophrenia (5.3% of total MNS DALYs)	■ Chronic or relapsing condition characterized by delusions, hallucinations, and disturbed behavior		■ Antipsychotic medication***	■ Family therapy/ support** Community-based rehabilitation* ■ Self-help and support groups*
Mood and anxiety disorders (41.9% of total MNS DALYs)	■ Group of conditions characterized by somatic, emotional, cognitive, and behavioral symptoms; bipolar disorder is associated with episodes of elated and depressed moods	■ Cognitive-behavioral therapy for persons with subthreshold symptoms**	■ Antidepressant, anxiolytic, mood stabilizer, and antipsychotic medications*** ■ Electroconvulsive therapy (ECT) for severe refractory depression**	■ Cognitive-behavioral therapy*** ■ Interpersonal therapy**
Mental and Developmental Disorders in Childhood and Adolescence				
Conduct disorder (2.2% of total MNS DALYs)	■ Pattern of antisocial behaviors that violate the basic rights of others or major age-appropriate societal norms	■ Life skills education to build social and emotional well-being and competencies** ■ Parenting skills training** ■ Maternal mental health interventions**		■ Parenting skills training** ■ Cognitive-behavioral therapy*
Anxiety disorders (2.3% of total MNS DALYs)	■ Excessive or inappropriate fear, with associated behavioral disturbances that impair functioning	■ Parenting skills training** ■ Maternal mental health interventions**		■ Cognitive-behavioral therapy**
Autism (1.6% of total MNS DALYs)	■ Severe impairment in reciprocal social interactions and communication skills, as well as the presence of restricted and stereotypical behaviors			■ Parental education and skills training* ■ Educational support*

Attention-deficit/ hyperactivity disorder (ADHD) (0.2% of total MNS DALYs)	■ Neurodevelopmental disorder characterized by inattention and disorganization, with or without hyperactivity–impulsivity, causing impairment of functioning	■ Psychosocial stimulation of infants and young children*	■ Methylphenidate**	■ Parenting skills training** ■ Cognitive-behavioral therapy**
Intellectual disability (idiopathic) (0.4% of total MNS DALYs)	■ Significantly impaired cognitive functioning and deficits in two or more adaptive behaviors	■ Psychosocial stimulation of infants and young children** ■ Perinatal interventions (e.g., screening for congenital hypothyroidism)** ■ Population-based interventions targeting intellectual disability risk factors (e.g., reducing maternal alcohol use)*		■ Parental education and skills training* ■ Educational support*
Neurological Disorders				
Migraine (8.7% of total MNS DALYs)	■ Episodic attacks where headache and nausea are the most characteristic attack features; the headache itself, lasting for hours to 2–3 days, is typically moderate or severe and likely to be unilateral, pulsating, and aggravated by routine physical activity	■ Prophylactic drug treatment with propranolol or amitriptyline***	■ Drug treatments: Aspirin or one of several other nonsteroidal anti-inflammatory drugs [NSAIDs]***	■ Behavioral and cognitive interventions*
Epilepsy (6.8% of total MNS DALYs)	■ A brain disorder traditionally defined as the occurrence of two unprovoked seizures occurring more than 24 hours apart with an enduring predisposition to generate further seizures	■ Population-based interventions targeting epilepsy risk factors (e.g., preventing head injuries; neurocysticercosis prevention)*	■ Standard antiepileptic medications (phenobarbital, phenytoin, carbamazepine, valproic acid)*** ■ Epilepsy surgery**	
Dementia (4.4% of total MNS DALYs)	■ A neuropsychiatric syndrome characterized by a combination of progressive cognitive impairment, behavioral, and psychological symptoms (BPSD) and functional difficulties	■ Cardiovascular risk factors management (healthy diet, physical activity, tobacco use cessation)*	■ Cholinesterase inhibitors and memantine for cognitive functions; medications for management of BPSD*	■ Caregiver education and support and behavioral training, and environmental modifications ** ■ Interventions to support carers of people with dementia**

(continues)

TABLE 10-6 Effective Interventions for the Prevention, Treatment, and Care of MNS Disorders *(continued)*

	Type of Disorder	Preventive Interventions	Drug and Physical Interventions	Psychosocial Interventions
Substance Use Disorders				
Alcohol use disorders (6.9% of total MNS DALYs)	■ Harmful use: "A pattern of alcohol use that causes damage to physical or mental health" ■ Alcohol dependence: A cluster of physiological, behavioral, and cognitive phenomena in which the use of a substance takes on a much higher priority for a given individual than other behaviors that once had greater value	■ Excise taxes*** ■ Restriction on sales** ■ Minimum legal age** ■ Measures to prevent drunk driving** ■ Advertising bans* ■ Restrictions on density* ■ Opening hours, closing hours, and days of sale** ■ Family interventions*	■ Naltrexone, acamprosate*	■ Family support* Motivational enhancement, brief advice, cognitive-behavioral therapy** ■ Screening and brief interventions*** ■ Self-help groups*
Illicit drug use disorders (7.8% total MNS DALYs)	■ A pattern of regular use of illicit drugs characterized by significantly impaired control over use and physiological adaptation to regular consumption as indicated by tolerance and withdrawal	■ Psychosocial interventions with primary school children (e.g., good behavior game or strengthening families program)*	■ Opioid substitution therapy (e.g., methadone, buprenorphine)***	■ Self-help groups, psychological interventions (e.g., cognitive-behavioral therapy)*
Suicide and Self-Harm				
Suicide and self-harm (1.47% of GBD; 22.5 million YLLs or 62.1% of suicide YLLs were attributed to mental and substance use disorders in 2010)	■ Suicide: The act of deliberately killing oneself ■ Suicide attempt: Any nonfatal suicidal behavior; includes intentional self-inflicted poisoning, injury, or self-harm that may or may not have a fatal intent or outcome	■ Policies and legislation to reduce access to the means of suicide (e.g., pesticides)** ■ Decriminalization of suicide* ■ Responsible media reporting of suicide*	■ Effective drug interventions for underlying MNS disorders** ■ Emergency management of poisoning**	■ Social support; psychological therapies for underlying MNS disorders*

Notes: Strength of evidence: *** evidence of cost-effectiveness (CE); ** Strong evidence of effectiveness (E) but not of CE; * modest evidence of E and either not CE or no evidence of CE.

TABLE 10-7 Intervention Priorities for MNS Disorders by Delivery Platform

Platform for Intervention Delivery						
			Healthcare Platforms			
Problem Area	**Population Platform**	**Community Platform**	**Self-Care and Management**	**Primary Health Care**	**First-Level Hospital**	**Specialized Care**
All MNS disorders	Awareness campaigns to increase mental health literacy and address stigma and discrimination Legislation on protection of human rights of persons affected by MNS disorders	Training of gatekeepers (e.g., community workers, police, teachers) in early identification of priority disorders, provision of low-intensity psychosocial support, and referral pathways Self-help and support groups (e.g., for alcohol use disorders, for epilepsy, parent support groups for children with developmental disorders, and for survivors of suicide)				
Adult mental disorders	Child protection laws		Physical activity Relaxation training Education about early symptoms and their management web- and smartphone-based psychological therapy for depression and anxiety disorders	Screening and proactive case finding for psychosis, depression, and anxiety disorders Diagnosis and management of depression and anxiety disorders* Continuing care of schizophrenia and bipolar disorder	***Diagnosis and management of acute psychoses*** Management of severe maternal depression* Management of depression and anxiety disorders in mothers, people with HIV, and people with other noncommunicable diseases*	*Electroconvulsive therapy for severe or refractory depression* Management of refractory psychosis with clozapine

(continues)

TABLE 10-7 Intervention Priorities for MNS Disorders by Delivery Platform *(continued)*

Platform for Intervention Delivery						
			Healthcare Platforms			
Problem Area	**Population Platform**	**Community Platform**	**Self-Care and Management**	**Primary Health Care**	**First-Level Hospital**	**Specialized Care**
Child mental and developmental disorders	Child protection laws	**Parenting programs in infancy to promote early child development** **Life skills training in schools to build social and emotional competencies** Parenting programs in early and middle childhood (2–14 years) Improve the quality of antenatal and perinatal care to reduce risk factors associated with intellectual disability	Web- and smartphone-based psychological therapy for depression and anxiety disorders	Screening for developmental disorders in children and maternal mental health interventions Parent skills training for developmental disorders Psychological treatment for mood, anxiety, ADHD, and disruptive behavior disorders*	**Diagnosis of childhood mental disorders such as autism and ADHD** Stimulant medication for severe cases of ADHD Newborn screening for modifiable risk factors for intellectual disability	
Neurological disorders	Policy interventions to address the risk factors for cardiovascular diseases (e.g., tobacco control)		**Self-managed treatment of migraine** Self-identification/management of seizure triggers Self-management of risk factors for vascular disease (e.g., healthy diet, physical activity, tobacco use)	Diagnosis and management of epilepsy and headaches Community-based screening for detection of dementia Interventions to support caregivers of patients with dementia *Management of prolonged seizures or status epilepticus*	Diagnosis of dementia and secondary causes of headaches	Surgery for refractory epilepsy

Alcohol and illicit drug use disorders	**Regulate the availability and demand for alcohol (e.g., increases in excise taxes on alcohol products, advertising bans)** Penalize risky behaviors associated with alcohol (e.g., enforcement of BAC limits)	Awareness campaigns to reduce maternal alcohol use during pregnancy	Self-monitoring of substance use	**Screening and brief interventions for alcohol use disorders** Opioid substitution therapy (e.g., methadone and buprenorphine) for opioid dependence	*Management of severe dependence and withdrawal*	Psychological treatments (e.g., CBT) for refractory cases*
Suicide and self-harm	**Control the sale and distribution of means of suicide (e.g., pesticides)** Decriminalize suicide	Safer storage of pesticides in the community and farming households	Web- and smartphone-based treatment for depression and self-harm	Primary health-care packages for underlying MNS disorders (as described above)* Planned follow-up and monitoring of suicide attempters* *Emergency management of poisoning*	Treatment of comorbid mood and substance use disorder*	Specialist health-care packages for underlying MNS disorders (as described above)

Abbreviations: ADHD, attention-deficit/hyperactivity disorder; BAC, blood alcohol concentration; CBT, cognitive-behavioral therapy; MNS = mental, neurological, and substance use.
Italic type denotes urgent care; underlined type denotes continuing care; normal font denotes routine care.
Recommendations in **bold** = best practice; recommendations in normal font = good practice.
* There is no fixed time period for the management of these complex conditions. For example, in the management of depression, some individuals need relatively short periods of engagement (e.g., 6–12 months for a single episode) at the one end, while others may need maintenance care for several years (e.g., when there is a relapsing course).

EXHIBIT 10-3 Community Care for Severe Mental Disorders in Low-Resource Settings

A community mental health program for severe mental disorders in a rural setting in India was initiated in partnership with Ashagram ("village of hope"), a nongovernmental organization (NGO) working toward the rehabilitation for people affected by leprosy. The NGO was located in Barwani, one of the poorest districts in India. Mental health care was routinely provided through an outpatient clinic that required patients to travel to the hospital to be assessed and to receive treatment.

A community-based rehabilitation (CBR) model was devised for patients with chronic schizophrenia, based on a three-tiered service delivery system. CBR relies on the active participation of the disabled and their families in rehabilitation and takes specific notice of prevailing social, economic, and cultural issues. The highest tier was outpatient (OP) care. All patients were started on antipsychotic medication. The second tier consisted of mental health workers (MHWs) drawn from the local community. After a 60-day training program, the MHWs worked with patients, families, and the local community in providing services. Each MHW served five or six contiguous villages and carried a total caseload of 25 to 30 patients, including some of the study subjects. The third tier consisted of family members and other key persons in the community who formed the local village health groups (*samitis*). These groups served as a forum for the members to plan relevant rehabilitation measures and reduce social exclusion.

The evaluation of the CBR program showed that, among patients who actively participated, this model was more effective than standard outpatient treatment, as determined by a range of clinical and functional outcomes (Chatterjee, Patel, Chatterjee, & Weiss, 2003). A four-year follow-up of the cohort of persons in the CBR care arm showed that adherence with medication and participation in self-help groups predicted a favorable outcome (Chatterjee, Pillai, Jain, Cohen, & Patel, 2009). A subsequent randomized controlled trial of this intervention in three sites confirmed its beneficial effects in reducing levels of disability in people with chronic schizophrenia in India (Chatterjee et al., 2014). Because a lack of professional resources is a reality in LMICs, the CBR method takes advantage of active local community participation and low levels of technical expertise to deliver services.

that are assumed to underlie adolescent drinking. Structured universal interventions for children in settings of armed conflict have positive effects for behavior, self-esteem, and coping.

The most effective promotion initiatives take a multicomponent approach by integrating their activities with programs on topics such as microfinance or reproductive and sexual health education. Standalone mental health promotions, such as only doing cognitive-behavioral therapy, have produced fewer benefits. Other good practices include parenting programs with infants for early child development.

Community care using locally available resources is a key strategy for enabling people with severe mental disorders and disabilities to remain in the community. One such example is the Ashagram model in rural India (**EXHIBIT 10-3**).

Preventive efforts directed toward reducing the risk factors for epilepsy and developmental disabilities have focused on improving prenatal care and promoting safe delivery. Other preventive strategies include better fever control in children; strategies aimed at reduction of the causes of brain injury, such as children's use of safety seats and helmets; control of infectious and parasitic diseases that infect the brain; genetic counseling; screening programs for conditions that are known to be associated with mental handicaps, such as hypothyroidism; micronutrient supplementation, such as with iodine; and reductions in environmental levels of toxins such as lead.

Healthcare-Level Recommendations

Best practices at the healthcare level include self-management psychological interventions, such as web-based psychological therapy for depression and anxiety. Given the tremendous treatment gap, the lack of available mental health specialists in most of the world, and the desire to move treatment from institutions to communities, numerous efforts are being made to deliver mental health through primary care (Gwaikolo et al., 2017; Lund et al., 2012; Petersen et al., 2016; WHO, 2010). A key innovation to attain this goal is to address the supply-side barrier of inadequate mental health specialists through task sharing of front-line interventions with nonspecialist health workers. A growing evidence base testifies to the effectiveness of nonspecialist providers in facilitating management of a range of mental disorders, typically in primary care and community settings (van Ginneken et al., 2013).

Perhaps the strongest evidence base supports task sharing of psychological treatments among primary care workers, community health workers, and non-healthcare providers in the community (**EXHIBIT 10-4**). In Uganda, interpersonal therapy (IPT) was adapted so that it could be delivered by highly supervised lay

EXHIBIT 10-4 Do Talking Treatments Work in Low-Income Countries?

> Things were already going pretty badly for Florence Manyande. Then one day last spring, while walking down the street, she was hit by a car.
>
> "This woman saw, and she pulled me out of the road," recalls Manyande, 50. "She tried to talk to me, but I couldn't talk then. I had a lot on my mind."
>
> Her run of bad luck had begun in 2010, when Manyande's husband skipped out on her and her three kids. "I had no way to pay school fees for my children," she says, and no way to pay rent. "Even my relatives were shunning me. They couldn't take me in because they said, 'We have our own problems.'"
>
> By the time Manyande had her accident, she was thinking about killing herself.
>
> Then her fortune took a turn. The woman who found her, injured, on the road happened to be a health worker. She took Manyande to the clinic to get bandaged up. "While I was there," Manyande says, "she introduced me to the 'Friendship Bench.'" (Singh, 2017)

There has been skepticism about the applicability of psychological treatments—such as cognitive-behavioral therapy (CBT), interpersonal therapy (IPT), and dialectical behavioral therapy (DBT)—for mental health problems in low-resource, non-Western cultural settings. Lack of clinical psychologists and different cultural belief systems have been seen as barriers to applying these psychological treatments. However, a rapidly growing evidence base suggests that lay persons can be trained and closely supervised to effectively deliver psychotherapies (Singla et al., 2017).

In developing the "Friendship Bench," Dixon Chibanda, a psychiatrist from Zimbabwe, adapted problem-solving therapy by integrating local cultural psychological concepts such as *kuvhura pfungwa* ("opening of the mind"), *kusimudzira* ("uplifting"), and *kusimbisa* ("strengthening"). Moreover, rather than focusing on psychiatric labels that could be stigmatizing, such as "depression" or "general anxiety disorder," the treatment employed culturally acceptable idioms of distress, such as *kufungisisa* ("thinking too much"). The intervention was delivered in a nonstigmatizing setting—a bench outside of health clinics, leading to the "Friendship Bench" name for the program.

Six months after receiving the Friendship Bench psychological intervention, only 13% of patients had depression, compared with 50% of persons in a basic treatment comparison group (Chibanda et al., 2016). Numerous other psychological treatments delivered by nonspecialists in LMICs—such as the Thinking Health Program, Healthy Activity Program, and Problem Management Plus (PM+)—have shown similar benefits. These interventions demonstrate both the effectiveness and the feasibility of psychotherapy for mood and anxiety disorders in settings that vary by culture and resource availability around the world.

community members for treatment of depression among war-affected adults and youth (Bolton et al., 2007; Bolton et al., 2003). Cognitive-behavioral therapy (CBT) was adapted for treatment of perinatal depression by female community health volunteers through the Thinking Healthy Program in Pakistan, which is now being replicated in diverse settings in LMICs (Rahman, Malik, Sikander, Roberts, & Creed, 2008). Behavioral activation has been adapted as a lay health worker–directed Healthy Activity Program for treatment of depression in India (Patel et al., 2017); Counseling for Alcohol Problems is a similar lay psychological treatment adapted from motivation interviewing for treatment of harmful alcohol use (Nadkarni et al., 2016). A Common Elements Treatment Approach (CETA) has also been developed so that closely supervised nonspecialists can provide trans-diagnostic care that incorporates elements from multiple psychological treatments (Murray et al., 2014). All of these nonspecialist interventions have demonstrated superiority to the usual treatments. In fact, when all nonspecialist psychological interventions for common mental disorders were reviewed (24 studies as of 2016), there was a pooled effect size of 0.49 (Singla et al., 2017), which approaches the effect sizes observed in psychological treatments delivered by specialists in high-income countries (Huhn et al., 2014).

In task sharing, nonspecialists take on detection, diagnosis, and management of health conditions in conjunction with training, supervision, and referral support by specialists (WHO, 2008). Thus, a key component of task sharing is the recognition that integration is not a simple matter of training nonspecialist health workers. A collaborative care delivery model, in which primary and community care practitioners work together with specialists, is an essential element for integrating mental health in primary care (Patel et al., 2013). Furthermore, primary care mental health services, although essential, cannot meet all of the mental health needs of any given population. This is especially true for people who are suffering from psychotic disorders and who need access to community mental health and rehabilitation programs, as well as emergency inpatient facilities (Patel, Farooq, & Thara, 2007).

Studies have explored the role of lay community health workers in primary and secondary prevention (Mutamba, van Ginneken, Smith Paintain, Wandiembe, & Schellenberg, 2013). For example, prevention studies in LMICs have shown that such interventions can reduce the burden of depression and PTSD in adults. Antenatal programs to prevent postpartum depression have led to improvements in other health outcomes related to children and mothers, such as mother–infant engagement (Rahman et al., 2013). There is also a growing evidence base on "early interventions," such as supportive psychotherapy and low-dose antipsychotic agents, that may delay or prevent progression to a first mental disorder episode (McGorry, 2015). Reducing the duration of untreated psychosis at the onset of schizophrenia can dramatically impact the life course of the illness (Fusar-Poli et al., 2009). This consideration is especially important in LMICs, where the duration of untreated psychosis can average two years, and it is not uncommon to go more than five years without initiating treatment (Thirthalli et al., 2011).

According to DCP-3 guidelines, at the healthcare level, hospital care is recommended as a best practice for delivery of MNS specialty services for severe, refractory, and emergency presentations of MNS disorders. In addition, mental health care should be integrated into other specialty services ranging from obstetric and gynecologic care to infectious diseases services to oncologic treatment.

Humanitarian Emergencies

Mental health services are especially important to address the acute and chronic needs of populations affected by complex humanitarian emergencies, such as war, environmental disasters, earthquakes, and other causes of forced displacement. Refugees, internally displaced persons, and other survivors of collective trauma are at an increased risk of MNS disorders. Unfortunately, populations affected by humanitarian emergencies live in settings where the health, and especially the mental health, services sector may have been limited or nonexistent even before the disaster. There is a heightened need to identify and allocate resources for providing mental health care and psychosocial support in these settings, both for those with disorders induced by the emergency and for those with preexisting disorders.

Guidance on mental health and psychosocial response is available through the Inter-Agency Standing Committee (IASC, 2007). WHO (2015) has also released a Humanitarian Intervention Guide version of mhGAP, which provides treatment recommendations for acute emergency response and includes trauma and grief-related conditions not covered in the basic mhGAP. In a number of countries, such emergencies have actually provided opportunities for systemic change or services reform in public mental health (WHO, 2013c).

Delivering Effective Treatments and Scale-up

Despite the evidence presented earlier, the fact remains that only a small proportion of people suffering from mental disorders receive effective interventions. This failure to provide effective treatments to all people who need them, which has been termed the "treatment gap," is evident in all countries, but is more marked in LMICs. Globally, only one in five people with depression in high-income countries receives minimally adequate treatment; however, the situation is far worse in lower-middle income countries, where only one in 27 people with depression receives minimally adequate care (Thornicroft et al., 2017).

Factors contributing to the treatment gap may be categorized as supply-side barriers and demand-side barriers. The supply-side barriers include lack of trained mental health workers, availability of culturally competent providers to address diverse ethnic and social groups, lack of transportation, and lack of affordable services. As mentioned earlier, there is a severe shortage of specialists in LMICs, where the bulk of the resources available for mental health services are devoted to large psychiatric hospitals and services are concentrated in a few urban settings (WHO, 2014)—this represents a major supply-side barrier in these countries. The demand-side barriers include lack of recognition among the public about mental disorders, lack of awareness that mental disorders can be treated, and stigma that prevents individuals from seeking care.

A challenge to implementing recommended prevention and treatment programs, as described in the DCP-3, is lack of political will—evidenced by less than 1% of the total health budget being allocated to MNS care in most national health systems. Similarly, less than 1% of development assistance goes toward MNS services. Given that 30% of the countries in Africa and 26% of countries in the Americas rank NGOs as the second major funding source for mental health care (WHO, 2015), it is clear that these development funds are stretched incredibly thin. In one out of five countries around the world, households

are the main source of funding for mental health care, mainly through direct out-of-pocket expenses and private health insurance coverage. Governments in Africa and the Americas are more likely to put the financial burden for such care on households. Additionally, health insurance in many countries specifically excludes or greatly restricts reimbursements for mental health care.

Even when government, employer, or private insurance policies do fund mental health care, there may not be *parity* with physical health services—that is, only a limited number of outpatient visits or inpatient treatment days may be covered by insurance. Mental health leaders from the United Kingdom, Canada, Australia, and New Zealand have led efforts to achieve parity in mental health care around the world (Royal College of Psychiatrists, 2012, 2013). In contrast, the United States has traditionally lagged behind other high-income countries, with major disparities being noted in coverage of mental health services compared to physical health: For outpatient services, 77% to 90% of plans place limits on mental health care that are not imposed on physical health care and 66% to 74% limit inpatient MNS services (Thalmayer, Friedman, Azocar, Harwood, & Ettner, 2017). The U.S. Mental Health Parity and Addiction Equity Act of 2008 required parity in insurers' coverage of mental health and physical health. In LMICs, recent national mental health policies in South Africa, India, and Liberia exemplify initiatives to increase government engagement in and support of mental health services.

Lack of financial investment in mental health services is related to high levels of stigma and lack of mental health literacy that lead to low demand from constituents, as well as lack of technical leadership to design and implement MNS treatment programs. Strategies for health-system strengthening include efforts in the following areas:

- To enact WHO's Comprehensive Mental Health Action Plan (WHO, 2013b); to adopt a mainstream rights-based perspective
- To update health policies, plans, and laws to be consistent with international human rights and standards, such as the United Nations' Convention on the Rights of Persons with Disability (UNCPRD)
- To address stigma and enhance mental health literacy to increase demand for care
- To increase advocacy by mental health service users
- To improve MNS services financing through diversion of taxes (alcohol, tobacco, marijuana), promotion of low-cost generic drugs, and de-implementation of harmful or ineffective treatments (e.g., benzodiazepines and vitamins in primary care)
- To include MNS disorders in health management information systems as national indicators

There have been a number of recent initiatives in implementation science to address both the mental health treatment gap and the efficacy–effectiveness gap. The Program for Improving Mental Health Care (PRIME) is integrating mental health services into primary and community care, based on WHO's mhGAP program, in South Africa, Uganda, Ethiopia, India, and Nepal. PRIME's objectives include developing packages of care for integration of mental health services into primary care, based on mhGAP; training primary care workers to deliver mental services; and advancing policy to create sustainable mental health services that can be scaled up from proof-of-concept districts to country-wide services (Lund et al., 2012). Emerging Mental Health Systems in Low- and Middle-Income Settings (EMERALD) is a six-country program (PRIME countries plus Nigeria) that addresses the gap in health systems' ability to scale up packages because of inadequate policy, human resources, funding, and infrastructure. Building on the work of PRIME, EMERALD's objectives include evaluating health-system inputs (e.g., resourcing with WHO's OneHealth tool, fair and sustainable financing), evaluating health-system processes (e.g., policy, legislation, governance, consumer participation), and evaluating health-system outputs (e.g., development, implementation, and monitoring of mental health indicators and performance) (Semrau et al., 2015). Another resource to support best practices in global mental health and advance from research to implementation is the Mental Health Innovation Network (MHIN; www.mhinnovation.net). These initiatives hold promise for addressing the needs of women, men, and children with mental disorders around the world. **EXHIBIT 10-5** outlines examples of national scaling-up in three countries across the spectrum of economic development.

Involving a range of sectors is a key aspect of responding to this call to action in all settings, and arguably even more so in LMICs, where formal mental healthcare systems are typically inadequately developed. Intersectoral, community-based action has been used to prevent alcohol and drug abuse and for rehabilitation models for schizophrenia. Another sector that has made important contributions to mental health care and reforms in high-income countries

EXHIBIT 10-5 Scaling Up Mental Health Initiatives

The "686 Project": China (Ma, 2012)

The "Central Government Support for the Local Management and Treatment of Severe Mental Disorders Project" was initiated in China in 2004 with a first financial allocation of 6.86 million renminbi ($829,000 in 2004 dollars). The program was subsequently referred to as the "686 Project." Modeled on WHO's recommended method for integrating hospital-based and community-based mental health services, this program provides care for a range of severe mental disorders through the delivery of a community-based packages by multidisciplinary teams. The interventions are functionally oriented, and are provided as free outpatient treatment through insurance coverage (New Rural Cooperative Medical Care system) and as subsidized inpatient treatment for poor patients. The program covered 30% of China's population by the end of 2011.

Program evaluation showed improved outcomes for the more than 280,000 registered patients. The proportion of patients with severe mental disorders who did not suffer a relapse for five years or longer increased from a baseline of 67% to 90%, and there were large reductions in the rates of "creating disturbances" and "causing serious accidents." The program investment by the government amounted to 280 million renminbi in 2011, and its key innovations were the increased availability of human resources, including both the involvement of non-mental health professionals and intensive capacity building; the latter has added one-third of all psychiatrists now working in China.

The National Depression Detection and Treatment Program: Chile (Alvarado & Rojas, 2011)

Programa Nacional de Diagnóstico y Tratamiento de la Depresión (The National Depression Detection and Treatment Program) is Chile's national mental health program that integrates detection and treatment of depression in primary care. The program is based on the scaling-up of an evidence-based collaborative stepped care intervention in which most patients diagnosed with depression are provided with medications and psychotherapy at primary care clinics, while only severe cases are referred to specialists. Launched in 2001, the program operates through a network of 500 primary care centers, and presently covers 50% of Chile's population. The program has added a large number of psychologists in the primary care system, with a 344% increase in the number of these providers occurring between 2003 and 2008. Enrollment of the patients in the program is growing steadily, with approximately 100,000 to 125,000 patients starting treatment each year from 2004 to 2006 and close to 170,000 patients starting treatment in 2007. Universal implementation of the program has led to a greater utilization by women and less-educated individuals, contributing to reduced health inequalities. The program's success can be attributed to the use of an evidence-based design that was made available to policy makers, teamwork, and proactive leadership, strategic alliances across sectors, sustained investment and ring-fencing of new and essential financial resources, program institutionalization, and sustained development of human resources that can implement the program.

Building Back Better: Burundi (WHO, 2013a)

Civil war in the last decade of the 20th century and the first decade of the 21st century resulted in widespread massacres and forceful migrations and internal displacement of approximately 1 million individuals in Burundi. To address this humanitarian crisis, Healthnet TPO (Transcultural Psychosocial Organization) started providing mental health services in Burundi during 2000, when the then Ministry of Public Health had no mental health policy, plan, or mental health unit, and when virtually all the psychiatric services in Burundi were provided by one psychiatric hospital. Healthnet TPO conducted a needs assessment first and then built a network of psychosocial and mental health services in communities in the national capital, Bujumbura, and in 7 of the country's 17 provinces. A new health worker cadre, consisting of psychosocial workers, played a pivotal role in delivery of these services.

Considerable progress has been made in the last decade, with the government now supplying essential psychiatric medications through its national drug distribution center, and outpatient mental health clinics being established in several provincial hospitals. From 2000 to 2008, more than 27,000 people were helped by the newly established mental health and psychosocial services. In the three years from 2006 to 2008, the mental health clinics in the provincial hospitals registered almost 10,000 people, who received more than 60,000 consultations. The majority (65%) were people with epilepsy.

In 2011, funding from the Dutch government enabled HealthNet TPO and the Burundian government to initiate a five-year project aimed at strengthening the country's health systems. One of the project's components is the integration of mental health care into primary care using WHO's mhGAP guidelines. The government has now established a National Commission for Mental Health, and appropriate steps are being taken to support provision of mental health care in general hospitals and follow-up within the community.

and has only recently achieved recognition in LMICs are consumer- and family-led movements. The World Fellowship for Schizophrenia and Allied Disorders and Alzheimer's Disease International are examples of NGOs that have their origins in high-income countries, where strong consumer movements led by families of persons with schizophrenia and Alzheimer's disease led to their establishment. In the past decade, both NGOs have established a growing presence in LMICs. Similarly, Befriender International, a voluntary group that provides support to persons who are suicidal, has spread to a number of LMICs. Local NGOs led by families of persons affected by mental disorders are also multiplying rapidly in LMICs. There are fewer examples of community movements that are led by persons who are themselves suffering from mental disorders. Perhaps the best example is Alcoholics Anonymous, which is widely represented internationally, and which is one of the most well-described examples of an effective community-based intervention for a mental disorder.

Advocacy to policy makers, the media, and other sectors in the health system is a core activity of these civil society organizations. Prominent examples of the success of these advocacy efforts include the inclusion of mental disabilities in the disability legislation passed by some countries. Many groups also provide services, usually in the form of support groups or networking for affected families, but the larger groups also support research activities and medical care. The Movement for Global Mental Health (mentioned earlier), which was launched in October 2008, took its inspiration from the global HIV/AIDS movements that have transformed HIV/AIDS care through a massive scaling-up of resources and services to provide a comprehensive continuum of care for people living with HIV (PLHIV) across the globe; the Movement is perhaps the largest global coalition of such civil society groups centered on mental health issues in existence.

One of the challenges going forward as part of scale-up of mental health services is empowering front-line health practitioners, along with considering how managers and administrators in governmental and nongovernmental health systems can make the leap into being active players in addressing gaps in mental health services. Clinical guidance, ranging from *Where There Is No Psychiatrist* (Patel & Hanlon, 2018) to the mhGAP second edition (WHO, 2016), seeks to empower health workers at any level of expertise to engage in mental health services provision. However, approaches are needed to synthesize lessons learned at the health organization and administration level. The DCP-3 focuses on MNS disorders and can help health system administrators determine the type and content of interventions across levels of care. For example, DCP-3 can guide administrators in determining which MNS conditions should be screened and at which level, as well as which types of pharmacologic and psychological interventions should be available at different health systems levels. The QualityRights Toolkit developed by WHO (2012) is another essential tool that can be used to assess health facilities and other social services institutions to assure human rights are protected and promoted in the delivery of mental health care. Case studies have been successfully used to identify common features for success and common challenges in implementing community mental health services in low-resource settings (Cohen et al., 2011).

Quality improvement tools that assess the fidelity of replicated interventions, competency of health workers, and outcomes and satisfaction among patients can also be used by health system administrators to guide implementation of mental health care services (Glisson & Williams, 2015). In Ethiopia, the Mental Health Services Satisfaction Scale (MHSSS) was developed as a self-report instrument for services users to evaluate interpersonal factors such as interactions with health workers, efficacy of treatment, communication within the health system, technical competency of providers, and adequacy of facilities (Mayston et al., 2017). This tool can be combined with community forums, which are government-established meetings for community members to provide feedback to local health facility workers. In Nepal, the Enhancing Assessment of Common Therapeutic Factors (ENACT) tool was developed as a way to rate nonspecialist health workers on basic mental health skills (Kohrt, Jordans, et al., 2015); it can be used by health administrators to determine when health workers have achieved sufficient competency to provide care and to guide supervision geared toward ongoing quality improvement. In Liberia, a combination of qualitative interviews and health facility infrastructure surveys have been used to determine health system preparedness for integration of mental health services into primary care (Gwaikolo, Kohrt, & Cooper, 2017). Based on PRIME's reception in sub-Saharan Africa and South Asia, a set of eight principles (**EXHIBIT 10-6**) have been identified for successful implementation of mental health services in community and primary-care settings (Davies & Lund, 2017).

EXHIBIT 10-6 Lessons for Integrating Mental Health into Primary Care in Low-Resource Settings

1. *Engage actively and collaboratively with local stakeholders*. Participatory methods involving health workers, health system administrators, ministry officials, community representatives, and mental health service users and family members are crucial to identify facilitators and barriers to successful implementation. Theory of Change (ToC) workshops are a methodology that have been developed and successfully implemented in global mental health to determine assumptions, interventions, and indicators for programs (Breuer et al., 2014).
2. *Use primary care systems to access vulnerable populations.* Health programs that have been successful in engaging vulnerable populations, such as pregnant women, orphan children, and persons with HIV, malaria, or tuberculosis, can be ideal platforms for the introduction of mental health services. Given the syndemic nature of comorbid physical and mental health problems and high-risk social environments, primary care be the entry point for comprehensive care models (Mendenhall, Kohrt, Norris, Ndetei, & Prabhakaran, 2017).
3. *Use cultural concepts of distress and narrative-based vignettes to identify persons with potential mental health problems*. Because of both the stigma and the lack of familiarity associated with biomedical psychiatric terminology, availability of treatment may not translate into engagement with treatment. The framing of psychoeducation and psychological and pharmacologic treatments around culturally salient ways of describing mental health problems can overcome the barriers associated with psychiatric terminology. An even more effective way to facilitate treatment engagement may be the use of culturally salient vignettes describing individuals with mental health problems. In Nepal, the Community Informant Detection Tool (CIDT) was developed to facilitate referrals by community stakeholders. Studies showed that more than two-thirds of individuals referred using the CIDT sought treatment, and the accuracy of the CIDT was better than screening tools such as the PHQ-9 (Jordans, Kohrt, Luitel, Komproe, & Lund, 2015; Jordans, Kohrt, Luitel, Lund, & Komproe, 2017).
4. *Use manual-based approaches to deliver care through nonspecialist health workers.* Extensive work in the field of global mental health has focused on adapting psychological and psychiatric treatment manuals for nonspecialist, cross-cultural audiences (Balaji et al., 2012). Some manuals even cater to providers with only basic literacy (Verdeli et al., 2003). Thoroughly adapted and piloted manuals can be used to optimize fidelity to evidence-based treatments. Increasingly, e-versions of manuals and intervention materials are being produced to facilitate delivery by nonspecialists (Zafar et al., 2016).
5. *Set up systems of ongoing supervision and support*. Training is inadequate in the absence of ongoing supervision when initiating mental health services in a new context. Supervision can be provided in person, via phone, or online by specialists and peers. Competency assessment tools and fidelity checklists can be used to identify key areas for remediation in supervision (Kohrt et al., 2015).
6. *Adequately compensate nonspecialist workers for their services.* For sustainability of mental health services, compensation of providers needs to be addressed. The PRIME program has evaluated the cost of scaling up services (Chisholm et al., 2016). When creating projections of costs, it is important to consider whether governments will fund providers rather than relying upon volunteer mental health workers.
7. *Respond to crises by "building back better."* Crises ranging from Ebola outbreaks in West Africa to earthquakes in Nepal have drawn attention to the inadequacy of LMICs' existing health systems to meet their populations' mental health needs. When international attention is drawn to the humanitarian emergencies, it can be an opportunity to make sustainable changes. For example, the earthquake in Nepal contributed to a reexamination of that government's essential drugs list, which was still relying on half-century-old classes of psychiatry medications with high side-effect burdens. Through a combination of international NGO funding, local NGO advocacy, and expertise from local psychiatrists, the government revised the psychiatric medications list.
8. *Make use of policy windows*. In South Africa, pilot implementation of mental health services as part of primary care was conducted in a district that was also a national pilot site for the planned national health insurance reforms. In another district in South Africa, the department of health was launching an initiative titled the "First 1000 Days" to promote maternal and child health and education. Mental health researchers then selected this district to promote and initiate maternal mental health interventions in routine maternal health services.

Modified from Davies, T., & Lund, C. (2017). Integrating mental health care into primary care systems in low- and middle-income countries: Lessons from PRIME and AFFIRM. *Global Mental Health, 4*, e7.

Global Mental Health: Looking Ahead

Ten years on from the first *The Lancet* series that propelled mental health into the global health spotlight, it is time to consider where the field should head in the next decade. It is evident that, while the existing agenda to improve access to care is still very far from being attained (e.g., treatment gaps remain very large in all countries), this alone is unlikely to lead to a measurable impact not only in reducing the treatment gap but ultimately in reducing the global burden of mental disorders. In this section, we outline research priorities for global mental health to tackle these needs, and then we conclude with an agenda to address key challenges in the field.

Research Priorities in Global Mental Health

Research priorities in global mental health are continuously evolving and reflect national and global policies, funders and their priorities, public attention to health issues, advocacy from human rights and service users' groups, and other trends in national and global health. There have been recent inflection points in the development of global mental health research priorities. Notably, the group of scholars who produced the 2007 *The Lancet* Global Mental Health series conducted a priority-setting exercise to identify the top 10 mental health research questions (Tomlinson et al., 2009). The resulting priority questions were related to health policy and systems research, where and how to deliver existing cost-effective interventions in a low-resource context, and epidemiologic research on the broad categories of child and adolescent mental disorders or those pertaining to alcohol and drug abuse.

A priority-setting activity for the Global Forum identified priorities among researchers and stakeholders in LMICs (Sharan et al., 2009). Priorities included determining the burden of MNS disorders, identification of risk factors, and studying health systems. Prioritized disorders were depression and anxiety, substance use disorders, and psychoses. Prioritized populations were children and adolescents, women, and persons exposed to violence and trauma. Social sciences methods have also been recommended to complement epidemiologic and intervention studies (Kohrt, Mendenhall, & Brown, 2016).

As mentioned at the beginning of the chapter, the Grand Challenges in Global Mental Health (Collins et al., 2011) include a range of research priorities, with an emphasis on implementation science (**TABLE 10-8**). A priority-setting activity was conducted for mental health and psychosocial support research in humanitarian settings (Tol et al., 2011). This exercise yielded key questions to be addressed in global mental health, including identifying stressors faced by populations in humanitarian settings, determining methods to assess mental health in humanitarian settings, documenting how affected populations describe mental health, selecting indicators for monitoring and evaluation, adapting interventions to different sociocultural settings, and determining the effectiveness of family- and school-based preventive interventions.

Emerging research priorities in global mental health include enhancing our understanding of mental health problems through global representation of populations in basic science and neuroscience research (Stein et al., 2015). Because neuroscience models of mental health are dominated by research in WEIRD populations (Henrich, Heine, & Norenzayan, 2010a, 2010b), generalizability of these models to LMIC populations is limited and the progress in understanding the nature of mental health problems has been slow. This is especially problematic when these neuroscience mechanisms are selected as the prioritized targets for intervention before confirmation of similar processes in populations with different child developmental histories, different socioeconomic conditions, and different genetic admixtures (Kirmayer & Gold, 2012a). Nevertheless, potential synergies exist between global mental health and clinical neuroscience research priorities, such as development of reliable criteria for diagnosis of schizophrenia across populations and contexts, understanding responses to trauma and adversity including early-life adversity, understanding resilience across contexts, and development of interventions for nonspecialist providers and preventive measures for resilience promotion (Stein et al., 2015).

As in other areas of global health, there is increasing interest in the application of digital technology and mHealth to increase the delivery, accessibility, and effectiveness of mental health services. In global mental health, digital technologies have been predominantly used in the domains of supporting clinical care and educating health workers, facilitating diagnosis and detection of mental disorders, promoting treatment adherence, and supporting recovery, online self-help, and programs for substance misuse prevention and treatment (Naslund et al., 2017). To date, however, there has been limited research into the effectiveness of these approaches, and data are lacking for both clinical outcomes and cost–benefit analyses of digital technologies compared to traditional approaches. The

TABLE 10-8 Grand Challenges for MNS Disorders

	Top 25 Challenges	Illustrative Research Questions
Goal A: Identify root causes, risk, and protective factors	■ Identify modifiable social and biological risk factors across the life course ■ Understand the impact of poverty, violence, war, migration, and disaster ■ Identify biomarkers	■ What is the relationship between early fetal and child development and the onset of MNS disorders? ■ What are the phenotypes and endophenotypes of MNS disorders across cultural settings? ■ Which gene environment interactions are associated with increased risk for mental disorders? ■ Which factors promote resilience and prevent mental disorders in persons at extreme social disadvantage? ■ What role does social context play in the persistence of MNS disorders throughout life?
Goal B: Advance prevention and implementation of early interventions	■ Support community environments that promote physical and mental well-being throughout life ■ Reduce the duration of untreated illness by developing culturally sensitive early interventions across settings ■ Develop interventions to reduce the long-term negative impact of low childhood socioeconomic status on cognitive ability and mental health ■ Develop an evidence-based set of primary prevention interventions for a range of MNS disorders ■ Develop locally appropriate strategies to eliminate childhood abuse and enhance child protection	■ Which behavioral skills can enhance executive function, resilience, and cognitive flexibility throughout life? ■ Which neuroprotective agents and/or cognitive retraining paradigms can be used during the period of rapid brain development to reduce vulnerability to disorders in adolescence? ■ How effective are home- and school-based interventions for child abuse and neglect?
Goal C: Improve treatments and expand access to care	■ **Integrate screening and core packages of services into routine primary health care** ■ **Reduce the cost and improve the supply of effective medications** ■ Develop effective treatments for use by nonspecialists, including lay health workers with minimal training ■ Incorporate functional impairment and disability into assessment ■ **Provide effective and affordable community-based care and rehabilitation** ■ **Improve children's access to evidence-based care by trained health providers in LMICs** ■ Develop mobile and information technologies (such as telemedicine) to increase access to evidence-based care	■ How effective are brief screening tools for the detection of MNS disorders in routine care settings? ■ How effective are interventions for serious mental disorders delivered by lay health workers? ■ How will increased understanding of neural circuits lead to alternatives to current pharmacologic interventions? ■ How can mobile-phone technology be used to monitor seizure frequency? ■ How can video games and other electronic media be used for cognitive remediation across cultural settings? ■ Which psychosocial interventions produce the best outcomes for community-based care for MNS disorders across cultural settings?

Goal D: Raise awareness of the global burden of mental health disorders	■ Develop culturally informed methods to eliminate the stigma, discrimination, and social exclusion of patients and families across cultural settings ■ Establish cross-national evidence on the cultural, socioeconomic, and services factors underlying disparities in incidence, diagnosis, treatment, and outcomes ■ Develop valid and reliable definitions, models, and measurement tools for quantitative assessment at the individual and population levels for use across cultures and settings ■ Establish shared, standardized global data systems for collecting surveillance data on the prevalence, treatment patterns, and availability of human resources and services		■ What are the components of effective interventions to reduce stigma associated with MNS disorders? ■ Which interventions to reduce stigma and discrimination can be targeted to and implemented in health and social services settings in different health-system environments? ■ What is the impact of macroeconomic factors (such as unemployment rates, international trade, national income) on the prevalence of MNS disorders over time? ■ What is the impact of policy initiatives on the coverage of treatment for MNS disorders? ■ Which measurement factors contribute to differences in the prevalence of mental disorders across ethnic groups within and between countries?	
Goal E: Build human resources capacity	■ Increase capacity in LMICs by creating regional centers for mental health research, education, training, and practice that incorporate the views and needs of local people ■ Develop sustainable models to train and increase the number of culturally and ethnically diverse lay and specialist providers to deliver evidence-based services ■ **Strengthen the mental health component in the training of all healthcare personnel**		■ What is the most effective way to train primary healthcare workers to deliver evidence-based care with adequate fidelity to guidelines? ■ What is the comparative effectiveness of care for MNS disorders by different cadres of healthcare providers? ■ What are the views of low-income communities in high- and low-income countries on the priority research questions for MNS disorders?	
Goal F: Transform health-system and policy responses	■ Establish and implement minimum healthcare standards for MNS disorders around the world ■ Redesign health systems to integrate MNS disorders with other chronic-disease care, and create parity between mental and physical illnesses in regard to investments into research, training, treatment, and prevention ■ Incorporate a mental health component into international aid and development programs		■ What can we learn from different approaches (and associated costs) to integrated delivery of care across health systems? ■ What are the most effective health-system-wide strategies to reduce consumption of alcohol and illicit drugs? ■ What is the impact of legislation that ensures parity between mental and other illnesses on access to mental health services?	
Summary principles	■ Use a life-course approach to study	■ Use system-wide approaches to address suffering	■ Use evidence-based interventions	■ Understand environmental influences

Note: **Bold type** denotes the top five challenges ranked by disease-burden reduction, impact on equity, immediacy of impact, and feasibility.
Modified from Collins, P. Y., Patel, V., Joestl, S., March, D., Insel, T. R., Daar, A. S. (2011). Grand challenges in global mental health. *Nature, 475*, 27–30.

abundance of studies focusing on digital technology that are currently in the research pipeline will likely transform the landscape of how mental health is studied, prevented, and treated in the coming decade.

With a focus on technological advances in methods and interventions through neuroscience and mobile innovations, it is important not to overlook important research priorities that have a major impact on the human rights and quality of life of persons with mental illness. Though the drive toward community mental health is certainly appropriate, there is a risk that the role and quality of inpatient and residential facilities might be overlooked in the haste to move services into outpatient settings. The focus on community services may falsely imply that such services can handle all possible cases and institutional services are not needed, or that there are appropriate facilities to which emergencies can be referred. Unfortunately, neither presumption is true. Physical restraint, such as chaining people in homes or to trees, is all too often the only option left to families and religious leaders when higher-level care services are not available in LMICs (Asher et al., 2017; Minas & Diatri, 2008). Research on the appropriate role of institutions in the protection and care of persons with mental illness is a major gap in global mental health initiatives.

Finally, a re-envisioning of mental health and economic development research priorities is needed. If we are to align global mental health with current international sustainable development policy (as formulated in the SDGs), then we need to demonstrate not only that providing mental health interventions (SDG3) can reduce the global burden of disease, but also that the SDGs have the potential to improve population mental health by addressing the social and economic determinants of mental health. Achieving this outcome requires a research agenda focused on the social determinants of mental health, such as elucidating the mechanisms by which social and economic conditions determine mental health across the life course, and generating evidence of interventions that might address the upstream risk factors for mental illness. This means linking mental health interventions to broader development agendas, particularly those focusing on poverty, violence, migration, and gender. In this context, mental health studies would need to demonstrate outcomes that go beyond mental health, such as reductions in violence and improvements in economic circumstances. Linking mental health care with other development interventions to demonstrate interaction effects could be very beneficial for showing the added value of including mental health services in broader development initiatives. This requires a focus on certain developmental phases, such as early childhood development and adolescence, that are critical periods for neurological development and neuroplasticity. Given the plasticity and associated vulnerability of these periods, interventions need to be tailored to adolescents according to their level of deprivation or trauma exposure. Studies that demonstrate awareness of this factor—for example, through flexible, tailored, multicomponent interventions—could capitalize on this developmental stage to optimize those interventions' clinical benefits and cost-effectiveness and to extend the longevity of positive benefits.

A Global Mental Health Agenda for the Next Decade

A number of barriers have prevented the transformation of the large body of science into actions, both at the global and local levels, and particularly in LMICs (Saraceno et al., 2007); these pose systemic and systematic threats that need to be acknowledged and addressed explicitly if we are to see radical change in the future.

First, compared to the experiences of other global health movements (e.g., HIV/AIDS and maternal and child health), advocacy for mental health has been hampered by the reliance on process indicators (e.g., increasing coverage of services) rather than outcome indicators (e.g., improved mental health). This orientation may cause reluctance among health authorities to support and prioritize mental health (WHO, 2003a; Howes, 2005).

Second, fragmentation of the advocacy by diverse constituencies and science from diverse disciplinary traditions poses a challenge. For example, approaches to mental health issues range from the happiness agenda promoted by economists, to the specialist care for mental disorders promoted by clinical professionals, to the quest to fight discrimination promoted by civil society activists, to the ventures related to mapping the human brain promoted by neuroscientists. This leads to contradictory messages being sent to governments by the diverse stakeholders concerned with mental health and may leave critical questions unanswered. For example, should we focus more on addressing social determinants or biomedical treatment? Do we need (and want) more or less hospital beds? Should we promote new-generation, more expensive drugs, or cheaper, older-generation agents?

Third, there has been an alienation of the mental health professional communities from the vision of global mental health. While reforms of mental health services from institutions led by psychiatrists and team-based approaches oriented to primary and community

care have been unanimously supported by many constituencies, they have been opposed, at times, by powerful national groups of psychiatrists and other mental health professionals. This opposition also risks derailing the process of reforming psychiatric institutions and, consequently, may prevent improvements in quality of care and greater respect for and enhancement of the human rights of people with mental disability. This resistance is problematic, because innovations in care delivery at all levels of the health system cannot happen without the full participation and support of psychiatrists and nurses (Berlin, 1969; Fisch, 1965; Johansson, Astrom, Kauffeldt, Helldin, & Carlstrom, 2014).

Fourth, there is the risk of global mental health becoming yet another silo, unlinked to other momentous initiatives in global health, such as Every Woman Every Child or Universal Health Care. This risk is exemplified by the lack of adequate engagement with mental health in the curricula of general health professionals and the continuing emphasis in psychiatric training on biomedical aspects of mental health, while often ignoring the important contributions coming from the social sciences and global health (Hilty et al., 2006; Lobo, de-la-Camara, Campos, & Saz, 2015; Wynaden, Orb, McGowan, & Downie, 2000).

The future requires a reimagining of the existing agenda of global mental health in a number of significant ways. First, the scope should be global, addressing concerns that are relevant in all countries. Indeed, as far as mental health is concerned, it would be fair to say that all countries are "developing" to one degree or another. Second, we must acknowledge that the biomedical approach to binary categories of diagnoses of mental disorders, while of utility to health workers, does not accurately reflect the multidimensional nature of mental health, and recommend a staged model in its stead (Patel, 2017). The theoretical framework underpinning this vision is that of the continuum of mental health spans from well-being to disability and health to ill health. Third, to emphasize a convergent model of mental health, we must recognize the complex interplay of environmental, biological, and developmental factors across the life course, from conception to the grave, across generations and genders. Finally, the global community must recognize mental health as a basic human right for all peoples, albeit with a specific focus on those who face the gravest danger of their rights being denied, such as populations living in institutions or facing severe social adversity.

Today, we have a historic opportunity to reimagine the global mental health agenda in light of the broader perception of mental health as an integral part of the SDGs, as envisioned in WHO's Comprehensive Mental Health Action Plan. There was a call for greater investment in mental health in the landmark summit hosted jointly by the World Bank and WHO in Washington, D.C., in April 2016. This call has been accompanied by explicit acknowledgment of mental health as a global development issue, the potential for a grand convergence of science across disciplines, and growing consensus between stakeholders. Improvements in mental health hold the potential to contribute to enhanced sustainable development and the improvement of people's lives. Indeed, the aging of populations across the world, and trauma and displacement consequent to conflict and climate change, are among the great social challenges of our time, and these translate into greater demand in society for higher-quality services as well as for greater social support for disability and elder care—areas where mental health care looms large. Over and above these concerns, the clearly evident effects of climate change, poverty, and inequality on mental health add urgency to the calls for broader social and economic transformation, and for a more equitable and sustainable world order.

In summary, there is an urgent need for a significant expansion of the agenda of global mental health, building on its achievements while also recognizing the limitations of its extant principles and strategies, by seeking to incorporate novel understandings of mental health and the alignment of mental health with the SDGs. At the heart of this reimagining of global mental health is a reframing of the goals of global mental health, from reducing the treatment gap for people affected by mental disorders to the more ambitious aim of enhancing the mental health of populations and reducing the actual global burden of mental disorders. Realizing this vision will require combined actions of promotion of population mental health, prevention of mental disorders, and effective medical and social care for people with mental disorders.

Discussion Questions

1. It is difficult to place mental health high on the public health agendas of LMICs that face an enormous burden from communicable diseases. Which evidence-based arguments might you make to challenge the notion that mental health is a luxury item on the health agendas of such countries?
2. The classification of mental disorders is mainly derived from the description of these disorders in high-income countries. Some argue that this fact limits the application of psychiatric knowledge and evidence to non-Western cultures.

How valid are these concerns? In which ways has the "Western" bias been addressed in classification of mental disorders in international public health?

3. Diseases that disproportionately affect the poor are typically prioritized by governments and donors. Some people believe that disorders such as depression are problems of the middle class and the affluent, or represent a "medicalization of misery," and, therefore, do not deserve a share of scarce resources. What is the evidence linking poverty with mental disorders? How might poverty interact with mental health?
4. While there is now a growing evidence base on effective treatments for most mental disorders, large treatment gaps can be found in all countries, especially in LMICs. What are the reasons for this treatment gap? How can they be addressed at the level of health policy and health service development?
5. Human rights are a major driver in global health. Even though people affected by mental disorders represent one of the most marginalized and discriminated groups in any context, their human rights are often ignored. Why is this the case, and what can be done to address this issue?

References

Aben, I., Lodder, J., Honig, A., Lousberg, R., Boreas, A., & Verhey, F. (2006). Focal or generalized vascular brain damage and vulnerability to depression after stroke: A 1-year prospective follow-up study. *International Psychogeriatrics, 18*(1), 19–35.

Alem, A. (2002). Community-based vs. hospital-based mental health care: The case of Africa. *World Psychiatry, 1*(2), 98.

Alem, A., Jacobsson, L., & Hanlon, C. (2008). Community-based mental health care in Africa: Mental health workers' views. *World Psychiatry, 7*(1), 54–57.

Alvarado, R., & Rojas, G. (2011). El programa nacional para el diagnóstico y tratamiento de depresión en atención primaria: una evaluación necesaria. *Revista médica de Chile, 139,* 592–599

Alzheimer's Disease International. (2015). *World Alzheimer report 2015: The global impact of dementia*. London, UK: Author.

American Psychiatric Association (APA). (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed). Washington, DC: Author.

American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Publishers.

Anda, R. F., Felitti, V. J., Bremner, J. D., Walker, J. D., Whitfield, C., Perry, B. D., . . . Giles, W. H. (2006). The enduring effects of abuse and related adverse experiences in childhood: A convergence of evidence from neurobiology and epidemiology. *European Archives of Psychiatry & Clinical Neuroscience, 256*(3), 174–186.

Anderson, I., Robson, B., Connolly, M., Al-Yaman, F., Bjertness, E., King, A., . . . Yap, L. (2016). Indigenous and tribal peoples' health (*The Lancet*–Lowitja Institute Global Collaboration): A population study. *The Lancet, 388*(10040), 131–157. doi:10.1016/S0140-6736(16)00345-7

Andreoli, S. B., Almeida-Filho, N., Martin, D., Mateus, M. D. M., & Mari, J. d. J. (2007). Is psychiatric reform a strategy for reducing the mental health budget? The case of Brazil. *Revista Brasileira de Psiquiatria, 29*(1), 43–46.

Asher, L., Fekadu, A., Teferra, S., De Silva, M., Pathare, S., & Hanlon, C. (2017). "I cry every day and night, I have my son tied in chains": Physical restraint of people with schizophrenia in community settings in Ethiopia. *Globalization and Health, 13*(1), 47. doi:10.1186/s12992-017-0273-1

Balaji, M., Chatterjee, S., Koschorke, M., Rangaswamy, T., Chavan, A., Dabholkar, H., . . . Patel, V. (2012). The development of a lay health worker delivered collaborative community based intervention for people with schizophrenia in India. *BMC Health Services Research, 12*, 42. doi:http://dx.doi.org/10.1186/1472-6963-12-42

Barry, M. M., Clarke, A. M., Jenkins, R., & Patel, V. (2013). A systematic review of the effectiveness of mental health promotion interventions for young people in low and middle income countries. *BMC Public Health, 13*(1), 835. doi:10.1186/1471-2458-13-835

Basu, S. (2016). Madras lunatic asylum: A remarkable history in British India. *Indian Journal of History of Science, 51*, 478–493.

Baxter, A., Scott, K., Vos, T., & Whiteford, H. (2013). Global prevalence of anxiety disorders: A systematic review and meta-regression *Psychological Medicine, 43,* 897–910.

Baxter, A. J., Vos, T., Scott, K. M., Ferrari, A. J., & Whiteford, H. A. (2014). The global burden of anxiety disorders in 2010. *Psychological Medicine, 44*(11), 2363–2374. doi:10.1017/S0033291713003243

Becker, A. E., Burwell, R. A., Gilman, S. E., Herzog, D. B., & Hamburg, P. (2002). Disordered eating behaviors and attitudes follow prolonged exposure to television among ethnic Fijian adolescent girls. *British Journal of Psychiatry, 180*, 509–514.

Bell, C., Williamson, J., & Chien, P. (2008). Cultural, racial and ethnic competence and psychiatric diagnosis. *Ethnicity and Inequalities in Health and Social Care, 1*(1), 34–41.

Berlin, I. N. (1969). Resistance to change in mental health professionals. *American Journal of Orthopsychiatry, 39*(1), 109–115.

Bilder, R. M. (2015). Dimensional and categorical approaches to mental illness: Let biology decide. In L. J. Kirmayer, R. Lemelson, & C. A. Cummings (Eds.), *Re-visioning psychiatry: Cultural phenomenology, critical neuroscience, and global mental health* (pp. 179–205). New York, NY: Cambridge University Press.

Björkenstam, E., Ljung, R., Burström, B., Mittendorfer-Rutz, E., Hallqvist, J., & Weitoft, G. R. (2012). Quality of medical care and excess mortality in psychiatric patients: A nationwide register-based study in Sweden. *BMJ Open, 2*(1). doi:10.1136/bmjopen-2011-000778

Blacker, D., & Tsuang, M. T. (1992). Contested boundaries of bipolar disorder. *American Journal of Psychiatry, 149*(11), 1473–1483.

Bloom, D. E., Cafiero, E. T., Jane-Llopis, E., Abrahams-Gessel, S., Bloom, L. R., Fathima, S., . . . Weinstein, C. (2011). *The global economic burden of non-communicable diseases.* Geneva, Switzerland: World Health Organization.

Bolton, P., Bass, J., Betancourt, T., Speelman, L., Onyango, G., Clougherty, K. F., . . . Verdeli, H. (2007). Interventions for depression symptoms among adolescent survivors of war and displacement in northern Uganda: A randomized controlled trial. *Journal of the American Medical Association, 298*(5), 519–527.

Bolton, P., Bass, J., Neugebauer, R., Verdeli, H., Clougherty, K., Wickramaratne, P., . . . Weissman, M. (2003). Group interpersonal psychotherapy for depression in rural Uganda: A randomized controlled trial. *Journal of the American Medical Association, 289*(23), 3117–3124.

Breuer, E., De Silva, M. J., Fekadu, A., Luitel, N. P., Murhar, V., Nakku, J., . . . Lund, C. (2014). Using workshops to develop theories of change in five low and middle income countries: Lessons from the Programme for Improving Mental Health Care (PRIME). *International Journal of Mental Health Systems, 8*, 15. doi:10.1186/1752-4458-8-15

Carstairs, G. M. (1973). Psychiatric problems of developing countries. *British Journal of Psychiatry, 123*(574), 271–277.

Chang-Quan, H., Zheng-Rong, W., Yong-Hong, L., Yi-Zhou, X., & Qing-Xiu, L. (2010). Education and risk for late life depression: A meta-analysis of published literature. *International Journal of Psychiatry in Medicine, 40*(1), 109–124.

Charlson, F. J., Baxter, A. J., Cheng, H. G., Shidhaye, R., & Whiteford, H. A. (2016). The burden of mental, neurological, and substance use disorders in China and India: A systematic analysis of community representative epidemiological studies. *The Lancet*. doi:10.1016/S0140-6736(16)30590-6

Charlson, F. J., Diminic, S., Lund, C., Degenhardt, L., & Whiteford, H. A. (2014). Mental and substance use disorders in sub-Saharan Africa: Predictions of epidemiological changes and mental health workforce requirements for the next 40 years. *PloS One, 9*(10). doi:10.1371/journal.pone.0110208

Charlson, F. J., Moran, A. E., Freedman, G., Norman, R. E., Stapelberg, N. J. C., Baxter, A. J., . . . Whiteford, H. A. (2013). The contribution of major depression to the global burden of ischemic heart disease: A comparative risk assessment. *BMC Medicine, 11*(1). doi:10.1186/1741-7015-11-250

Charlson, F. J., Stapelberg, N. J. C., Baxter, A. J., & Whiteford, H. A. (2011). Should global burden of disease estimates include depression as a risk factor for coronary heart disease? *BMC Medicine, 9*, 47. doi:10.1186/1741-7015-9-47

Chatterjee, S., Naik, S., John, S., Dabholkar, H., Balaji, M., Koschorke, M., . . . Thornicroft, G. (2014). Effectiveness of a community-based intervention for people with schizophrenia and their caregivers in India (COPSI): A randomised controlled trial. *The Lancet*. doi:10.1016/S0140-6736(13)62629-X

Chatterjee, S., Patel, V., Chatterjee, A., & Weiss, H. (2003). Evaluation of a community based rehabilitation model for chronic schizophrenia in a rural region of India. *British Journal of Psychiatry, 182*, 57–62.

Chatterjee, S., Pillai, A., Jain, S., Cohen, A., & Patel, V. (2009). Outcomes of people with psychotic disorders in a community-based rehabilitation programme in rural India. *British Journal of Psychiatry, 195*(5), 433–439.

Chibanda, D., Weiss, H. A., Verhey, R., Simms, V., Munjoma, R., Rusakaniko, S., . . . Araya, R. (2016). Effect of a primary care–based psychological intervention on symptoms of common mental disorders in Zimbabwe: A randomized clinical trial. *Journal of the American Medical Association, 316*(24), 2618–2626. doi:10.1001/jama.2016.19102

Chisholm, D., Heslin, M., Docrat, S., Nanda, S., Shidhaye, R., Upadhaya, N., . . . Lund, C. (2016). Scaling-up services for psychosis, depression and epilepsy in sub-Saharan Africa and South Asia: Development and application of a mental health systems planning tool (OneHealth). *Epidemiology and Psychiatric Sciences, 26*(3), 234–244. doi:10.1017/S2045796016000408

Chisholm, D., Naci, H., Hyder, A. A., Tran, N. T., & Peden, M. (2012). Cost effectiveness of strategies to combat road traffic injuries in sub-Saharan Africa and South East Asia: Mathematical modelling study. *BMJ, 344*, e612.

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: A global return on investment analysis. *The Lancet Psychiatry*. doi:10.1016/S2215-0366(16)30024-4

Clifford, T. (1990). *Tibetan Buddhist medicine and psychiatry*. York Beach, ME: Samuel Weiser.

Cohen, A., Eaton, J., Radtke, B., George, C., Manuel, B. V., De Silva, M., & Patel, V. (2011). Three models of community mental health services in low-income countries. *International Journal of Mental Health Systems, 5*(1), 3. doi:10.1186/1752-4458-5-3

Cohen, A., & Minas, H. (2016). Global mental health and psychiatric institutions in the 21st century. *Epidemiology and Psychiatric Sciences*, 1–6. doi:10.1017/S2045796016000652

Cohen, A., Padmavati, R., Hibben, M., Oyewusi, S., John, S., Esan, O., . . . Morgan, C. (2016). Concepts of madness in diverse settings: A qualitative study from the INTREPID project. *BMC Psychiatry, 16*(1), 388. doi:10.1186/s12888-016-1090-4

Cole, S. W. (2014). Human social genomics. *PLoS Genetics, 10*(8), e1004601. doi:http://dx.doi.org/10.1371/journal.pgen.1004601

Collins, P. Y., Patel, V., Joestl, S., March, D., Insel, T. R., & Daar, A. S. (2011). Grand challenges in global mental health. *Nature, 475*, 27–30.

Compton, M. T., Broussard, B., Munetz, M., Oliva, J. R., & Watson, A. C. (2011). *The crisis intervention team (CIT) model of collaboration between law enforcement and mental health*. New York, NY: Novinka/Nova Science Publishers.

Corrigan, P. W., Morris, S. B., Michaels, P. J., Rafacz, J. D., & Rusch, N. (2012). Challenging the public stigma of mental illness: A meta-analysis of outcome studies. *Psychiatric Services, 63*(10), 963–973. doi:http://dx.doi.org/10.1176/appi.ps.201100529

Creed, F. H., Davies, I., Jackson, J., Littlewood, A., Chew-Graham, C., Tomenson, B., . . . McBeth, J. (2012). The epidemiology of multiple somatic symptoms. *Journal of Psychosomatic Research, 72*(4), 311–317. doi:10.1016/j.jpsychores.2012.01.009

Davidson, L., Hoge, M. A., Godleski, L., Rakfeldt, J., & Griffith, E. E. H. (1996). Hospital or community living? Examining consumer perspectives on deinstitutionalization. *Psychiatric Rehabilitation Journal, 19*(3), 49–58.

Davies, T., & Lund, C. (2017). Integrating mental health care into primary care systems in low- and middle-income countries: Lessons from PRIME and AFFIRM. *Global Mental Health, 4*, e7. doi:10.1017/gmh.2017.3

Davies, T., Schneider, M., Nyatsanza, M., & Lund, C. (2016). "The sun has set even though it is morning": Experiences and explanations of perinatal depression in an urban township, Cape Town. *Transcult Psychiatry, 53*(3), 286–312. doi:10.1177/1363461516632389

Degenhardt, L., Charlson, F., Stanaway, J., Larney, S., Alexander, L. T., Hickman, M., . . . Vos, T. (2016). Estimating the burden

of disease attributable to injecting drug use as a risk factor for HIV, hepatitis C, and hepatitis B: findings from the Global Burden of Disease Study 2013. *The Lancet Infectious Diseases, 16*(12), 1385–1398. doi:10.1016/s1473-3099(16)30325-5

Degenhardt, L., & Hall, W. (2012). Extent of illicit drug use and dependence, and their contribution to the global burden of disease. *The Lancet, 379*(9810), 55–70.

de Jong, J. T. V. M., Komproe, I. H., & Ommeren, M. V. (2003). Common mental disorders in postconflict settings. *The Lancet, 361*, 2128–2130.

Digby, A. (1985). Moral treatment at the Retreat, 1796–1846. In W. F. Bynum, R. Porter, & M. Shepherd (Eds.), *The anatomy of madness: Essays in the history of psychiatry* (Vol. II, pp. 52–72). London, UK: Tavistock Publications.

Diminic, S., Hielscher, E., Lee, Y. Y., Harris, M., Schess, J., Kealton, J., & Whiteford, H. (2016). The economic value of informal mental health caring in Australia: Technical report. Retrieved from https://www.mindaustralia.org.au/sites/default/files/Mind_value_of_informal_caring_full_report.pdf

Dipnall, J. F., Pasco, J. A., Meyer, D., Berk, M., Williams, L. J., Dodd, S., & Jacka, F. N. (2015). The association between dietary patterns, diabetes and depression. *Journal of Affective Disorders, 174*, 215–224. doi:10.1016/j.jad.2014.11.030

Dols, M. W. (1987). Insanity and its treatment in Islamic society. *Medical History, 31*(1), 1–14.

Druss, G. B., Rosenheck, A. R., Desai, M. M., & Perlin, B. J. (2002). Quality of preventive medical care for patients with mental disorders. *Medical Care, 40*(2), 129–136. doi:10.1097/00005650-200202000-00007

Edington, C. (2013). Going in and getting out of the colonial asylum: Families and psychiatric care in French Indochina. *Comparative Studies in Society and History, 55*(3), 725–755.

Ehsan, A. M., & De Silva, M. J. (2015). Social capital and common mental disorder: A systematic review. *Journal of Epidemiology and Community Health, 69*(10), 1021–1028. doi:10.1136/jech-2015-205868

El-Islam, M. F. (1982). Arabic cultural psychiatry. *Transcultural Psychiatry, 19*(1), 5–24.

Escobar, J. I., Waitzkin, H., Silver, R. C., Gara, M., & Holman, A. (1998). Abridged somatization: A study in primary care. *Psychosomatic Medicine, 60*(4), 466–472.

Fakhoury, W., & Priebe, S. (2007). Deinstitutionalization and reinstitutionalization: Major changes in the provision of mental healthcare. *Psychiatry, 6*(8), 313–316.

Fazel, M., Patel, V., Thomas, S., & Tol, W. (2014). Mental health interventions in schools in low-income and middle-income countries. *The Lancet Psychiatry, 1*(5), 388–398. doi:http://dx.doi.org/10.1016/S2215-0366(14)70357-8

Ferrari, A., Charlson, F., Norman, R., Patten, S., Freedman, G., Murray, C., . . . Whiteford, H. (2013). Burden of depressive disorders by country, sex, age, and year: Findings from the Global Burden of Disease Study 2010. *PLoS Medicine, 10*(11), e1001547. doi:10.1371/journal.pmed.1001547

Ferrari, A. J., Norman, R. E., Freedman, G., Baxter, A. J., Pirkis, J. E., Harris, M. G., . . . Whiteford, H. A. (2014). The burden attributable to mental and substance use disorders as risk factors for suicide: Findings from the Global Burden of Disease Study 2010. *PloS One, 9*(4), e91936. doi:10.1371/journal.pone.0091936

Fisch, R. (1965). Resistance to change in the psychiatric community: Alternatives. *Archives of General Psychiatry, 13*(4), 359–366.

Flaherty, J. A., Gaviria, F. M., Pathak, D., Mitchell, T., Wintrob, R., Richman, J. A., & Birz, S. (1988). Developing instruments for cross-cultural psychiatric research. *Journal of Nervous & Mental Disease, 176*(5), 257–263.

Fusar-Poli, P., Meneghelli, A., Valmaggia, L., Allen, P., Galvan, F., McGuire, P., & Cocchi, A. (2009). Duration of untreated prodromal symptoms and 12-month functional outcome of individuals at risk of psychosis. *British Journal of Psychiatry, 194*(2), 181–182. doi:10.1192/bjp.bp.107.047951

GBD 2015 DALYs and HALE Collaborators. (2016). Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388*(10053), 1603–1658. doi:10.1016/S0140-6736(16)31460-X

GBD 2015 Risk Factors Collaborators. (2016). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388*(10053), 1659–1724. doi:10.1016/S0140-6736(16)31679-8

German, G. A. (1972). Aspects of clinical psychiatry in sub-Saharan Africa. *British Journal of Psychiatry, 121*, 461–479.

Gilbert, B. J., Patel, V., Farmer, P. E., & Lu, C. (2015). Assessing development assistance for mental health in developing countries: 2007–2013. *PLoS Medicine, 12*(6), e1001834. doi:10.1371/journal.pmed.1001834

Glisson, C., & Williams, N. J. (2015). Assessing and changing organizational social contexts for effective mental health services. *Annual Review of Public Health, 36*, 507–523. doi:10.1146/annurev-publhealth-031914-122435

Goddard, M. (1992). Bedlam in paradise: A critical history of psychiatry in Papua New Guinea. *Journal of Pacific History, 27*, 55–72.

Goldman, H. H. (1983). The demography of deinstitutionalization. *New Directions for Mental Health Services, 17*, 31–40.

Goldstein, J. (2001). *Console and classify: The French psychiatric profession in the nineteenth century*. Chicago, IL: University of Chicago Press.

Gorenstein, E. E. (1992). *The science of mental illness*. San Diego, CA: Academic Press.

Greenblatt, M. (1992). Deinstitutionalization and reinstitutionalization of the mentally ill. In M. J. Robertson & M. Greenblatt (Eds.), *Homelessness: A national perspective* (pp. 47–56). New York, NY: Plenum Press.

Grigsby, A. B., Anderson, R. J., Freedland, K. E., Clouse, R. E., & Lustman, P. J. (2002). Prevalence of anxiety in adults with diabetes: A systematic review. *Journal of Psychosomatic Research, 53*(6), 1053–1060.

Grob, G. N. (1994). *The mad among us: A history of the care of America's mentally ill*. New York, NY: Free Press.

Gureje, O., Simon, G. E., Ustun, T. B., & Goldberg, D. P. (1997). Somatization in cross-cultural perspective: A World Health Organization study in primary care. *American Journal of Psychiatry, 154*(7), 989–995.

Gwaikolo, W. S., Kohrt, B. A., & Cooper, J. L. (2017). Health system preparedness for integration of mental health services in rural Liberia. *BMC Health Services Research, 17*(1), 508. doi:10.1186/s12913-017-2447-1

Hacking, I. (1998). *Mad travelers: Reflections on the reality of transient mental illnesses*. Charlottesville, VA: University Press of Virginia. Retrieved from http://proxy.library.emory.edu/login?url=http://www.netLibrary.com/urlapi.asp?action=summary&v=1&bookid=17064

Hahn, R. A. (1995). *Sickness and healing: An anthropological perspective*. New Haven, CT: Yale University Press.

Hanlon, C., Wondimagegn, D., & Alem, A. (2010). Lessons learned in developing community mental health care in Africa. *World Psychiatry, 9*(3), 185–189.

Harding, T. W. (1973). Psychosis in a rural West African community. *Social Psychiatry, 8*, 198–203.

Harding, T. W., De Arango, M. V., Baltazar, J., Climent, C. E., Ibrahim, H. H. A., Ladrigo-Ignacio, L., . . . Wig, N. N. (1980). Mental disorders in primary health care: A study of their frequency in four developing countries. *Psychological Medicine, 10*, 231–241.

Haroz, E. E., Ritchey, M., Bass, J. K., Kohrt, B. A., Augustinavicius, J., Michalopoulos, L., . . . Bolton, P. (2017). How is depression experienced around the world? A systematic review of qualitative literature. *Social Science and Medicine, 183*, 151–162. doi:10.1016/j.socscimed.2016.12.030

Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). Most people are not WEIRD. *Nature, 466*(7302), 29–29.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). The weirdest people in the world? *Behavioral & Brain Sciences, 33*(2–3), 61–83; discussion 83–135. doi:http://dx.doi.org/10.1017/S0140525X0999152X

Hert, M., Cohen, D., Bobes, J., Cetkovich-Bakmas, M., Leucht, S., Ndetei, D. M., . . . Correll, C. U. (2011). Physical illness in patients with severe mental disorders. II. Barriers to care, monitoring and treatment guidelines, plus recommendations at the system and individual level. *World Psychiatry, 10*(2), 138–151. doi:10.1002/j.2051-5545.2011.tb00036.x

Hilty, D. M., Alverson, D. C., Alpert, J. E., Tong, L., Sagduyu, K., Boland, R. J., . . . Yellowlees, P. M. (2006). Virtual reality, telemedicine, web and data processing innovations in medical and psychiatric education and clinical care. *Academic Psychiatry, 30*, 528–533.

Hinton, D. E., Hofmann, S. G., Pollack, M. H., & Otto, M. W. (2009). Mechanisms of efficacy of CBT for Cambodian refugees with PTSD: Improvement in emotion regulation and orthostatic blood pressure response. *CNS Neuroscience & Therapeutics, 15*(3), 255–263. doi:10.1111/j.1755-5949.2009.00100.x

Hinton, D. E., & Lewis-Fernandez, R. (2011). The cross-cultural validity of posttraumatic stress disorder: Implications for DSM-5. *Depression and Anxiety, 28*(9), 783–801. doi:10.1002/da.20753

Hinton, D. E., Rivera, E. I., Hofmann, S. G., Barlow, D. H., & Otto, M. W. (2012). Adapting CBT for traumatized refugees and ethnic minority patients: Examples from culturally adapted CBT (CA-CBT). *Transcultural Psychiatry, 49*(2), 340–365. doi:http://dx.doi.org/10.1177/1363461512441595

Holt, R. I., Bushe, C., & Citrome, L. (2005). Diabetes and schizophrenia 2005: Are we any closer to understanding the link? *Journal of Psychopharmacology, 19*(6 suppl), 56–65.

Howes, D. (2005). *Empire of the senses: The sensual culture reader*. Oxford, UK: Berg.

Hruschka, D. J., & Hadley, C. (2008). A glossary of culture in epidemiology. *Journal of Epidemiology & Community Health, 62*(11), 947–951.

Huhn, M., Tardy, M., Spineli, L. M., Kissling, W., Foerstl, H., Pitschel-Walz, G., . . . Leucht, S. (2014). Efficacy of pharmacotherapy and psychotherapy for adult psychiatric disorders: A systematic overview of meta-analyses. *JAMA Psychiatry, 71*(6), 706–715. doi:10.1001/jamapsychiatry.2014.112

Humanitas. (1827). Lunatic asylums. *The Lancet, 9*(224), 432.

Human Rights Watch. (2015). Callous and cruel: Use of force against inmates with mental disabilities in US jails and prisons. Retrieved from https://www.hrw.org/report/2015/05/12/callous-and-cruel/use-force-against-inmates-mental-disabilities-us-jails-and

Human Rights Watch. (2016). Double punishment: Inadequate conditions for prisoners with psychosocial disabilities in France. Retrieved from https://www.hrw.org/report/2016/04/05/double-punishment/inadequate-conditions-prisoners-psychosocial-disabilities-france

Imai, H., Hosomi, J., Nakao, H., Tsukino, H., Katoh, T., Itoh, T., & Yoshida, T. (2005). Characteristics of psychiatric hospitals associated with length of stay in Japan. *Health Policy, 74*(2), 115–121.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., . . . Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry, 167*(7), 748–751.

Institute of Medicine. (2001). *Neurological, psychiatric and developmental disorders: Meeting the challenge in the developing world*. Washington, DC: National Academy Press.

Inter-Agency Standing Committee (IASC). (2007). *Guidelines on mental health and psychosocial support in emergency settings*. Geneva, Switzerland: Author.

Jablensky, A., Sartorius, N., Ernberg, G., Anker, M., Korten, A., Cooper, J. E., . . . Bertelsen, A. (1992). Schizophrenia: Manifestations, incidence and course in different cultures: A World Health Organization ten-country study. *Psychological Medicine, 20*(suppl 20), 1–97.

Jackson, L. A. (1999). The place of psychiatry in colonial and early postcolonial Zimbabwe. *International Journal of Mental Health, 28*(2), 38–71.

Jilek, W. G. (1995). Emil Kraepelin and comparative sociocultural psychiatry. *European Archives of Psychiatry and Clinical Neuroscience, 245*(4–5), 231–238. doi:10.1007/BF02191802

Jilek, W. G., & Jilek-Aall, L. (1970). Transient psychoses in Africans. *Psychiatrica Clinica, 3*, 337–364.

Johansson, C., Astrom, S., Kauffeldt, A., Helldin, L., & Carlstrom, E. (2014). Culture as a predictor of resistance to change: A study of competing values in a psychiatric nursing context. *Health Policy, 114*(2–3), 156–162. doi:10.1016/j.healthpol.2013.07.014

Jordans, M. J., Kohrt, B. A., Luitel, N. P., Komproe, I. H., & Lund, C. (2015). Accuracy of proactive case finding for mental disorders by community informants in Nepal. *British Journal of Psychiatry, 207*(6), 501–506. doi:10.1192/bjp.bp.113.141077

Jordans, M. J. D., Kohrt, B. A., Luitel, N. P., Lund, C., & Komproe, I. H. (2017). Proactive community case-finding to facilitate treatment seeking for mental disorders, Nepal. *Bulletin of the World Health Organization, 95*(7), 531–536. doi:10.2471/BLT.16.189282

Kahn, R. S., Wise, P. H., Kennedy, B. P., & Kawachi, I. (2000). State income inequality, household income, and maternal mental and physical health: cross-sectional national survey. *British Medical Journal, 321*, 1315–1331.

Kaiser, B. N., Haroz, E. E., Kohrt, B. A., Bolton, P. A., Bass, J. K., & Hinton, D. E. (2015). "Thinking too much": A systematic review of a common idiom of distress. *Social Science and Medicine, 147*, 170–183. doi:10.1016/j.socscimed.2015.10.044

Kakuma, R., Minas, H., van Ginneken, N., Dal Poz, M. R., Desiraju, K., Morris, J. E., . . . Scheffler, R. M. (2011). Human resources for mental health care: Current situation and strategies for action. *The Lancet, 378*(9803), 1654–1663.

Keller, R. C. (2008). *Colonial madness: Psychiatry in French North Africa.* Chicago, IL: University of Chicago Press.

Kendler, K. S., & Gardner, C. O., Jr. (1998). Boundaries of major depression: An evaluation of DSM-IV criteria. *American Journal of Psychiatry, 155*(2), 172–177.

Kim, A. M. (2017). Why do psychiatric patients in Korea stay longer in hospital? *International Journal of Mental Health Systems, 11*(1), 2. doi:10.1186/s13033-016-0110-6

Kirkbride, J. B., Fearon, P., Morgan, C., Dazzan, P., Morgan, K., Tarrant, J., . . . Leff, J. P. (2006). Heterogeneity in incidence rates of schizophrenia and other psychotic syndromes: Findings from the 3-center Aesop Study. *Archives of General Psychiatry, 63*(3), 250–258. doi:10.1001/archpsyc.63.3.250

Kirmayer, L. J., & Gold, I. (2012a). Critical neuroscience and the limits of reductionism. In S. Choudhary & J. Slaby (Eds.), *Critical neuroscience: A handbook of the social and cultural contexts of neuroscience* (pp. 307–330). Oxford, UK: Blackwell.

Kirmayer, L. J., & Gold, I. (2012b). Re-socializing psychaitry: Critical neuroscience and the limits of reductionism. In S. Choudhary & J. Slaby (Eds.), *Critical neuroscience: A handbook of the social and cultural contexts of neuroscience* (pp. 307–330). Oxford, UK: Blackwell.

Kirmayer, L. J., & Minas, H. (2000). The future of cultural psychiatry: An international perspective. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie, 45*(5), 438–446.

Kirmayer, L. J., & Swartz, L. (2014). Culture and global mental health. In V. Patel, H. Minas, A. Cohen, & M. J. Prince (Eds.), *Global mental health: Principles and practice* (pp. 41–62). New York, NY: Oxford University Press.

Kleinman, A. (1980). *Patients and healers in the context of culture: An exploration of the borderland between anthropology, medicine, and psychiatry*. Berkeley, CA: University of California Press.

Kleinman, A. (1987). Anthropology and psychiatry: The role of culture in cross-cultural research on illness. *British Journal of Psychiatry, 151*, 447–454.

Kleinman, A. (1988). *Rethinking psychiatry: From cultural category to personal experience*. New York, NY: Free Press/Collier Macmillan.

Kleinman, A., & Good, B. (1985). *Culture and depression: Studies in the anthropology and cross-cultural psychiatry of affect and disorder*. Berkeley, CA: University of California Press.

Knapp,M.,McDaid,D.,&Parsonage,M.(2011).*Mentalhealthpromotion and mental illness prevention: The economic case. IDEAS Working Paper Series from RePEc*. London, UK: Department of Health. Retrieved from http://www.lse.ac.uk/businessAndConsultancy/LSEEnterprise/pdf/PSSRUfeb2011.pdf

Kohrt, B. A., Blasingame, E., Compton, M. T., Dakana, S. F., Dossen, B., Lang, F., . . . Cooper, J. (2015). Adapting the crisis intervention team (CIT) model of police–mental health collaboration in a low-income, post-conflict country: Curriculum development in Liberia, West Africa. *American Journal of Public Health, 105*(3), e73–e80. doi:10.2105/AJPH.2014.302394

Kohrt, B. A., Hruschka, D. J., Kohrt, H. E., Panebianco, N. L., & Tsagaankhuu, G. (2004). Distribution of distress in post-socialist Mongolia: A cultural epidemiology of yadargaa. *Social Science & Medicine, 58*(3), 471–485.

Kohrt, B. A., Jordans, M. J. D., Rai, S., Shrestha, P., Luitel, N. P., Ramaiya, M., . . . Patel, V. (2015). Therapist competence in global mental health: Development of the Enhancing Assessment of Common Therapeutic Factors (ENACT) rating scale. *Behaviour Research and Therapy, 69*, 11–21. doi:http://dx.doi.org/10.1016/j.brat.2015.03.009

Kohrt, B. A., Jordans, M. J., Tol, W. A., Luitel, N. P., Maharjan, S. M., & Upadhaya, N. (2011). Validation of cross-cultural child mental health and psychosocial research instruments: Adapting the Depression Self-Rating Scale and Child PTSD Symptom Scale in Nepal. *BMC Psychiatry, 11*(1), e127.

Kohrt, B. A., Mendenhall, E., & Brown, P. J. (2016). How anthropological theory and methods can advance global mental health. *The Lancet Psychiatry, 3*(5), 396–398. doi:10.1016/S2215-0366(16)00046-8

Kohrt, B. A., Rasmussen, A., Kaiser, B. N., Haroz, E. E., Maharjan, S. M., Mutamba, B. B., . . . Hinton, D. E. (2014). Cultural concepts of distress and psychiatric disorders: Literature review and research recommendations for global mental health epidemiology. *International Journal of Epidemiology, 43*(2), 365–406. doi:10.1093/ije/dyt227

Konnopka, A., Schaefert, R., Heinrich, S., Kaufmann, C., Luppa, M., Herzog, W., & Konig, H.-H. (2012). Economics of medically unexplained symptoms: A systematic review of the literature. *Psychotherapy and Psychosomatics, 81*(5), 265–275. doi:10.1159/000337349

Kulhara, P., & Chandiramani, K. (1988). Outcome of schizophrenia in India using various diagnostic systems. *Schizophrenia Research, 1*(5), 339–349.

Kulhara, P., & Wig, N. N. (1978). The chronicity of schizophrenia in North West India: Results of a follow-up study. *British Journal of Psychiatry, 132*, 186–190.

Lambo, T. A. (1960). Further neuropsychiatric observations in Nigeria with comment on the need for further epidemiological study in Africa. *British Medical Journal, 2*, 1696–1704.

Latthe, P., Mignini, L., Gray, R., Hills, R., & Khan, K. (2006). Factors predisposing women to chronic pelvic pain: Systematic review. *BMJ, 332*(7544), 749–755.

Laursen, T. M., Nordentoft, M., & Mortensen, P. B. (2014). Excess early mortality in schizophrenia. *Annual Review of Clinical Psychology, 10*, 425–448. doi:10.1146/annurev-clinpsy-032813-153657

Lawrence, D., Hancock, K. J., & Kisely, S. (2013). The gap in life expectancy from preventable physical illness in psychiatric patients in Western Australia: Retrospective analysis of population based registers. *British Medical Journal, 346*(1). doi:10.1136/bmj.f2539

Lee, S. (2001). From diversity to unity: The classification of mental disorders in 21st century China. *Psychiatric Clinics, 24*(3), 421-431. doi:10.1016/S0193-953X(05)70238-0.

Leighton, A. H., Lambo, T. A., Hughes, C. C., Leighton, D. C., Murphy, J. M., & Macklin, D. B. (1963). *Psychiatric disorder among the Yoruba: A report from the Cornell–Aro mental health research project in the Western Region, Nigeria*. Ithaca, NY: Cornell University Press.

Leon, C. A. (1972). Psychiatry in Latin America. *British Journal of Psychiatry, 121*, 121–136.

Levy, C. J. (2002a, April 28). For mentally ill, death and misery. *New York Times,* p. 28.

Levy, C. J. (2002b, April 29). Here, life is squalor and chaos. *New York Times,* p. 29.

Levy, C. J. (2002c, April 28). Ingredients of a failing system: A lack of state money, a group without a voice. *New York Times,* p. 32.

Levy, C. J. (2002d, April 30). Voiceless, defenseless and a source of cash. *New York Times,* p. A1.

Li, Z., Page, A., Martin, G., & Taylor, R. (2011). Attributable risk of psychiatric and socio-economic factors for suicide from individual-level, population-based studies: A systematic review. *Social Science & Medicine, 72*(4), 608–616. doi:10.1016/j.socscimed.2010.11.008

Lichtman, H. J., Froelicher, S. E., Blumenthal, A. J., Carney, M. R., Doering, V. L., Frasure-Smith, E. N., . . . Wulsin, S. L. (2014). Depression as a risk factor for poor prognosis among patients with acute coronary syndrome: Systematic review and recommendations: A scientific statement from the American Heart Association. *Circulation, 129*(12), 1350–1369. doi:10.1161/CIR.0000000000000019

Lin, E. H. B., Rutter, C. M., Katon, W., Heckbert, S. R., Ciechanowski, P., Oliver, M. M., . . . Von Korff, M. (2010). Depression and advanced complications of diabetes: A prospective cohort study. *Diabetes Care, 33*(2), 264. doi:10.2337/dc09-1068

Lin, T. Y., Rin, H., Yeh, E. K., Hsu, C. C., & Chu, H. M. (1969). Mental disorders in Taiwan, fifteen years later. In W. Caudill & T. Y. Lin (Eds.), *Mental health research in Asia and Pacific* (pp. 66–91). Honolulu, HI: East–West Centre Press.

Littlewood, R. (2002). *Pathologies of the West: An anthropology of mental illness in Europe and America*. Ithaca, NY: Cornell University Press.

Lobo, A., de-la-Camara, C., Campos, R., & Saz, P. (2015). Innovative methods in teaching psychiatry to medical students. *European Journal of Psychiatry, 29*(1), 67–78.

Louden, D. (1995). The epidemiology of schizophrenia among Caribbean-born and first- and second-generation migrants in Britain. *Journal of Social Distress and the Homeless, 4*(3), 237–253. doi:10.1007/BF02088020

Lovell, A. M. (2014). The World Health Organization and the contested beginnings of psychiatric epidemiology as an international discipline: One rope, many strands. *International Journal of Epidemiology, 43*(suppl 1), i6–i18. doi:10.1093/ije/dyu125

Lund, C., Breen, A., Flisher, A. J., Kakuma, R., Corrigall, J., Joska, J. A., . . . Patel, V. (2010). Poverty and common mental disorders in low and middle income countries: A systematic review. *Social Science & Medicine, 71*(3), 517–528. doi:10.1016/j.socscimed.2010.04.027

Lund, C., De Silva, M., & Stansfeld, S. (2014). Social determinants of mental health. In V. Patel, M. Prince, A. Cohen, & H. Minas (Eds.), *Global mental health: Principles and practice* (pp. 116–136). London, UK: Oxford University Press.

Lund, C., Tomlinson, M., De Silva, M., Fekadu, A., Shidhaye, R., Jordans, M., . . . Prince, M. (2012). PRIME: A programme to reduce the treatment gap for mental disorders in five low-and middle-income countries. *PLoS Medicine, 9*(12), e1001359.

Ma, H. (2012). Integration of hospital and community services—the "686 Project"—is a crucial component in the reform of China's mental health services. *Shanghai Archives of Psychiatry, 24*(3), 172–174. doi:10.3969/j.issn.1002-0829.2012.03.007.

Makanjuola, R. O. (1987). "Ode Ori": A culture-bound disorder with prominent somatic features in Yoruba Nigerian patients. *Acta Psychiatrica Scandinavica, 75*(3), 231–236.

Martin, E. (2007). *Bipolar expeditions: Mania and depression in American culture.* Princeton, NJ: Princeton University Press.

Mayston, R., Habtamu, K., Medhin, G., Alem, A., Fekadu, A., Habtamu, A., . . . Hanlon, C. (2017). Developing a measure of mental health service satisfaction for use in low income countries: A mixed methods study. *BMC Health Services Research, 17*(1), 183. doi:10.1186/s12913-017-2126-2

McGorry, P. D. (2015). Early intervention in psychosis: Obvious, effective, overdue. *Journal of Nervous and Mental Disease, 203*(5), 310–318. doi:10.1097/nmd.0000000000000284

McGrath, J., Saha, S., Welham, J., El Saadi, O., MacCauley, C., & Chant, D. (2004). A systematic review of the incidence of schizophrenia: The distribution of rates and the influence of sex, urbanicity, migrant status and methodology. *BMC Medicine, 2*(1), 13. doi:10.1186/1741-7015-2-13

Mendenhall, E., Kohrt, B. A., Norris, S. A., Ndetei, D., & Prabhakaran, D. (2017). Non-communicable disease syndemics: Poverty, depression, and diabetes among low-income populations. *The Lancet, 389*(10072), 951–963. doi:10.1016/S0140-6736(17)30402-6

Mental Health Innovation Network (MHIN). (n.d.). Program for screening, diagnosis and comprehensive treatment of depression. Retrieved from http://mhinnovation.net/innovations/program-screening-diagnosis-and-comprehensive-treatment-depression#.VVYpd46qqkp

Meyer, M. (2010). Sanity in the South Atlantic: The mythos of Philippe Pinel and the asylum movement in nineteenth-century Rio de Janeiro. *Atlantic Studies, 7*(4), 473–492.

Minas, H., & Diatri, H. (2008). Pasung: Physical restraint and confinement of the mentally ill in the community. *International Journal of Mental Health Systems, 2*(1), 8. doi:10.1186/1752-4458-2-8

Mora, G. (2008). Mental disturbances, unusual mental states, and their interpretation during the Middle Ages. In E. R. Wallace & J. Gach (Eds.), *History of psychiatry and medical psychology* (pp. 199–226). New York, NY: Springer US.

Morosini, P. L., Repetto, F., De Salvia, D., & Cecere, F. (1985). Psychiatric hospitalization in Italy before and after 1978. *Acta Psychiatrica Scandinavica, 71*(S316), 27–43. doi:10.1111/j.1600-0447.1985.tb08511.x

Moussaoui, D., & Glick, I. D. (2015). The Maristan "Sidi Fredj" in Fez, Morocco. *American Journal of Psychiatry, 172*(9), 838–839. doi:10.1176/appi.ajp.2015.15010051

Mundt, A. P., Chow, W. S., Arduino, M., Barrionuevo, H., Fritsch, R., Girala, N., . . . Priebe, S. (2015). Psychiatric hospital beds and prison populations in South America since 1990: Does the Penrose hypothesis apply? *JAMA Psychiatry, 72*(2), 112–118. doi:10.1001/jamapsychiatry.2014.2433

Murphy, H. B., & Raman, A. C. (1971). The chronicity of schizophrenia in indigenous tropical peoples: Results of a twelve-year follow-up survey in Mauritius. *British Journal of Psychiatry, 118*(546), 489–497.

Murray, C. J., Vos, T., Lozano, R., Naghavi, M., Flaxman, A. D., Michaud, C., . . . Memish, Z. A. (2012). Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. *The Lancet, 380*(9859), 2197–2223. doi:S0140-6736(12)61689-4

Murray, L. K., Dorsey, S., Haroz, E., Lee, C., Alsiary, M. M., Haydary, A., . . . Bolton, P. (2014). A common elements treatment approach for adult mental health problems in low- and middle-income countries. *Cognitive and Behavioral Practice, 21*(2), _111–123. doi:http://dx.doi.org/10.1016/j.cbpra.2013.06.005

Mutamba, B. B., van Ginneken, N., Smith Paintain, L., Wandiembe, S., & Schellenberg, D. (2013). Roles and effectiveness of lay community health workers in the prevention of mental, neurological and substance use disorders in low and middle income countries: A systematic review. *BMC Health Services Research, 13*(1), 412. doi:10.1186/1472-6963-13-412

Nadkarni, A., Weobong, B., Weiss, H. A., McCambridge, J., Bhat, B., Katti, B., . . . Patel, V. (2016). Counselling for Alcohol Problems (CAP), a lay counsellor–delivered brief psychological treatment for harmful drinking in men, in primary care in India: A randomised controlled trial. *The Lancet, 389*(10065), 186–195. doi:10.1016/S0140-6736(16)31590-2

Nanni, M. G., Caruso, R., Mitchell, A. J., Meggiolaro, E., & Grassi, L. (2015). Depression in HIV infected patients: A review. *Current Psychiatry Reports, 17*(1), 530. doi:10.1007/s11920-014-0530-4

Naslund, J. A., Aschbrenner, K. A., Araya, R., Marsch, L. A., Unützer, J., Patel, V., & Bartels, S. J. (2017). Digital technology for treating and preventing mental disorders in low-income and middle-income countries: A narrative review of the literature. *The Lancet Psychiatry, 4*(6), 486–500. doi:10.1016/S2215-0366(17)30096-2

Neki, J. S. (1973). Psychiatry in South-East Asia. *British Journal of Psychiatry, 123*(574), 257–269.

Ng, B. Y., & Chee, K. T. (2006). A brief history of psychiatry in Singapore. *International Review of Psychiatry, 18*(4), 355–361.

Nichter, M. (1981). Idioms of distress: Alternatives in the expression of psychosocial distress: A case study from South India. *Culture, Medicine & Psychiatry, 5*(4), 379–408.

Organisation for Economic Co-operation and Development (OECD). (2015). *Mental health and work: Fit mind, fit job: From evidence to practice in mental health and work*. Paris, France: OECD Publishing.

Pan, A., Sun, Q., Okereke, O. I., Rexrode, K. M., & Hu, F. B. (2011). Depression and risk of stroke morbidity and mortality: A meta-analysis and systematic review. *Journal of the American MedicalAssociation,306*(11),1241–1249.doi:10.1001/jama.2011.1282

Patel, V. (2017). Talking sensibly about depression. *PLoS Medicine, 14*(4), e1002257. doi:10.1371/journal.pmed.1002257

Patel, V. & Hanlon, C. (2018) *Where There Is No Psychiatrist* (2nd edition). Cambridge University Press, UK.

Patel, V., Araya, R., de Lima, M., Ludermir, A., & Todd, C. (1999). Women, poverty and common mental disorders in four restructuring societies. *Social Science & Medicine, 49*(11), 1461.

Patel, V., Belkin, G. S., Chockalingam, A., Cooper, J., Saxena, S., & Unutzer, J. (2013). Integrating mental health services into priority health care platforms: Addressing a grand challenge in global mental health. *PLoS Medicine, 10*, e1001448.

Patel, V., Chisholm, D., Parikh, R., Charlson, F. J., Degenhardt, L., Dua, T., . . . Whiteford, H. (2016). Addressing the burden of mental, neurological, and substance use disorders: Key messages from Disease Control Priorities, 3rd edition. *The Lancet, 387*(10028), 1672–1685. doi:10.1016/S0140-6736(15)00390-6

Patel, V., Farooq, S., & Thara, R. (2007). What is the best approach to treating schizophrenia in developing countries? *PLoS Medicine, 4*(6), e159.

Patel , V., & Prince, M. (2010). Global mental health: A new global health field comes of age. *Journal of the American Medical Association, 303*, 1976–1977.

Patel, V., Rodrigues, M., & DeSouza, N. (2002). Gender, poverty, and postnatal depression: A study of mothers in Goa, India. *American Journal of Psychiatry, 159*(1), 43–47.

Patel, V., Simon, G., Chowdhary, N., Kaaya, S., & Araya, R. (2009). Packages of care for depression in low- and middle-income countries. *PLoS Medicine, 6*(10), e1000159.

Patel, V., Simunyu, E., & Gwanzura, F. (1995). Kufungisisa (thinking too much): A Shona idiom for non-psychotic mental illness. *Central African Journal of Medicine, 41*(7), 209–215.

Patel, V., Weiss, H. A., Chowdhary, N., Naik, S., Pednekar, S., Chatterjee, S., . . . Kirkwood, B. R. (2010). Effectiveness of an intervention led by lay health counsellors for depressive and anxiety disorders in primary care in Goa, India (MANAS): A cluster randomised controlled trial. *The Lancet, 376*(9758), 2086–2095.

Patel, V., Weiss, H. A., Kirkwood, B. R., Pednekar, S., Nevrekar, P., Gupte, S., & Mabey, D. (2006). Common genital complaints in women: The contribution of psychosocial and infectious factors in a population-based cohort study in Goa, India. *International Journal of Epidemiology, 35*, 1478–1485.

Patel, V., Weobong, B., Weiss, H. A., Anand, A., Bhat, B., Katti, B., . . . Fairburn, C. G. (2017). The Healthy Activity Program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: A randomised controlled trial. *The Lancet, 389*(10065), 176-185. doi:10.1016/s0140-6736(16)31589-6

Patel, V., Xiao, S., Chen, H., Hanna, F., Jotheeswaran, A. T., Luo, D., . . . Saxena, S. (2017). The magnitude of and health system responses to the mental health treatment gap in adults in India and China. *The Lancet, 388*(10063), 3074–3084. doi:10.1016/S0140-6736(16)00160-4

Petersen, I., Fairall, L., Bhana, A., Kathree, T., Selohilwe, O., Brooke-Sumner, C., . . . Patel, V. (2016). Integrating mental health into chronic care in South Africa: The development of a district mental healthcare plan. *British Journal of Psychiatry, 208*(suppl 56), s29–s39. doi:10.1192/bjp.bp.114.153726

Pierloot, R. A. (1975). Belgium. In J. G. Howells (Ed.), *World history of psychiatry* (pp. 136–149). New York, NY: Brunner/Mazel.

Pols, H. (2006). The development of psychiatry in Indonesia: From colonial to modern times. *International Review of Psychiatry, 18*(4), 363–370. doi:10.1080/09540260600775421

Priebe, S., Badesconyi, A., Fioritti, A., Hansson, L., Kilian, R., Torres-Gonzales, F., . . . Wiersma, D. (2005). Reinstitutionalisation in mental health care: Comparison of data on service provision from six European countries. *BMJ, 330*(7483), 123–126. doi:10.1136/bmj.38296.611215.AE

Prince, M., Patel, V., Saxena, S., Maj, M., Maselko, J., Phillips, M. R., & Rahman, A. (2007). No health without mental health. *The Lancet, 370*, 859–877.

Pyne, H. H., Claeson, M., & Correia, M. (2002). *Gender dimensions of alcohol consumption and alcohol-related problems in Latin America and the Caribbean* (Discussion Paper No. 433). Washington, DC: World Bank.

Rahman, A., Fisher, J., Bower, P., Luchters, S., Tran, T., Yasamy, M. T., . . . Waheed, W. (2013). Interventions for common perinatal mental disorders in women in low- and middle-income countries: A systematic review and meta-analysis. *Bulletin of the World Health Organization, 91*(8), 593–601i. doi:10.2471/blt.12.109819

Rahman, A., Malik, A., Sikander, S., Roberts, C., & Creed, F. (2008). Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: A cluster-randomized controlled trial. *The Lancet, 372*, 902–909.

Reed, G. M., Roberts, M. C., Keeley, J., Hooppell, C., Matsumoto, C., Sharan, P., . . . Medina-Mora, M. E. (2013). Mental health professionals' natural taxonomies of mental disorders: Implications for the clinical utility of the ICD-11 and the DSM-5. *Journal of Clinical Psychology, 69*(12), 1191–1212. doi:10.1002/jclp.22031

Rehm, J. (2011). The risks associated with alcohol use and alcoholism. *Alcohol Research & Health, 34*(2), 135. doi:Fea-AR&H-65

Rehm, J., Mathers, C., Popova, S., Thavorncharoensap, M., Teerawattananon, Y., & Patra, J. (2009). Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. *The Lancet, 373*(9682), 2223–2233.

Rotella, F., & Mannucci, E. (2013). Depression as a risk factor for diabetes: A meta-analysis of longitudinal studies. *Journal of Clinical Psychiatry, 74*(1), 31–37. doi:10.4088/JCP.12r07922

Royal College of Psychiatrists. (2012). *Joint call for parity of esteem on World Mental Health Day*. London, UK: Author.

Royal College of Psychiatrists. (2013). *Whole-person care: From rhetoric to reality*. London, UK: Author.

Rule, N. O., Freeman, J. B., & Ambady, N. (2013). Culture in social neuroscience: A review. *Social Neuroscience, 8*(1), 3–10. doi:10.1080/17470919.2012.695293

Sadowsky, J. (1997). Psychiatry and colonial ideology in Nigeria. *Bulletin of the History of Medicine, 71*, 94–111.

Saraceno, B., van Ommeren, M., Batniji, R., Cohen, A., Gureje, O., Mahoney, J., . . . Underhill, C. (2007). Barriers to improvement of mental health services in low-income and middle-income countries. *The Lancet, 370*(9593), 1164–1174. doi:10.1016/S0140-6736(07)61263-X

Sartorius, N., & Harding, T. W. (1983). The WHO collaborative study on strategies for extending mental health care, I: The genesis of the study. *American Journal of Psychiatry, 140*(11), 1470–1473. doi:10.1176/ajp.140.11.1470

Sartorius, N., Kaelber, C. T., Cooper, J. E., Roper, M. T., Rae, D. S., Gulbinat, W., . . . Regier, D. A. (1993). Progress toward achieving a common language in psychiatry. *Archives of General Psychiatry, 50*, 115–124.

Saxena, S., Funk, M., & Chisholm, D. (2013). World Health Assembly adopts Comprehensive Mental Health Action Plan 2013–2020. *The Lancet, 381*(9882), 1970–1971. doi:10.1016/S0140-6736(13)61139-3

Saxena, S., Thornicroft, G., Knapp, M., & Whiteford, H. (2007). Resources for mental health: Scarcity, inequity, and inefficiency. *The Lancet, 370*(9590), 878–889.

Scarfone, M. (2016). Italian colonial psychiatry: Outlines of a discipline, and practical achievements in Libya and the Horn of Africa. *History of Psychiatry, 27*(4), 389–405. doi:10.1177/0957154X16659853

Schmidt, K. E. (1967). Mental health services in a developing country in South-East Asia (Sarawak). In H. C. Freeman & J. Farndale (Eds.), *New aspects of the mental health services* (pp. 213–236). Oxford, UK: Pergamon Press.

Schofield, D. J., Shrestha, R. N., Percival, R., Passey, M. E., Callander, E. J., & Kelly, S. J. (2011). The personal and national costs of mental health conditions: Impacts on income, taxes, government support payments due to lost labour force participation. *BMC Psychiatry, 11*, 72–72. doi:10.1186/1471-244X-11-72

Scull, A. (1985). Deinstitutionalization and public policy. *Social Science & Medicine, 20*(5), 545–552.

Scull, A. (1989). *Social order/mental disorder: Anglo-American psychiatry in historical perspective*. Berkeley, CA: University of California Press.

Scull, A. (2015). *Madness in civilization.* Princeton, NJ: Princeton University Press.

Selten, J.-P., & Hoek, H. (2008). Does misdiagnosis explain the schizophrenia epidemic among immigrants from developing countries to Western Europe? *Social Psychiatry and Psychiatric Epidemiology, 43*(12), 937–939. doi:10.1007/s00127-008-0390-5

Semrau, M., Barley, E. A., Law, A. N. N., & Thornicroft, G. (2011). Lessons learned in developing community mental health care in Europe. *World Psychiatry, 10*(3), 217–225. doi:10.1002/j.2051-5545.2011.tb00060.x

Semrau, M., Evans-Lacko, S., Alem, A., Ayuso-Mateos, J. L., Chisholm, D., Gureje, O., . . . Thornicroft, G. (2015). Strengthening mental health systems in low- and middle-income countries: The Emerald programme. *BMC Medicine, 13*(1), 1–9. doi:10.1186/s12916-015-0309-4

Sharan, P., Gallo, C., Gureje, O., Lamberte, E., Mari, J. J., Mazzotti, G., . . . Saxena, S. (2009). Mental health research priorities in low- and middle-income countries of Africa, Asia, Latin America and the Caribbean. *British Journal of Psychiatry, 195*(4), 354–363. doi:10.1192/bjp.bp.108.050187

Simons, R. C., & Hughes, C. C. (Eds.). (1985). *The culture bound syndromes: Folk illnesses of psychiatric and anthropological interest*. Boston, MA: D. Reidel Publishing.

Singh, M. (2017). The friendship bench can help chase the blues away. Retrieved from http://www.npr.org/sections/goatsandsoda/2017/01/10/508588401/the-friendship-bench-can-help-chase-the-blues-away

Singh, S. P., Greenwood, N., White, S., & Churchill, R. (2007). Ethnicity and the Mental Health Act 1983. *British Journal of Psychiatry, 191*(2), 99–105. doi:10.1192/bjp.bp.106.030346

Singla, D., Kohrt, B. A., Murray, L. K., Anand, A., Chorpita, B. F., & Patel, V. (2017). Psychological treatments for the world: Lessons from low- and middle-income countries. *Annual Review of Clinical Psychology, (13)*, 5.1-5.33. doi:10.1146/annurev-clinpsy-032816-045217

Sornpaisarn, B., Shield, K., Cohen, J., Schwartz, R., & Rehm, J. (2013). Elasticity of alcohol consumption, alcohol-related harms, and drinking initiation in low-and middle-income countries: A systematic review and meta-analysis. *International Journal of Alcohol and Drug Research, 2*(1), 45–58.

Stein, A., Pearson, R. M., Goodman, S. H., Rapa, E., Rahman, A., McCallum, M., . . . Pariante, C. M. (2014). Effects of perinatal mental disorders on the fetus and child. *The Lancet, 384*(9956), 1800–1819. doi:10.1016/S0140-6736(14)61277-0

Stein, D. J., He, Y., Phillips, A., Sahakian, B. J., Williams, J., & Patel, V. (2015). Global mental health and neuroscience: potential synergies. *The Lancet Psychiatry, 2*(2), 178–185. doi:http://dx.doi.org/10.1016/S2215-0366(15)00014-0

Stein, D. J., Lund, C., & Nesse, R. M. (2013). Classification systems in psychiatry: Diagnosis and global mental health in the era of DSM-5 and ICD-11. *Current Opinion in Psychiatry, 26*(5), 493–497. doi:10.1097/YCO.0b013e3283642dfd

Strik, J. J., Lousberg, R., Cheriex, E. C., & Honig, A. (2004). One year cumulative incidence of depression following myocardial infarction and impact on cardiac outcome. *Journal of Psychosomatic Research, 56*(1), 59–66.

Swartz, S. (1999). "Work of mercy and necessity": British rule and psychiatric practice in the Cape Colony, 1891–1910. *International Journal of Mental Health, 28*(2), 72–90.

Tansella, M., & Thornicroft, G. (2001). The principles underlying community care. In G. Thornicroft & G. Szmukler (Eds.), *Textbook of community psychiatry* (pp. 155–165). Oxford, UK: Oxford University Press.

Thalmayer, A. G., Friedman, S. A., Azocar, F., Harwood, J. M., & Ettner, S. L. (2017). The Mental Health Parity and Addiction Equity Act (MHPAEA) evaluation study: Impact on quantitative treatment limits. *Psychiatric Services, 68*(5), 435–442. doi:10.1176/appi.ps.201600110

Thara, R., & Padmavati, R. (2013). Community mental health care in South Asia. *World Psychiatry, 12*(2), 176–177. doi:10.1002/wps.20042

The Lancet. (1857). The crime of lunacy, and how we punish it. *The Lancet, 70*(1768), 62–64.

The Lancet. (1875). Report of *The Lancet* Commission on lunatic asylums. *The Lancet, 106*(2727), 816–820. doi:http://dx.doi.org/10.1016/S0140-6736(02)30721-9

The Lancet. (1877). Report of *The Lancet* Commission on lunatic asylums. *The Lancet, 109*(2796), 464–467.

The Lancet Global Mental Health Group. (2007). Scaling up services for mental disorders: A call for action. *The Lancet, 370*, 1241–1252.

Thirthalli, J., Channaveerachari, N., Subbakrishna, D., Cottler, L., Varghese, M., & Gangadhar, B. (2011). Prospective study

of duration of untreated psychosis and outcome of never-treated patients with schizophrenia in India. *Indian Journal of Psychiatry, 53*(4), 319–323. doi:10.4103/0019-5545.91905

Thornicroft, G., Alem, A., Antunes Dos Santos, R., Barley, E., Drake, R. E., Gregorio, G., . . . Wondimagegn, D. (2010). WPA guidance on steps, obstacles and mistakes to avoid in the implementation of community mental health care. *World Psychiatry, 9*(2), 67–77.

Thornicroft, G., Alem, A., Dos Santos, R. A., Barley, E., Drake, R. E., Gregorio, G., . . . Wondimagegn, D. (2010). WPA guidance on steps, obstacles and mistakes to avoid in the implementation of community mental health care. *World Psychiatry, 9*(2), 67–77. doi:10.1002/j.2051-5545.2010.tb00276.x

Thornicroft, G., & Bebbington, P. (1989). Deinstitutionalisation—from hospital closure to service development. *British Journal of Psychiatry, 155*(6), 739–753. doi:10.1192/bjp.155.6.739

Thornicroft, G., Chatterji, S., Evans-Lacko, S., Gruber, M., Sampson, N., Aguilar-Gaxiola, S., . . . Kessler, R. C. (2017). Undertreatment of people with major depressive disorder in 21 countries. *British Journal of Psychiatry, 210*(2), 119–124. doi:10.1192/bjp.bp.116.188078

Thornicroft, G., Mehta, N., Clement, S., Evans-Lacko, S., Doherty, M., Rose, D., . . . Henderson, C. (2016). Evidence for effective interventions to reduce mental-health–related stigma and discrimination. *The Lancet, 387*(10023), 1123–1132. doi:10.1016/S0140-6736(15)00298-6

Thornicroft, G., Tansella, M., & Law, A. (2008). Steps, challenges and lessons in developing community mental health care. *World Psychiatry, 7*(2), 87–92.

Tol, W. A., Patel, V., Tomlinson, M., Baingana, F., Galappatti, A., Panter-Brick, C., . . . van Ommeren, M. (2011). Research priorities for mental health and psychosocial support in humanitarian settings. *PLoS Medicine, 8*(9), e1001096.

Tomlinson, M., Rudan, I., Saxena, S., Swartz, L., Tsai, A. C., & Patel, V. (2009). Setting priorities for global mental health research. *Bulletin of the World Health Organization, 87*, 438–446.

Truong, K. D., & Ma, S. (2006). A systematic review of relations between neighborhoods and mental health. *Journal of Mental Health Policy and Economics, 9*, 137–154.

Tsuang, M. T., Stone, W. S., & Faraone, S. V. (2000). Toward reformulating the diagnosis of schizophrenia. *American Journal of Psychiatry, 157*(7), 1041–1050. doi:10.1176/appi.ajp.157.7.1041

van Ginneken, N., Tharyan, P., Lewin, S., Rao, G. N., Meera, S., Pian, J., . . . Patel, V. (2013). Non-specialist health worker interventions for the care of mental, neurological and substance-abuse disorders in low- and middle-income countries. *Cochrane Database of Systematic Reviews, 11*, CD009149. doi:10.1002/14651858.CD009149.pub2

Van Niel, C., Pachter, L. M., Wade, R., Jr., Felitti, V. J., & Stein, M. T. (2014). Adverse events in children: Predictors of adult physical and mental conditions. *Journal of Developmental & Behavioral Pediatrics, 35*(8), 549–551. doi:http://dx.doi.org/10.1097/DBP.0000000000000102

Van Ommeren, M., Sharma, B., Thapa, S., Makaju, R., Prasain, D., Bhattaria, R., & de Jong, J. T. V. M. (1999). Preparing instruments for transcultural research: Use of the translation monitoring form with Nepali-speaking Bhutanese. *Transcultural Psychiatry, 36*(3), 285–301.

Veling, W. (2013). Ethnic minority position and risk for psychotic disorders. *Current Opinion in Psychiatry, 26*(2), 166–171. doi:10.1097/YCO.0b013e32835d9e43

Verdeli, H., Clougherty, K., Bolton, P., Speelman, L., Lincoln, N., Bass, J., . . . Weissman, M. M. (2003). Adapting group interpersonal psychotherapy for a developing country: Experience in rural Uganda. *World Psychiatry, 2*(2), 114–120.

Vigo, D., Thornicroft, G., & Atun, R. (2016). Estimating the true global burden of mental illness. *The Lancet Psychiatry, 3*(2), 171–178. doi:10.1016/S2215-0366(15)00505-2

Wakefield, J. C., & Demazeux, S. (2015). *Sadness or depression? International perspectives on the depression epidemic and its meaning*. Amsterdam, Netherlands: Springer.

Walker, E., McGee, R. E., & Druss, B. G. (2015). Mortality in mental disorders and global disease burden implications: A systematic review and meta-analysis. *JAMA Psychiatry*. doi:10.1001/jamapsychiatry.2014.2502

Wang, P. S., Aguilar-Gaxiola, S., Alonso, J., Angermeyer, M. C., Borges, G., Bromet, E. J., . . . Wells, J. E. (2007). Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *The Lancet, 370*(9590), 841–850.

Waxler, N. E. (1979). Is outcome for schizophrenia better in nonindustrial societies? The case of Sri Lanka. *Journal of Nervous & Mental Disease, 167*, 144–158.

Weaver, L. J. (2017). Tension among women in North India: An idiom of distress and a cultural syndrome. *Culture, Medicine, and Psychiatry, 41*(1), 35–55. doi:10.1007/s11013-016-9516-5

Weich, S., & Lewis, G. (1998). Poverty, unemployment, and common mental disorders: Population based cohort study. *British Medical Journal, 317*, 115–119.

Weiner, D. B. (1992). Philippe Pinel's "Memoir on Madness" of December 11, 1794: A fundamental text of modern psychiatry. *American Journal of Psychiatry, 149*, 725–732.

Weiss, M. G. (1983). The treatment of insane patients in India in the lunatic asylums of the nineteenth century. *Indian Journal of Psychiatry, 25*(4), 312–316.

Weiss, M. G. (1997). Explanatory Model Interview Catalogue (EMIC): Framework for comparative study of illness. *Transcultural Psychiatry, 34*(2): 235–263.

Weiss, M. G., Desai, A., Jadhav, S., Gupta, L., Channabasavanna, S. M., Doongaji, D. R., & Behere, P. B. (1988). Humoral concepts of mental illness in India. *Social Science & Medicine, 27*(5), 471–477.

Weobong, B., Akpalu, B., Doku, V., Owusu-Agyei, S., Hurt, L., Kirkwood, B., & Prince, M. (2009). The comparative validity of screening scales for postnatal common mental disorder in Kintampo, Ghana. *Journal of Affective Disorders, 113*(1–2), 109–117. doi:http://dx.doi.org/10.1016/j.jad.2008.05.009

Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., . . . Vos, T. (2013). Global burden of disease attributable to mental and substance use disorders: Findings from the Global Burden of Disease Study 2010. *The Lancet, 382*(9904), 1575–1586. doi:10.1016/S0140-6736(13)61611-6

Whiteford, H., Ferrari, A., & Degenhardt, L. (2016). Global burden of disease studies: Implications for mental and substance use disorders. *Health Affairs, 35*(6), 1114–1120. doi:10.1377/hlthaff.2016.0082

Whiteford, H. A., Ferrari, A. J., Degenhardt, L., Feigin, V., & Vos, T. (2015). The global burden of mental, neurological and substance use disorders: An analysis from the Global Burden of Disease Study 2010. *PloS One, 10*(2), e0116820. doi:10.1371/journal.pone.0116820

Whiteford, H. A., Ferrari, A. J., & Vos, T. (2016). Challenges to estimating the true global burden of mental disorders. *The Lancet Psychiatry, 3,* 402–403.

Winkler, P., Barrett, B., McCrone, P., Csémy, L., Janoušková, M., & Höschl, C. (2016). Deinstitutionalised patients, homelessness and imprisonment: systematic review. *British Journal of Psychiatry, 208*(5), 421–428. doi:10.1192/bjp.bp.114.161943

Wolf, J. P., & Midanik, L. T. (2013). Alcohol policy. In M. G. Vaughn & B. E. Perron (Eds.), *Social work practice in the addictions* (pp. 231–248). New York, NY: Springer.

World Bank. (1993). *World development report 1993: Investing in health*. New York, NY: Oxford University Press.

World Health Organization (WHO). (1960). *Epidemiology of mental disorders: Eighth report of the Expert Committee on Mental Health*. Geneva, Switzerland: Author.

World Health Organization (WHO). (1973). *Report of the International Pilot Study of Schizophrenia*. Geneva, Switzerland: Author.

World Health Organization (WHO). (1975). *Organization of mental health services in developing countries: Sixteenth report of the WHO Expert Committee on Mental Health*. Geneva, Switzerland: Author.

World Health Organization (WHO). (1990). *International classification of diseases: 10th revision*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2001). *Atlas: Mental health resources in the world*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2003a). *Advocacy for mental health: Mental health policy and service guidance package*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2003b). *Mental health financing: Mental health policy and service guidance package*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2003c). *Organisation of services for mental health: Mental health policy and service guidance package*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2005a). Child and adolescent mental health policies and plans. Retrieved from http://www.who.int/mental_health/policy/Childado_mh_module.pdf

World Health Organization (WHO). (2005b). *Mental health policy, plans and programmes: Mental health policy and service guidance package*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2008). *Task shifting: Rational redistribution of tasks among health workforce teams: Global recommendations and guidelines*. Geneva, Switzerland: Author.

World Health Organisation (WHO). (2010). *mhGAP intervention guide for mental, neurological and substance use disorders in non-specialized health settings: Mental health Gap Action Programme (mhGAP)*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2012). WHO QualityRights Tool Kit. Retrieved from: www.who.int/mental_health/publications/QualityRights_toolkit/en/

World Health Organization (WHO). (2013a). *Building back better: Sustainable mental health care after emergencies*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2013b). Mental health action plan 2013–2020. Retrieved from http://www.who.int/mental_health/action_plan_2013/en/

World Health Organization (WHO). (2013c). WHO humanitarian response compendium of health priorities and WHO projects in consolidated appeals and response plans. Retrieved from http://www.who.int/hac/cap_2013_Feb.pdf

World Health Organization (WHO). (2014). *Mental health atlas*. Geneva, Switzerland: Author

World Health Organization (WHO). (2015). Mental health global action programme humanitarian intervention guide (mhGAP-HIG): Clinical management of mental, neurological and substance use conditions in humanitarian emergencies. Retrieved from http://www.who.int/mental_health/publications/mhgap_hig/en/

World Health Organization (WHO) & Gulbenkian Global Mental Health Platform. (2014). *Innovation in deinstitutionalization: A WHO expert survey* (9241506814). Geneva, Switzerland: WHO.

Wynaden, D., Orb, A., McGowan, S., & Downie, J. (2000). Are universities preparing nurses to meet the challenges posed by the Australian mental health care system? *Australian and New Zealand Journal of Mental Health Nursing, 9*(3), 138–146.

Young, A. (1995). *The harmony of illusions: Inventing post-traumatic stress disorder*. Princeton, NJ: Princeton University Press.

Zafar, S., Sikander, S., Hamdani, S. U., Atif, N., Akhtar, P., Nazir, H., . . . Rahman, A. (2016). The effectiveness of technology-assisted cascade training and supervision of community health workers in delivering the Thinking Healthy Program for perinatal depression in a post-conflict area of Pakistan: Study protocol for a randomized controlled trial. *Trials, 17*(1), 188. doi:10.1186/s13063-016-1308-2

CHAPTER 11

Environmental and Occupational Health

Alex Macmillan, Tord Kjellstrom, Kirk R. Smith, Ajay Pillarisetti, and Alistair Woodward

Introduction

This chapter describes the important current issues in environmental and occupational health, including the scope of those issues and the methods used to assess them. It first explores the definition of "environment" and its ways of affecting human health. In doing so, it takes note of (1) several key disciplinary perspectives, (2) the international spectrum and burden of environmental health issues, and (3) the ongoing emergence of larger-scale environmental problems. The significant conceptual development in approaches to environmental health is made apparent—namely, the shift from considering the environment as a source of hazard to considering the Earth as habitat, a perspective that requires more integrative approaches to ecological sustainability and health. The main conceptual and methodological issues that relate to environmental health research and public health action are described. Five of the most pressing current issues linking environment and health are examined in illustrative case studies. We finish by considering future prospects, issues, and priorities for environmental health research and practice.

The word "environment" is broad and elastic in scope. In this chapter, "environment" refers to those external physical, chemical, and biological exposures and processes that shape the health and well-being of both individuals and populations. Environmental health research and practice have evolved significantly over the past three decades, from a primary focus on viewing the environment as a unidirectional source of risk or hazard, to a greater reconnection with understanding humans as part of natural ecological systems that support human health, and with which we have bidirectional relationships. Such systems include hydrological, atmospheric, climate, and biological systems. Human-created built environment systems can also support flourishing, yet be a source of hazards for health.

These definitions of "environmental health" exclude social, cultural, and economic environments that are also important for health, and which are dealt with elsewhere in this text (see the *Culture, Behavior, and Health*; *Understanding and Acting on Social Determinants of Health and Health Equity*; and *Health and 'The Economy'* chapters). However, these different environments are closely intertwined, and they often mediate relationships between physical, microbiological, and chemical exposures and health. Although injury (e.g., road traffic injury, workplace injury, and injury in the home) is a significant outcome of human interactions with the physical environment, it has developed its own public health specialization and is addressed separately in this text (see the *Unintentional Injuries and Violence* chapter). Occupational health, which is most often focused on environmental health

in the workplace setting, remains closely linked with environmental health and is considered together with environmental health in this chapter.

Because of their philosophical, geographic, and topical breadth, environmental health research and practice are undertaken by a wide range of professional experts. These range from epidemiologists to qualitative researchers, from biochemists to ecologists, and from water quality officers to urban planners.

Environmental Hazards, Exposures, and Risks

When considering environmental hazards, the environment can be categorized in several ways for easier consideration, including in relation to environmental media (air, water, soil, and food), the economic sector (transport, land use, and energy generation), physical scale (local, regional, and global), setting (household, workplace, and urban environment), and disease outcomes (cancers, congenital anomalies, and others). A useful classification that comprises five categories is used here, defined jointly by physical scale and by setting:

1. Household
2. Workplace
3. Community
4. Regional
5. Global

This categorization is not strictly bounded; many relationships between environment and health occur across these scales, so an additional "cross-scale" category should be kept in mind. This final category recognizes that the scale at which an environmental health impact eventually occurs may not be the scale at which the hazard was initiated.

To demonstrate this categorization, consider the hierarchy of environmental health consequences of energy use. The impacts of energy production and use contribute significantly to the total human impact on the environment at each of the six previously mentioned levels. That is, the extraction, harvesting, processing, distribution, and use of fuels and other energy sources have major environmental impacts at all scales, from individual households to the globe itself. Combustion occurs locally, causing local air pollution, both indoors and outdoors, but it also contributes to regional acid rain and, on a global scale, to the accumulation of carbon dioxide as a heat-trapping greenhouse gas in the lower atmosphere.

Of note, the workplace has some characteristics that make it a special setting for environmental exposure (the combination of hazard and people being present), which has led to occupational health and safety developing as its own specialization. These characteristics include higher than ambient levels of chemical and mechanical hazards in confined environments; exposure to hazards that are not present outside the workplace (such as chemicals that are used only in manufacturing); and workers who are overall healthier and arguably less vulnerable to negative impacts of exposures than other sectors of the population (such as older adults and young children). Nevertheless, the links between occupational health and environmental health remain vital: Occupational studies have commonly led to a greater understanding of the health impacts of non-workplace environmental exposures, while the impacts of workplace environmental hazards almost universally have repercussions at the community level (either via hazardous wastes or through the wider impacts of workplace injury and death). We therefore consider occupational and environmental health and hazards together in this chapter.

In defining these hazards, it is worth noting that exposures to some hazards arise because of natural conditions, whereas others are due to human interventions. Natural exposures arise from seasonal, latitudinal, or altitudinal gradients in solar irradiation, extremes of hot and cold weather, the occurrence of physical disasters, and local micronutrient deficiencies in soil. While many of these natural hazards can be mitigated by specific intervention and design, some—such as the 2011 earthquakes in New Zealand and Japan, and the devastating tsunami in Japan—cannot be "kept under control" even in the highest-income countries.

TABLE 11-1 provides examples of the types of environmental health hazard exposures that can occur at the different scales discussed in this chapter. Note that exposures may come via air, water, food, or other media, depending on the specific situation. Studies of environmental health impacts will, therefore, require an understanding of a number of specialized scientific disciplines.

The Burden and Distribution of Disease from Environmental and Occupational Hazards

Conceptualizing the Public Health Impact of Environmental Exposures

There are many ways of attempting to capture and express the idea of public health impact. Suppose the results of a study show that the mortality rate within a

TABLE 11-1 Examples of Sources of Unhealthy Environmental Exposures

Scale	Household	Workplace	Community	Regional	Global
Types of Environmental Exposure					
Microbiological	Contaminated drinking water; poor sanitation and food safety	Spread of infectious diseases via air or needlesticks in health services	Contamination of local rivers and lakes from animal and human feces	Spread of contamination to major rivers	Large-scale spread of infectious diseases (e.g., malaria) due to climate change
Chemical	Exposure of children to lead from household paint	Solvents exposure from glues or paints in manufacturing industries	Urban air pollution from motor vehicles or industry	Confluence of urban air pollution into larger areas (e.g., Asian brown cloud)	Global spread of persistent organic pollutants; greenhouse gas emissions
Physical	Too-cold or too-hot indoor temperatures in households	Noisy machinery and processes; excessive workplace heat exposure	Traffic noise	Ionizing radiation from nuclear accidents (e.g., Chernobyl, 1986)	Increased ultraviolet (UV) radiation due to ozone layer depletion
Ergonomic	Injury risks in badly designed dwellings	Heavy lifting and injury risks caused by machinery	Traffic safety risks due to poor urban design		

community for a particular disease has increased by a factor of 1.3, with a 95% confidence interval of 1.1 to 1.6. What does this mean? If the disease in question is rare, the 1.3-fold increase may represent no more than a few extra cases per million exposed people per year. If the disease is common, however, the same relative risk carries a much greater population-level health impact. Approaches to quantify this impact include estimation of (1) the number of cases per year that are due to the exposure (also known as attributable cases), or (2) the number of cases that would be prevented if present exposures ceased, or (3) the number of cases that would occur over a longer time period (e.g., a lifetime) under different scenarios relating to population susceptibility and exposures. These types of measures are sometimes referred to as health impact estimates (Talbot et al., 2009).

In the last two decades, a considerable amount of work has been done to estimate the global and national burden of disease attributable to environmental exposures (Gakidou et al., 2017; Landrigan et al., 2017; Prüss-Üstün & Corvalán, 2006; Prüss-Üstün, Wolf, Corvalán, Bos, & Neira, 2016). Such estimates are difficult to construct, both because knowledge about disease etiology is incomplete and because of the latency between environmental exposures and chronic health outcomes. The complex relationships among environmental conditions, socioeconomic circumstances, demographic change, human movement, and health present further difficulties in estimating the environmental contribution to disease burden (Prüss-Üstün et al., 2016).

For example, the combination of population pressure and poverty among rural populations in low- and middle-income countries (LMICs) often leads to land degradation, deforestation, flooding, further impoverishment, and increased risks to health from infectious disease, food shortages, and nutritional deficiencies.

The situation of sub-Saharan Africa—with its persistent poverty, environmental stresses, and marginalization in the global economy—illustrates these complex relationships. Although significant improvements have occurred in sub-Saharan Africa's health, education, and living standards in the past quarter-century, the region remains far behind the rest of the world in advancing economic, environmental, and health goals. The final report on the United Nations' Millennium Development Goals noted that 41% of people in this region live in extreme poverty, 42% lack safe drinking water, and 69% lack proper sanitation (United Nations, 2015a). Child mortality rates in sub-Saharan African countries are more than 50% higher than those in the next highest region, and average life expectancy is about a decade less (World Health Organization [WHO], Global Health Observatory, 2016).

The Environmental Risk Transition

As discussed in the *Measures of Health and Disease in Populations* chapter, certain changes in disease patterns accompany economic development (the epidemiologic transition). In general, there is also a tendency for environmental health risks to shift during the economic development process in terms of geographical scale, types and severity of hazards, and disease outcomes. First, as described in the "environmental risk transition" (Mock et al., 2017; Smith & Ezzati, 2005), environmental hazards shift from the household to the community and then to regional and global scales. Environmental risks in low- and middle-income societies are dominated by poor food, water, and air quality at the household level from inadequate sanitation, contaminated water, and combustion of low-quality fuels (WHO, 1997). Some of the activities that help solve these problems act to transfer problems to the community level in the form of urban air pollution, hazardous waste, and chemical pollution. In high-income countries (HICs), where many community and household problems have been subjected to considerable control, problems have to some extent shifted to the global scale—for example, through greenhouse gas emissions and regional scale air pollution; at the same time, low-income communities within these countries are often still at risk from local environmental hazards (Mock et al., 2017; WHO, 2008).

A shift in disease patterns also occurs during the epidemiologic transition, from mainly infectious diseases to cancer and other non-infectious diseases. This disease shift is accompanied by a temporal shift. Many important infectious diseases—diarrhea, malaria, and measles, for example—have relatively short latency periods (hours to weeks) between exposure and development of disease. Cancer and other chronic non-infectious diseases, by comparison, often entail delays of several decades before they become manifest. Global processes such as anthropogenic climate change may involve even longer time periods. This temporal shift has important implications for research and social policy (McMichael, 2009), especially since political cycles result in short policy horizons.

Current Estimates of the Environmental Burden of Disease

Over time, estimates of the global burden of disease due to environmental factors have become more consistent, with these factors estimated to account for one-fourth to one-third of the total burden of disease in disability-adjusted life years (DALYs) (Landrigan et al., 2017; Prüss-Üstün & Corvalán, 2006; Prüss-Üstün et al., 2016; WHO, 1997). This standardized metric combines mortality and morbidity data in a manner suitable for international comparisons (see the *Measures of Health and Disease in Populations* chapter). In this section, three such estimates are discussed.

One of the most recent integrated comparative risk assessments of the global and regional burden of disease due to environmental risk factors was published by WHO in 2016 (Prüss-Üstün et al., 2016), based on data from 2012. Compared to previous estimates, the authors of this report included a broader range of environmental factors in their assessment—for instance, the built environment and climate change. They also employed a combination of methods, including comparative risk assessment where possible, limited epidemiological data, and transmission pathways. Where none of these were available, they utilized expert surveys. Their estimates of the total burden of disease attributable to environmental risk factors were consistent with the previous comprehensive assessments: 23% of deaths and 22% of DALYs. Like previous estimates, the WHO analysis almost certainly underestimates the environmental burden of disease, since only a small proportion of all risks possible are covered in the literature. For example, estimates of the health impacts of climate change have been limited in their scope thus far (WHO, 2014). **TABLE 11-2** summarizes the population-attributable fractions of disease from environmental factors from the WHO assessment, by stage of development (LMICs and HICs) and by disease group. Note the significant contribution of occupational risk factors across diseases.

In 2016, the Global Burden of Disease Risk Factor Collaborators reported their most recent analysis of the "behavioural, environmental, occupational, and metabolic" risk factors for disease—a global and national comparative risk assessment of 83 such

TABLE 11-2 Health Burden Due to Environmental Risk Factors in 2012 by Development Status (population-attributable fraction [PAF] of total disability-adjusted lost life years due to the environment, with diseases ranked by combined total global burden of disease in DALYs and environmental contribution)

Disease or Disease Cluster	PAF (%): LMICs	PAF (%): HICs	Main Environmental Risk Factors
Cardiovascular Diseases Ischemic heart disease Stroke	 39 46	 23 22	Household and ambient air pollution, secondhand tobacco smoke, chemicals, and environments encouraging physical inactivity (e.g., urban design)
Diarrheal diseases	58	13	Water, sanitation, hygiene, agricultural practices
Lower respiratory tract infections	37	1	Household and ambient air pollution, secondhand tobacco smoke
Cancers	16–48	13–21	Air pollution, chemical exposures, radiation, and occupational safety
Chronic obstructive pulmonary disease	38	12	Household air pollution, secondhand tobacco smoke, occupational safety
Road traffic injuries	42	17–42	Traffic system environments and land-use planning
Neonatal conditions	11	6	Air pollution, secondhand tobacco smoke, water and sanitation in birth settings
Malaria*	40–64	40–64	Environmental management of vector habitats, contact between humans and vector habitats
Non-Road Traffic Injuries Falls Fires, heat Drownings Other	 31 78 74 45	 26 42 54 30	Home, community, and occupational safety, climate change
Musculoskeletal disorders*	15–27	15–27	Occupational stressors and ergonomics, prolonged sitting, carrying water and solid fuels for household needs
Asthma*	44	44	Household and ambient air pollution, secondhand tobacco smoke, indoor damp and mold, occupational allergens
Depression*	12	12	Occupational stress, work–life imbalance, environmental degradation

*no comparative data available by income

Modified from Prüss-Üstün, Wolf et al. 2016. PAF Population Attributable Fraction, LMICs low-middle income countries, HICs high-income countries.

risk factors (Gakidou et al., 2017). The distinction between behavioral and environmental risk factors is a difficult one to make, especially since it is increasingly understood that individual behaviors are heavily influenced by physical environments. Within the top 30 risk factors contributing most heavily to the global burden of disease in terms of DALYs, the authors identified 9 environmental and occupational risk factors: ambient air pollution (7); household air pollution (10); unsafe water (16); unsafe sanitation (21); occupational injury (22); no access to handwashing facility (23); occupational carcinogens (24); lead exposure (28); and secondhand tobacco smoke (30). We would argue from the evidence, though, that other top risk factors also have major environmental components—for example, high body mass index (6) and low physical activity (25), which are increasingly understood as being underpinned by physical environments that build in unhealthy eating patterns and exclude activity from people's daily lives; and unsafe sex (17), an occupational risk for sex workers.

Another study conducted to estimate the size of the health problem due to environmental factors was the 2017 *The Lancet* Commission on Pollution

(**EXHIBIT 11-1**). In this case, environmental pollutants were defined as "unwanted, often dangerous, material that is introduced into the Earth's environment as the result of human activity, that threatens human health, and that harms ecosystems." Based on the work of the Global Burden of Disease project, *The Lancet* Commission estimated that approximately 9 million premature deaths per year occurred worldwide in 2015 that were attributable to pollution (Landrigan et al., 2017).

Across all recent burden of disease estimates, certain outcomes and diseases are either entirely or significantly associated with modifiable environmental conditions. Some diseases, such as those caused by intestinal nematodes (hookworm, roundworm, whipworm) and trachoma (*Chlamydia trachomatis*, the main infectious cause of blindness), are considered to be 100% attributable to modifiable environmental factors. In both of these examples, the diseases are related to sanitation, the availability of clean freshwater, hygiene, and housing conditions. Other outcomes that are significantly attributable to modifiable environmental factors include road traffic accidents, drowning, venomous animal bites, falls, and fires/heat as well as malaria, dengue, and asthma. In addition, the prognosis for nearly every disease will be better when people live in cleaner and well-organized home and community environments; thus, environmental quality is a factor to some extent in many, if not all, risk factors. Exhibit 11-1 provides information on more recent estimates of environmental factors and their contributions to the burden of disease, which includes some discussion of issues surrounding how such estimates are done.

Finally, an important caveat should be noted: The burden of disease due to factors in the environment is not directly related to statistical significance. That is, a small *p* value does not guarantee an important result (in terms of how strong the effect, or the impact on a population). Moreover, a large burden of disease does not always translate into a high priority in policy and prevention: Other relevant considerations are whether interventions are available to reduce exposures, and what the costs and concomitant benefits of putting these interventions into action would be.

EXHIBIT 11-1 *The Lancet* Commission on Pollution and Health

The Lancet Commission on Pollution and Health is one of several dozen such commissions over the past decade, though it was greeted by more media attention than were the reports of any of the previous commissions (Cohen, 2017; Nesbit, 2017; Scutti, 2017). This effort focused on the health impacts of just the most well-understood types of pollution exposures in LMICs and found that, in total, pollution exposures are responsible for a large part of the global burden of disease today, essentially similar in size to the burden associated with high blood pressure, smoking, and poor diets worldwide, for example. Its report also included a section on economics, indicating a large impact on the world's per capita income from the ill health that is created by pollution.

As indicated in the following table, *The Lancet* Commission found that close to 9 million premature deaths can be attributed to pollution exposures in recent years, nearly all in LMICs. This is approximately 17% of all annual deaths.

Estimated Global Deaths by Pollution Risk Factor and Age at Death, 2015

	GBD Study Best Estimate (95% CI)	WHO Best Estimate (95% CI)
Air (total)	6.5 (5.7–7.3)	6.5 (5.4–7.4)
Household air	2.9 (2.2–3.6)	4.3 (3.7–4.8)
Ambient particulate	4.2 (3.7–4.8)	3.0 (3.7–4.8)
Ambient ozone	0.3 (0.1–0.4)	-
Water (total)	1.8 (1.4–2.2)	0.8 (0.7–1.0)
Unsafe sanitation	0.8 (0.7–0.9)	0.3 (0.1–0.4)
Unsafe source	1.3 (1.0–1.4)	0.5 (0.2–0.7)
Occupational	0.8 (0.8–0.9)	0.4 (0.3–0.4)
Carcinogens	0.5 (0.5–0.5)	0.1 (0.1–0.1)

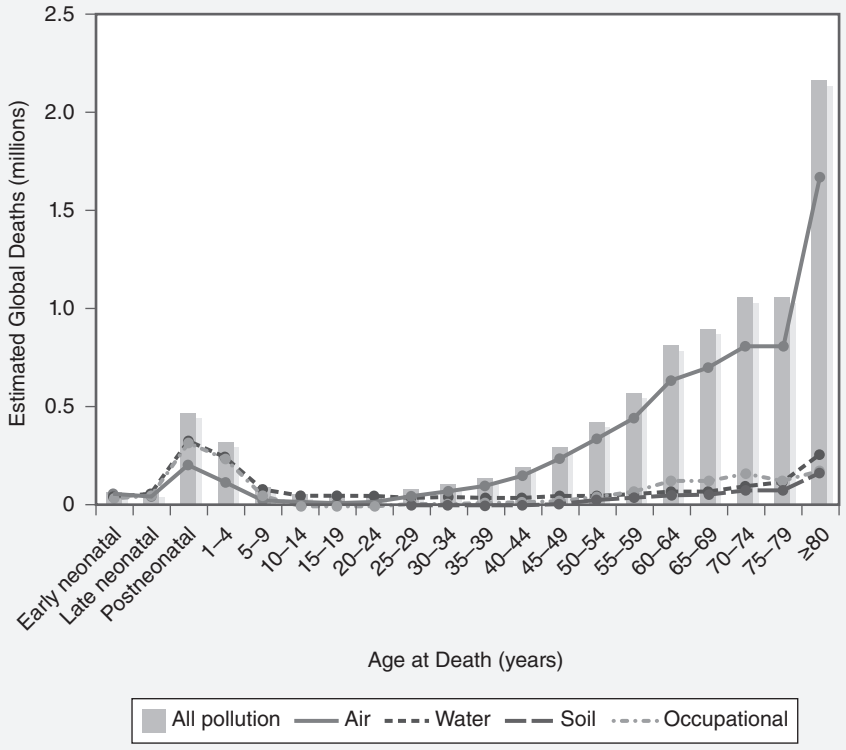
Particulates	0.4 (0.3–0.4)	0.2 (0.2–0.3)
Soil, heavy metals, and chemicals	0.5 (0.2–0.8)	0.7 (0.2–0.8)
Lead	0.5 (0.2–0.8)	0.7 (0.2–0.8)
Total	9	8.4

Note that the totals for air pollution, water pollution, and all pollution are less than the arithmetic sum of the individual risk factors within each of these categories because these have overlapping contributions—e.g., household air pollution also contributes to ambient air pollution and vice versa.

Also shown in the table, are the somewhat different estimates provided by the two main organizations doing global burden estimates, the WHO and the Institute for Health Metrics and Evaluation (IHME). The stated uncertainty bounds do not always overlap for an individual risk factor, although as it turns out the totals are similar. Of course, it is impossible to know exactly what these numbers are, so a better approach is probably to take the overall spread as the real range of estimates and not to pick precise values. The two organizations use some of the same methods and databases, but also differ in others. Thus, it woud be extremely difficult to decide which organization is more accurate for any one estimate.

Note that the totals in each major category are less than the total of the parts within that category. For example, the totals for air pollution from each organization are less than the sum of the individual components of air pollution. This difference arises because there is some overlap in estimated effects, due to both physical and epidemiologic reasons. The degree of overlap is not well known. Therefore, the totals are probably more uncertain than indicated.

Finally, although the press and the public are thought to best understand the number of deaths as an indicator of ill health, it is actually quite a poor indicator, particularly when combining effects across the age spectrum. Put simply, no child should die, but all adults must die, albeit sometimes at ages less than desirable. This is why the international health community tends to use a measure of lost time as the best indicator, which weights child deaths much more heavily. The difference can be substantial, as indicated in the accompanying figures from the report.

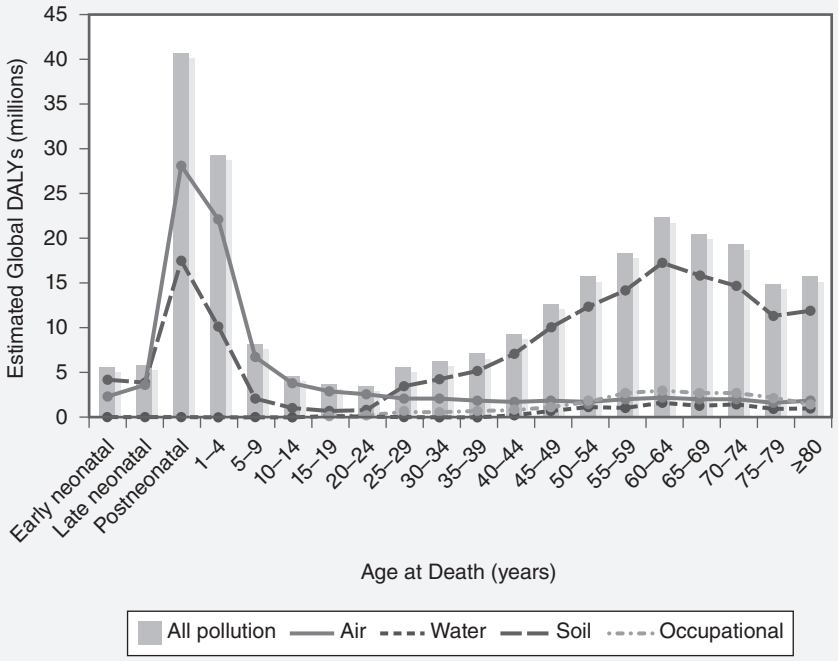


The distribution of lost lives from pollution exposures globally.

Data from Landrigan, P. J., Fuller, R., Acosta, N. J. R., Adeyi, O., Arnold, R., Basu, N., et al. (2017). The Lancet Commission on pollution and health. *The Lancet*.

(continues)

EXHIBIT 11-1 *The Lancet* Commission on Pollution and Health *(continued)*



The distribution of lost DALYS from pollution exposures globally.

Data from Landrigan, P. J., Fuller, R., Acosta, N. J. R., Adeyi, O., Arnold, R., Basu, N., et al. (2017). The Lancet Commission on pollution and health. *The Lancet*.

While the impacts look most important at older ages, when deaths alone are considered, they are distributed more widely across the age spectrum when considered as lost healthy life years (DALYs), as the impact on children dominates when this perspective is adopted. Arguably, the graph depicting the impact in this way is a better representation of the problem from a global health perspective.

Distribution of and Factors Influencing the Environmental Burden of Disease

As shown in Table 11-2, the largest environmental health burdens occur in LMICs with significant household-level risks and tend to affect young children in particular. Indeed, as a percentage of total burden as well as in absolute terms (e.g., DALYs per capita), environmental risks are most important in the poorest populations. Both the Millennium Development Goals (agreed to in 2000 by all UN member states) and now a second WHO environmental burden of disease assessment allow for longitudinal analysis of environmental hazards. These demonstrate well the environmental risk transition, with globally declining burdens of infectious, neonatal, and nutritional disease deaths attributable to environmental factors, and increasing deaths from noncommunicable diseases and injuries. In addition to between-country differences by income, low- and middle-income populations within all types of countries are the most vulnerable to the health impacts of environmental degradation and change. They are typically more exposed, in terms of residential and occupational location, and have fewer resources for taking protective or adaptive action.

Inequities between and within countries by income and ethnicity, relative to both environmental hazards that cause harm and positive aspects of environments that support good health, have led to the development of civic movements for environmental justice. What started as a 1980s civil rights movement in the (United States protesting the local-scale siting of toxic waste facilities in ethnic minority communities has since grown into global-scale research and activism. For example, notions of "climate justice" were central in the negotiations that led to the Paris Agreement on climate change (Agyeman, Schlosberg, Craven, & Matthews, 2016), and the term is recognized on the first page of that agreement (United Nations Framework Convention on Climate Change, 2015).

Changes in the age structure of populations also affect vulnerability. For instance, as average life expectancies increase, populations become more vulnerable to many environmental stressors because of the increasing proportion of elderly persons, although there is also somewhat less population vulnerability due to the

lower proportion of the very young. Many physiological and biochemical functions of the human body—such as kidney function, liver function, eyesight, and hearing—decline with age, even after early adulthood.

At an individual level, constitutional characteristics frequently influence susceptibility to environmental exposures. Readily seen examples include skin pigmentation modulating the risk of solar-induced skin cancer, and age modulating the efficiency of intestinal absorption of lead. Individual variations in metabolic phenotype, which are substantially determined by genotype, are also important. Many enzyme pathways are known to be involved in the activation or deactivation of potentially carcinogenic or other toxic chemicals, such as the various oxidizing enzymes of the mitochondrial P450 system, the acetylation pathway (which yields phenotypically "fast" and "slow" acetylators), the glutathione transferase pathway, and the alpha-1 antitrypsin pathway. A steadily accumulating body of epidemiologic evidence indicates that these polymorphisms modify the disease-inducing effects of an external environmental exposure, such that a gene–environment interaction is observable at the individual level. Discovery of such linkages opens up important opportunities for higher-resolution research. When study subjects are stratified on a metabolic polymorphic characteristic relevant to the external exposure, then the effect within the susceptible subgroup will become more evident than when the effect is diluted (averaged) across the susceptible and nonsusceptible subgroups.

Emerging Issues in Environmental Health: Moving from Hazard to Habitat

Most analyses of environmental health effects have focused on specific single, direct-acting hazards within a localized setting. As described in the introduction to this chapter, hazards can be investigated (1) based on which of these categories they fall into, (2) through their route of exposure (ingestion, inhalation, dermal), or (3) through the locations in which they occur (home, school, outdoors). Occupational or workplace hazards are often considered separately from the previously mentioned categories, partly because workplace settings are higher-risk environments, with both increased intensity and longer duration of exposure, and partly because working populations are usually different from the overall population, having fewer vulnerable groups such as the very young, old, or ill, and because they personally gain, at least conceptually, in the processes that lead to risk.[1]

As part of an environmental health assessment, exposures are assessed at either the individual level or the group level, as are health outcomes. From the resulting exposure and outcome data, exposure–response[2] relationships can be estimated, usually by fitting the data to statistical models. Where data are sparse, model fitting may be guided by theoretical considerations. Once exposure–response relationships have been satisfactorily estimated, and if the causal interpretation is convincing, the results can be used to guide environmental policy. There are some caveats, however: Exposures have often been higher in workplaces than in the ambient environment, and many of the published dose–response relationships are based on occupational epidemiologic studies. As such, they sometimes do not reflect the exposures experienced by the broader population.

During recent decades, this mode of environmental and occupational health research, which culminates in formal risk assessment (National Academy of Sciences, 2009; Samet, Schnatter, & Gibb, 1998), has prevailed in high-income countries. Using this system, ambient and workplace environmental exposure standards have been set for several hundred specific environmental exposure agents. There are currently more than 140,000 human-made chemical substances in commercial use worldwide, and 5,000 of these are in widespread use (Landrigan et al., 2017). Further, thousands of naturally occurring chemicals are in general use, including many in LMICs. Insufficient epidemiologic and toxicologic data are available to evaluate the potential health effects of most of these chemicals (Moochhala, Shahi, & Cote, 1997; National Academy of Sciences, 2009), and the risks tend to be evaluated for individual agents only, if at all. As such, the regulatory and monitoring paradigms for dealing with these chemicals may be insufficient to deal with the complex mixtures to which individuals are exposed, through multiple routes, on a regular basis (National Research Council, 2012).

Meanwhile, larger-scale environmental hazards to human health began to emerge in the final decades of the 20th century. All around the globe, a spectacular shift of human populations into the world's cities has been occurring. The urbanized proportion of the world's population has skyrocketed from approximately 5% to more than 50% in the past two centuries and is still rising (United Nations Habitat, 2016). This

[1] For instance, consider a factory worker using dangerous, job-specific machinery involving repetitive motion for many hours each day, 5 days a week for 30 or 40 years. Similarly, agricultural workers may perform repetitive tasks and be exposed to heat, lack of water, and pesticides for many hours each day for many years.

[2] These are sometimes called "dose–response" relationships, a term from toxicological studies in which a known dose of toxin is administered to animals under study. For most environmental hazards, however, exposure is the best metric that can be determined.

urban migration reflects, to varying degrees in particular geographic areas, the advent of industrialization, the contraction of rural employment, the flight from insecurity, and the search for jobs, amenities, or a stimulating environment. With such vast movement to urban areas comes new challenges, including overcrowding; exposures to both traditional and modern environmental health risks, which are amplified in urban slums; occupational and economic risks, as many jobs in cities are unregulated and sometimes exploitative; and environmental degradation, as many urban centers expand without adequate planning.

A major consequence of the increasing scale of the human enterprise is the potentially important health impact of global environmental changes. Humankind is now disrupting some of the biosphere's life-support systems—the natural processes of stabilization, production, cleansing, and recycling that our predecessors were able to take for granted in a less human-dominated world (McMichael, Woodruff, & Hales, 2006; Millennium Ecosystem Assessment, 2005). We no longer live in such a world. The scale and intensity of these human-initiated global environmental changes has led to the assertion that we are shifting out of the Holocene and now live in a new, human-dominated geological epoch, the Anthropocene (Lewis & Maslin, 2015).

In 2009, Johan Rockstrom and his colleagues introduced the concept of planetary boundaries, a framework at a global level for considering ecological levels of risk to human health, well-being, and society. They identified nine such boundaries and have been quantifying the current status of determining variables (Steffen et al., 2015), summarized in **FIGURE 11-1** and discussed in more detail in this section.

With increased emissions of greenhouse gases and various ozone-destroying gases, the basic composition of the world's lower and middle atmospheres is being altered by human actions. The world's great geochemical cycles of sulfur and nitrogen are also being significantly perturbed by human actions. Human-induced environmental changes are causing worldwide depletion of soil fertility, aquifers, ocean fisheries, and biologic diversity. The ongoing mass extinction might cause approximately one-third of all species alive in the 1800s to disappear before the end of the 21st century (Ceballos, Ehrlich, & Dirzo, 2017).

The loss of key species would weaken whole ecosystems, creating adverse consequences to humans, such as disturbing the ecology of vector-borne infections and food-producing systems that depend on pollinators and the predation of pests, and impairing the cleansing of water and the circulation of nutrients that normally pass through ecosystems. A rich repertoire of genetic and phenotypic material would also be lost. To maintain the hybrid vigor and environmental resilience of food species, a diversity of wild species needs to be preserved as a source of genetic diversity (Chivian & Bernstein, 2008). Similarly, many modern medicinal drugs in Western medicine have natural origins, and many defy synthesis in the laboratory. Indeed, scientists continue to test thousands of novel natural chemicals each year, seeking new drugs to treat human immunodeficiency virus (HIV), malaria, drug-resistant tuberculosis, and cancers. And, more fundamentally, the species with which we share our planet have profound connections with human well-being via complex relationships that extend back through evolution.

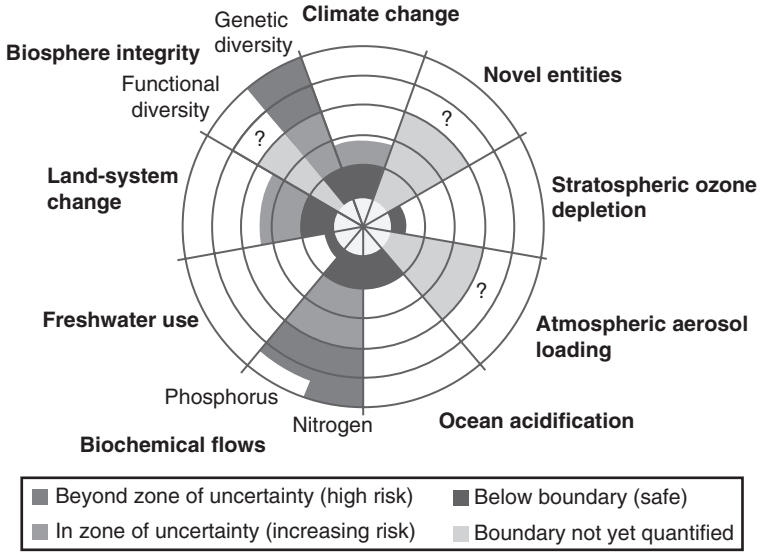


FIGURE 11-1 Nine planetary boundaries determining a "safe operating space" for human well-being, including the current status of control variables for seven boundaries.

Reproduced from Steffen, W., K. Richardson, J. Rockström, S. E. Cornell, I. Fetzer, E. M. Bennett, et al. (2015). "Sustainability. Planetary boundaries: guiding human development on a changing planet." Science (New York, N.Y.) 347(6223): 1259855. Reprinted with permission from AAAS.

The opposite side of this coin is the accelerating spread of invasive species, as long-distance trade, tourism, and migration increase in intensity. For example, the proliferation of water hyacinth (a decorative plant from Brazil) in Lake Victoria, eastern Africa, has extended the breeding grounds for the water snail that transmits schistosomiasis. The planting of *Lantana camerata* as a garden border shrub in Uganda, and its subsequent dispersed spread, has increased the habitat for the tsetse fly, which transmits the trypanosome that causes African sleeping sickness.

As a result of these changes to the ecosystem, serious consequences should be expected, such as changes in patterns of infectious diseases, in regional agricultural and aquatic yields, and in the effects of economic hardship or demographic displacement. Such changes are already beginning to change the conditions of the biosphere—our habitat—and pose long-term, and unfamiliar, risks to human population and ecosystem health.

Recognizing the intertwined nature of economic development, environmental sustainability, and human health, the Millennium Development Goals have recently been superseded by the Sustainable Development Goals, a set of 17 goals that inextricably link economic development, human health, and environmental sustainability (United Nations, 2015b). While all these Sustainable Development Goals have some relevance to modern environmental health, some are of obvious, direct concern to a habitat approach, including sustainable water management (Goal 6); access to affordable, clean energy (Goal 7); resilient, livable cities (Goal 11) with sustainable industrialization (Goal 9) and consumption and production patterns (Goal 12); and ecosystem health (Goals 13, 14, and 15).

Given these considerations, the scope of the environmental health framework must be extended to emphasize the impairment of our global habitat (McMichael & Bambrick, 2011) and its complex relationship with human health. Such a shift in focus should emphasize the positive roles that intact ecosystems and well-designed built environments play in supporting human health. While these "big picture" goals are incorporated into the environmental health framework, "traditional" environmental health priorities must evolve to incorporate modern realities. Hazard-oriented research should continue, but expand to incorporate a more thorough understanding of complex, interacting, cumulative exposures. New linkages need to be explored between urban planning and health, as large populations move into growing cities and face a combination of new and old environmental health threats.

Responses to the need for an expanded philosophical and methodological scope have burgeoned over the past three decades, giving birth to a number of important new approaches to environment and health. The One Health Commission (2017), which emerged from the medical and veterinary sciences (Zinsstag, Schelling, Waltner-Toews, & Tanner, 2011), acknowledges that issues such as zoonotic disease, food safety, and antimicrobial resistance require a greater understanding of human–animal–environment interactions, with a heavy emphasis placed on the human–animal nexus and on interdisciplinary collaborations among the medical, veterinary, and environmental sciences (see the *Public Health Infrastructure* chapter for further information on One Health). Ecosystem Health, or EcoHealth, has been developing in practice and research since the beginning of this century from roots in ecology and public and global health. Charron (2012) synthesizes the following set of EcoHealth research and practice principles, based on values of environmental sustainability and social equity: (1) systems thinking; (2) transdisciplinary research (involvement of policy makers and community representatives with scientists) and participation; and (3) translation of knowledge into policy action.

The evolving global approaches to rabies prevention and management provide a useful contrast between the One Health and EcoHealth methodologies. Since rabies is almost always fatal once contracted, prevention strategies are vital. Partnerships between medical and veterinary scientists in a One Health approach have led to the successful eradication of human rabies in many countries through the development of animal vaccines and mass-vaccination programs for dogs (WHO, 2013c). Further, One Health partnerships between medicine and veterinary science have been used to understand the transmission of rabies from wild carnivores and bats to humans (WHO, 2013c). However, in some parts of the world (such as India and parts of Africa and South Asia), attempts at eliminating rabies through mass dog vaccinations alone have proved unsuccessful. One of these places is Bali, which experienced a resurgence of rabies cases from 2008 despite having been previously rabies-free. Initial centralized mass dog vaccination was unsuccessful at achieving high rates of vaccination (Putra et al., 2013), which led to a project based on EcoHealth principles (Jatikusumah et al., 2013). This program involved wider partnerships among health, veterinary science, local government, livestock management, and communities. A mixture of methods were used to understand free-roaming dog ecology, demographics, and behavior; social and cultural relationships between humans and dogs; and local leadership and governance processes. This information was then used to create a much more context-specific, local, and community-led response to education, dog vaccination and management, dog registration, and rapid response to dog bites.

More recently, the planetary boundaries concept has stimulated renewed discussions in global health about Earth systems, social systems, and health

(Horton & Lo, 2015), with *The Lancet* introducing a new journal on the subject titled *Planetary Health*. This new concept is focused on "the interplay between health and the determinants of health in our living world" (http://www.thelancet.com/lanplh/about) and appears to be bringing ecosystem health concepts into the mainstream global health discourse.

Methods for Environmental Health Research and Practice

Research Scope and Strategies

Environmental health research seeks to understand how the environment influences states of health, to help prioritize and develop appropriate interventions to reduce risks to health, and to investigate the effectiveness of such interventions. Methods derived from a wide range of scientific disciplines are applied to these endeavors. Toxicology and microbiomics, audiology and neuropsychology, ecology and systems science, and others make important contributions.

We focus first on the methods and strategies applied in epidemiology, which has traditionally been the basic quantitative science of environmental health research. On their own, traditional research approaches are insufficient to engage with environmental problems of large scale and high complexity, especially as the field evolves from studies of discrete hazards to investigations of disordered systems. We outline some of the research tools that are being applied to studies of modern environmental health problems such as climate change and loss of biodiversity.

In essence, epidemiologic research describes and explains variations and temporal changes in the patterns of illness and disease between and within populations. Most environmental epidemiology is observational (that is, nonexperimental), which introduces some important issues in research design and data interpretation (Baker & Nieuwenhuijsen, 2008; Morgenstern & Thomas, 1993). However, where health benefits are anticipated from exposure-reducing interventions, experimental studies may be carried out.

Historically, epidemiology played an important role in identifying specific, localized hazards such as severe air pollution (e.g., the "London Fog" of 1952), waterborne infections, and asbestos in workplaces. While problems of this kind persist, we now face more subtle exposures, which are less well defined and more diffuse, such as electromagnetic fields and chemical exposures that mimic hormones and act cumulatively over decades (Diamanti-Kandarakis et al., 2009).

Many environmental exposures occur at levels that are low by comparison with what was experienced in the past in highly polluted settings. For example, in terms of fine particulates and noxious gases, exposures in many cities in high-income countries are commonly two orders of magnitude less than the "London Fog." This makes it more difficult to detect increments in risk that are, in absolute terms, modest. Even so, the effects are important, for the following reasons:

- They typically impinge on many persons, perhaps whole populations, thereby causing a large aggregate health impact (an economic–political criterion).
- They are encountered on an essentially involuntary, and often unequal, basis (an ethical criterion).
- They are often amenable to control at the source (a practical criterion).

The environmental health researcher faces two other recurrent difficulties. First, these real-world exposures are likely to be accompanied (and potentially confounded) by other exposures or risk factors—some of which may be unknown to the investigator or, indeed, to science. Second, the exposure–effect relationships often involve long-term, chronic, and sometimes subtle causal processes.

Because of these complexities, environmental health research must often be tackled in a multidisciplinary fashion so as to develop a sufficiently broad basis of evidence from which to make causal inferences. For example, causal inferences about the effect of low-level environmental lead exposure on the cognitive development of young children required the integrated consideration of the results of epidemiologic studies, animal experimental research, and neuropathological and molecular toxicological studies.

Extra leverage may be gained via interdisciplinary research in which the techniques of several disciplines are combined. For example, the development of molecular biology over the past several decades has yielded many new techniques for measuring internal exposure, especially in relation to carcinogenesis. Despite these challenges, complexities, and advances, there remains an important role for small-scale local studies that identify useful preventive policy recommendations.

Causation and Other Methodological Issues

For a detailed account of epidemiological reasoning, concepts, and methods, see a textbook such as *Modern Epidemiology* (Rothman, 2012). In this section, we cover only the most important elements relevant to environmental health.

Over recent decades, epidemiologists have developed a set of criteria specifically suited to their predominantly nonexperimental, bias-prone, confounding-rich research, with particular emphasis placed on the temporality of the relationship, its strength, the presence of

a plausible dose–response relationship, the consistency of findings in diverse studies, and coherence with other disciplinary findings and biomedical theory (Bonita, Beaglehole, & Kjellstrom, 2007). Nevertheless, etiologic research in environmental epidemiology entails several distinctive methodological issues:

- Choice of study design for data collection and analysis.
- Definition of exposure and choice of the mode of exposure assessment.
  - Choice of the relevant reference exposure (the theoretical minimum exposure level that a society could achieve).
  - Approach to dealing with multiple coexistent, potentially interacting environmental exposures.
  - Choice of the appropriate level of comparison (population, local community, or individual). Many environmental exposures (e.g., ambient air pollution or fluoride levels in drinking water) impinge on whole communities, with minimal exposure differences between individuals.

Study Design Options

The same basic set of study designs that are used in general epidemiology are also used in environmental epidemiology—that is, descriptive, analytic, and experimental (Baker & Nieuwenhuijsen, 2008; Bonita et al., 2007).

Descriptive Studies: Trends and Correlations

In descriptive studies, the pattern of variation in a population's or community's environmental exposure or health status (or both) is described, usually in relation to time, place, or category of person (e.g., age, sex, ethnicity). If the data are appropriate, the relationship between exposure factors and health status may be described. Such studies aid in identifying research priorities and in guiding the design of etiologic studies. For example, a study may show that the exposure levels are not high enough to warrant closer investigation. Time trends may be informative (e.g., seasonal variations in mortality that may or may not be related to temperature). Spatial distributions also provide clues to possible causes and sources of vulnerability.

Descriptive studies may struggle to reach definite conclusions about etiology. Commonly, this difficulty arises because inadequate information is available for confounding factors, or the temporal relationship between exposure and health outcomes is unclear.

Time-series studies have a special role in environmental epidemiology. Some environmental exposures such as urban air pollution and weather conditions vary on a short-term basis. Intrinsically interesting questions may arise, such as whether asthma attacks increase on high-pollution days or whether daily death rates increase on days of extreme temperature. The statistical techniques used within time-series analyses have been acquired largely from econometrics and engineering research; they include tools to adjust for lower-frequency (e.g., seasonal) cyclical variations, background secular trends, and autocorrelation. Time-series studies benefit from the fact that ongoing characteristics of the study population—such as age distribution, socioeconomic profile, and smoking habits—remain essentially constant over time. Further, because the comparison is made entirely within the chosen population, between population confounding factors do not apply.

Analytic Studies

Analytic studies examine formal statistical associations between an exposure variable and a health outcome variable at the level of the individual or a small homogeneous exposure group (Bonita et al., 2007). Are individuals with higher exposure to indoor air pollution more likely to develop respiratory disease than those with low exposure to such pollution? Are individuals who develop diarrheal disease in a coastal city more likely to have been swimming in contaminated seawater recently than individuals without such disease? Studies that seek to answer questions such as these can be designed to start either from exposures (cohort studies, such as the indoor air pollution example described earlier) or from the health effect (case-control studies, such as the diarrheal disease example described earlier).

The study of the relationship of early-life environmental lead exposure to child cognitive–intellectual development, for example, began with various types of cross-sectional studies in the 1970s. However, it was not possible to establish from those studies the temporal relationship between occurrence of exposure and occurrence of intellectual or behavioral deficit. To determine this relationship, cohort studies were required, in which infants and children were followed from birth, with systematic documentation of their early-life lead exposure history and cognitive–intellectual development. An investigation of this kind was carried out in and around the lead smelting town of Port Pirie, South Australia (Tong, Baghurst, Sawyer, Burns, & McMichael, 1998), and provided the data necessary to estimate by how much a given exposure to lead impaired neurologic function among young children (**EXHIBIT 11-2**).

Experimental and Intervention Studies

Experimental and intervention studies begin with sets of reasonably similar populations or groups, which can then be allocated, preferably randomly, to "intervention" or "control" categories. The statistical analysis

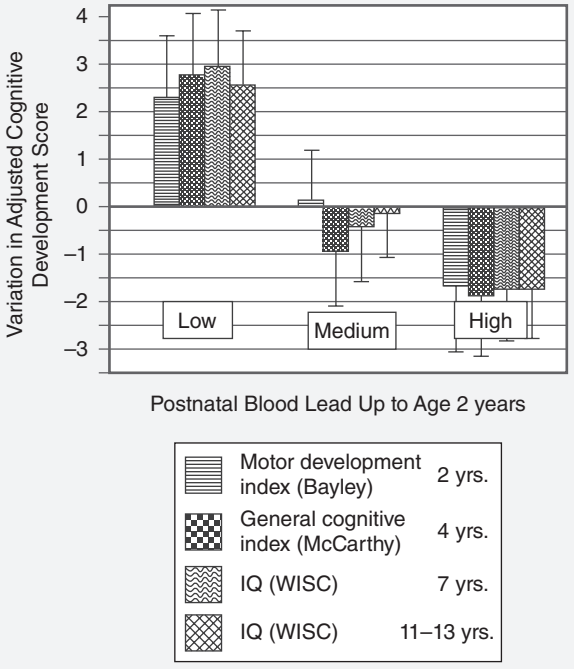
EXHIBIT 11-2 Environmental Lead Exposure and Childhood Cognitive Development

Lead is the most abundant heavy metal and may have been the first metal smelted, around 6500 B.C.E. The ancient civilizations of Phoenicia, Egypt, Greece, Rome, China, and India used lead for vessels, roofs, water ducts, utensils, ornaments, and weights. The use of this metal increased greatly during the Industrial Revolution, given the many uses of lead in manufacturing processes. The rise of the automobile meant that large of quantities of lead were required both in lead-acid batteries and as an antiknock additive in gasoline. This history is evident in the environmental record: The lead content of Greenland ice, for instance, reached 100 times the natural (pre-industrial) background level in the mid-1990s.

Lead affects many organ systems—most importantly, the gastrointestinal tract, central nervous system, kidneys, and blood (through the impairment of hemoglobin synthesis). Very high levels of exposure (leading to blood lead concentrations of 50 to 100 mg/dL) may occur in the workplace, particularly in lead smelting, battery recycling, lead soldering, and various lead-based craft activities. Such exposures are uncommon now in high-income countries, but are widespread in many LMICs.

Epidemiologic evidence shows that exposure to lead in early childhood causes a deficit in neurocognitive development, as illustrated in the nearby figure (Tong et al., 1998). Evidence from animal experimental studies and neuropathologic analyses backs up the epidemiology. A doubling in blood lead concentration from 10 to 20 mg/dL—roughly equivalent to the difference between high and low tertiles in poorly controlled urban environments—is associated with a deficit in intelligence quotient of 1 to 3 points, or 1% to 3% of the expected average IQ score of 100 (Pocock, Smith, & Baghurst, 1994). Subtle effects on IQ may result from blood lead levels as low as 5 mg/dL. Note that a loss of IQ points will have greater impact in children on the threshold of significant neuro-psychological impairment (e.g., mild mental retardation, defined as an IQ score of 50 to 69) than in children with a higher IQ.

Many high-income countries, including the United States and Australia, have set relatively stringent standards for environmental lead levels, in an effort to protect young children from exposure to this metal. Even so, lead may still present a serious environmental health problem, including in wealthy societies. For instance, widespread contamination of the drinking water of Flint, Michigan, occurred in the United States in 2016, caused by a shift to a more corrosive water source in a city with an old reticulation system that included lead piping (Hanna-Attisha, LaChance, Sadler, & Champney Schnepp, 2016). Likewise, childhood lead poisoning occurs in many LMICs. For example, the lead content of gasoline sold in Africa is the highest in the world and is associated with high lead concentrations in the atmosphere, dust, and soils. More than 90% of the children in Cape Province, South Africa, in the 1990s had blood lead levels in excess of 10 mg/dL (Nriagu, Blankson, & Ocran, 1996). In 2017, it was estimated that lead causes approximately 14 million lost DALYs and 500,000 premature deaths each year worldwide (Landrigan et al., 2017).



Lead exposure and cognitive development in childhood. One point approximates to a 1% change. Low, medium, and high lead exposure categories have mean blood lead concentrations of 12.5, 18.5, and 25.9 mg/dL, respectively.

Modified from Tong, S. et al. (1996). Lifetime exposure to environmental lead and children's intelligence at 11–13 years: The Pt. Pirie cohort study. *British Medical Journal*, 312, pp. 1569–1570.

Some difficulties arise in including cognitive and intellectual impairment in the risk assessment equation for evaluating and controlling environmental lead as a public health hazard. This particular functional health deficit—a subtle neurologic impairment that will have its most marked effects on the social and psychological development of children who already have a low IQ score—does not readily translate into the standard currency of deficit due to disease, disability, or death (the DALY). Hence, the burden of disease caused by relatively low but widespread exposures to lead may be underestimated by policy makers. Although overt lead-induced toxicity is apparent in individuals experiencing high levels of exposure, the full public health impact of widespread exposure to a range of environmental lead levels requires evaluation of more subtle health, behavioral, and developmental effects (Moore, 2003).

In developing a policy on an environmental health risk such as lead exposure and taking the evidence, such as that on cognitive impairment, into account, two questions commonly arise. First, does the deficit persist over time? It appears that the effect of lead on the developing brain does, indeed, persist through late childhood and into early adulthood. Second, is there a threshold—that is, an exposure level below which no neurotoxicologic effect occurs? It is difficult to answer this question because few data are available in the very low exposure range, so conclusions about "safe exposures" must be framed cautiously. Nevertheless, a strong case can be made for adopting public health measures to prevent lead exposure in early childhood. Because lead exposure tends to be widespread within a population, even a modest health impact upon each individual could yield a substantial aggregate impact for the total population. For instance, through its effect on blood pressure, it is estimated that lead is responsible for 2.5% of the global burden of ischemic heart disease (Landrigan et al., 2017).

compares outcome rates in the two or more groups (the intervention may be applied at more than one level).

The clinical randomized controlled trial (RCT) is commonly regarded as the "gold standard" for health research, but is seldom applicable to the study of environmental diseases among populations. It was not possible, for instance, to conduct an RCT of the effects of removing lead from gasoline, and indeed such a study was not necessary, given the weight of observational evidence justifying such a policy, and the rapid fall in body lead levels that occurred following its implementation. That said, there are sometimes opportunities for studies with an element of random allocation—for example, testing whether the installation of household humidifiers reduces the prevalence of respiratory symptoms, or whether the provision of masks to reduce inhalation of fine particles prevents acute changes in the cardiovascular system. In these cases, interventions are randomly allocated at the individual or household level. In some circumstances, randomized allocation of interventions at the population level is possible. A well-known historical example from the 1940s was the addition of fluoride to the drinking water of four towns in North America for the purpose of preventing dental caries, but not in four other similar towns (Fawell et al., 2006). However, the effectiveness of random allocation as a device for reducing the influence of confounding is much reduced when the number of observational units is small (8 in the case of the study of fluoride).

The need for a broader range of study designs to explore the effects of interventions becomes more apparent as the focus of environmental health is scaled upward. In the 21st century, studies of discrete toxins and physical exposures must be accompanied by investigations into the health effects of sociotechnical and physical system change. These include social constructs (e.g., the ingredients of a successful city, a well-functioning transport system) as well as elements in the natural world (e.g., the global climate, ecosystems that modulate risk of infectious diseases) and systems which mix the two (e.g., food systems).

Another type of intervention study often utilized in environmental health research is the "before and after" study. With this design, assessments are made before and after the introduction of an intervention. Before and after studies of interventions may struggle to control for time-varying confounders (i.e., factors that may differ before and after deployment of the intervention). However, research that includes multiple sampling points and well-chosen external reference groups can provide high-quality evidence. An example is the investigation undertaken to determine the effect of smoke-free legislation in Scotland on hospital admissions (Pell et al., 2008).

Exposure and Dose: Assessment and Definitions

Paracelsus, the 16th-century German physician/alchemist who is often credited with being the founder of environmental health and toxicology, is credited with writing: "Poison is in everything, and nothing is without poison. The dosage makes it either poison or remedy" (Deichmann, Henschler, Holmstedt, & Keil, 1986). This statement lies at the heart of environmental health science.

Measurements of toxic agents at the source or in emissions or in the ambient environment may be a useful guide, but do not directly indicate how much of the toxin is absorbed and so can cause harm. There is a huge amount of toxic mercury in the oceans, for example, but very little of it reaches people in normal circumstances. Volcanoes emit vast amounts of toxic gases, but fortunately few persons are usually nearby to breathe them. Dose is the amount of toxin that has actually reached

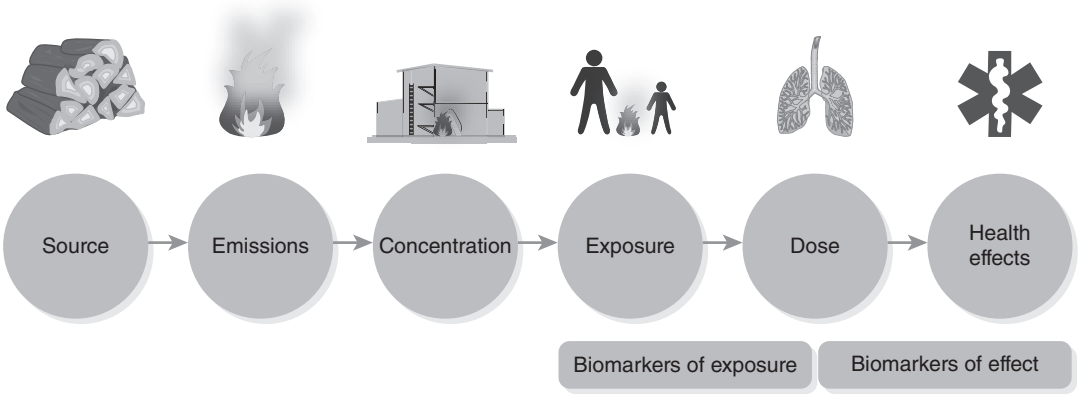


FIGURE 11-2 Environmental pathway.

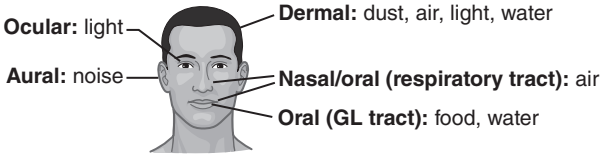


FIGURE 11-3 Routes of exposure.

the vulnerable parts of the body. Unfortunately, it is difficult to measure dose directly for most hazards (whether chemical or physical), either because it involves sophisticated, expensive, and invasive procedures (e.g., extracting and analyzing blood or tissue samples) or because it is beyond current scientific abilities.

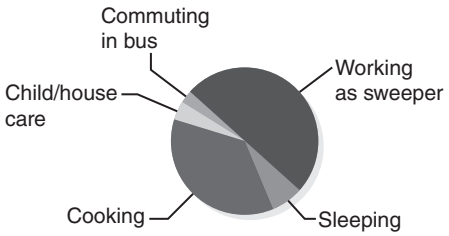
For this reason, scientific and policy attention is mostly focused on the intermediate segment of the environmental pathway—that labeled "exposure" in **FIGURE 11-2**. Exposure refers to the quantity and duration of the hazardous factor that humans encounter in the course of their activities. More explicitly, it is the amount of material or energy in the air, water, food, and soil that reaches the body's protective barriers of the respiratory and digestive systems, skin, eyes, and ears (**FIGURE 11-3**).

Total Exposure Assessment

To understand the full impact of a pollutant, it is necessary to examine all the ways it might reach people, rather than simply relying on measurements made in the most convenient places. This consideration is especially important for pollutants that can reach people through several different routes. For example, lead pollution may affect vulnerable groups, such as children, through airborne dusts or contaminated water, soil, and food. Confining attention to a single route may underestimate the true size of the problem (see Exhibit 11-2): It is total exposure, of course, that determines the risk to health.

Consider, for example, the woman with the daily pattern of activities shown in **FIGURE 11-4**. What is her health risk due to exposure to particulate air pollution? She lives in an urban slum of a low-income country where outdoor air pollution levels are fairly high. Her total exposure is even higher, however, because she spends considerable time in locations where particulate concentrations are even higher than the outdoor levels. During the working day, she works as a sweeper on busy streets, where pollution is especially severe due to large volumes of nearby traffic. In the morning and evening, she is exposed indoors to toxic products of combustion from an unvented cook stove, made worse because her family can afford only low-quality cooking fuels such as briquettes made from coal dust. In the evening, she is also exposed to secondhand smoke from her husband's cigarettes. Her total exposure over the day is best estimated by the sum of the pollutant concentration in each major microenvironment, weighted by the fraction of time she spends in it.

Total exposure assessment is necessary to accurately record the burden of disease attributable to environmental factors. Such information may change the relative importance of sources of pollution, or uncover important new sources of personal risk, such as secondhand smoke, that do not appreciably affect concentrations in the ambient environment. This metric also reveals a new dimension of potential control measures. For example, chimneys for household stoves that would not change emissions at all—and might even increase outdoor concentrations—can lower exposures substantially by separating the people from the pollution. Laws to reduce smoking in public places can lower exposure of the general population to secondhand smoke, even if there are no changes in the amount of smoking overall. In some cases, the cost-effectiveness of such exposure-control measures can be much higher than the exposure reduction achieved through generalized control of outdoor sources. Although ideally all pollution sources should be controlled, in reality there is a limit to what can be achieved in the short term.



Activity	Hours/day	Particulate concentration ($\mu g/m^3$)	Daily exposure equivalent ($\mu g/m^3$)	Proportion of total exposure
Sleeping	7	100	29.2	0.07
Cooking	3	1,200	150.0	0.36
Child/house care	3	120	15.0	0.04
Commuting in bus	1	300	12.5	0.03
Working as sweeper	10	500	208.3	0.50
Total	24	Mean = 415	415	1

FIGURE 11-4 Total particulate exposure for a woman in an urban slum.

Biological Markers of Dose

Environmental health researchers may use many different biological markers of dose: heavy metals in hair, nails, and blood; metabolites in urine of chemicals such as nicotine; chlorinated organic chemicals in adipose tissue; radionuclides in bone; core body temperature after excessive heat exposure; or antibody titers in relation to infectious agents. For instance, fingernails and hair have been used to study exposure to organophosphate compounds and polybrominated ethers (Liu, He, Hites, & Salamova, 2016).

Biological assays integrate the physiological and metabolic processes that influence absorption and clearance. These processes vary between individuals, and even in the same person, there may be fluctuations over time. Biomarkers are attractive ways to assess dose because they provide precise measures and imply plausible pathways for pathogenesis, but their relevance to disease causation and policy requires careful thought. Precise measures of irrelevant parameters are less helpful than imprecise data collected on variables that are directly applicable to research and practice. Useful monitoring and regulation of ambient air pollution, for instance, is more likely to be based on measures of exposure than on bioassays.

Consider the field of molecular epidemiology, which became prominent during the 1990s as an approach to studying the cause—especially the environmental cause—of cancer (McMichael, 1994). This same field has also become important in modern infectious disease epidemiology, particularly for the determination of environmental sources and transmission pathways for infections such as Legionnaires' disease, tuberculosis, influenza, cholera, and food-poisoning organisms.

Molecular biomarkers make use of the variations in the structures of macromolecules, particularly DNA. One area in which DNA measurement has assisted the conventional epidemiologic study of causation has been in investigations of dietary aflatoxin and liver cancer (McMichael, 1994). Aflatoxin is a biotoxin produced by the *Aspergillus flavus* mold in foods stored in warm, humid environments. Initially it was not possible to measure directly an individual's level of aflatoxin intake. Epidemiologists were limited to demonstrating ecological correlations, in eastern Africa and within China, between average aflatoxin concentrations in local diets and rates of liver cancer mortality. Now it is possible to measure the concentration of excised, excreted, aflatoxin-DNA adducts in urine and to use this as a measure of recent individual exposure. Such an approach was taken in a cohort study of 18,000 Chinese men in Shanghai that found a positive association, at the individual level, between adduct level at the commencement of the study and subsequent cancer occurrence (Qian et al., 1994).

Exposure Assessment at Individual and Population Levels

Exposure assessment at the individual level can occur at varying levels of intensity, ranging from exposure inferred via questionnaire, to personal exposure assessment (in which an individual's contact with environmental hazards is quantified, often using measurement devices), to measurement of individual biomarkers in blood, urine, hair, or tissue. Each type of assessment has its advantages and disadvantages.

For example, in studies of the health effects of secondhand smoke, it has been common practice to classify

exposures according to whether an adult has a smoking spouse or a child has a smoking parent. Clearly, someone who lives with a smoker is more likely to be exposed to tobacco smoke than someone who does not, but this is not necessarily so in every case. A child with no smoking parent, for example, may have four smoking grandparents who visit every day. Another child may have smoking parents who are careful to refrain from smoking anywhere near their child. In such cases, classifying children based on the smoking status of their parents would lead to exposure misclassification and, consequently, an attenuated estimate of the true effect. This type of error may be avoided by careful questionnaire design and by use of biomarkers such as cotinine in urine or saliva for at least a sample of the study population.

Often studies are based on a measure of exposure of whole population. This is obviously appropriate if the exposure is truly ecological—that is, it pertains to the whole population rather than individuals. Examples might be the level of income inequality or uptake of immunizations. But if the population measure is a proxy for what is happening at the level of individuals, there is a risk of misclassification, leading to bias in the measure of effect (Blakely & Woodward, 2000). It is common, for example, to conduct air pollution epidemiology studies by dividing urban populations into exposure classes according to the measurements made at the nearest outdoor air pollution monitoring station or even to use one or a few monitors to represent the exposure of an entire city for comparison with other cities. In reality, local exposure sources can contribute to great variation between individual exposures and the estimated population average exposure. The same problem arises in studies of heat, which commonly rely on measures of air temperature taken at a small number of weather stations, as there are many factors in the micro-environment that affect local temperatures and, in turn, an individual's experience of heat (Kjellstrom, Holmer, & Lemke, 2009).

Health Outcome Assessment

Many studies of the health impacts of environmental factors such as fine particles or heat have examined associations with mortality. However, the underlying pyramid of nonfatal health effects has a broad base: A full understanding of the effects of environments on health requires information on hospitalizations, primary care consultations, long-term chronic conditions, impaired organ function, and self-assessed symptoms, to name just a few items.

Researchers and data-collection agencies naturally tend to prefer "hard" endpoints that are well defined clinically and amenable to clear-cut counting or measurement. Yet community surveys indicate that the most commonly experienced effects of environmental exposures are frequently social, behavioral, and psychological disruptions that are not well suited to precise measurement. Examples are the mental stress of noisy environments, symptoms elicited by unpleasant odors, and the constellation of health-damaging exposures that affect children in car-dependent cities (Kjellstrom, Ferguson, & Taylor, 2009).

Environmental Health Indicators and Monitoring and Their Implications for Policy

Environmental health issues are recognized in the UN Sustainable Development Goals as an important part of the broader picture of human welfare, sustainability, and social and economic development (Griggs, 2013). To achieve a lasting improvement in global health, an intervention must influence the factors that created the exposure in the first place.

Most high-income countries have established a range of environmental performance indicators that attempt to be both scientifically robust and readily applicable to policy and practice. For example, the Organisation of Economic Co-operation and Development (OECD, 1993) developed the pressure–state–response (PSR) model to meet these criteria. Other frameworks have been constructed that elaborate the relationships between human activities, environmental change, and human health (Corvalán, Briggs, & Kjellstrom, 1995). It is possible to distinguish "driving forces" that lead to pressures on environments and population health, affecting the state of the environment and resulting in human exposures, and ultimately leading to health effects and actions or interventions in response. This is known as the DPSEEA model, where the acronym stands for the hierarchy of Driving forces, Pressures, State of environment, Exposures, Effects, and Actions (**FIGURE 11-5**).

Driving force indicators are likely to be qualitative and are often expressed as yes/no answers. For example:

- Is there a policy to redirect all storm water to treatment plants?
- Are sedimentation dams in operation upstream in potentially cadmium-contaminated rivers?
- Are safety regulations for nuclear power stations adhered to?

Pressure indicators are usually quantitative. For example:

- The amount of sewage-contaminated storm water entering a beach or river after heavy rain

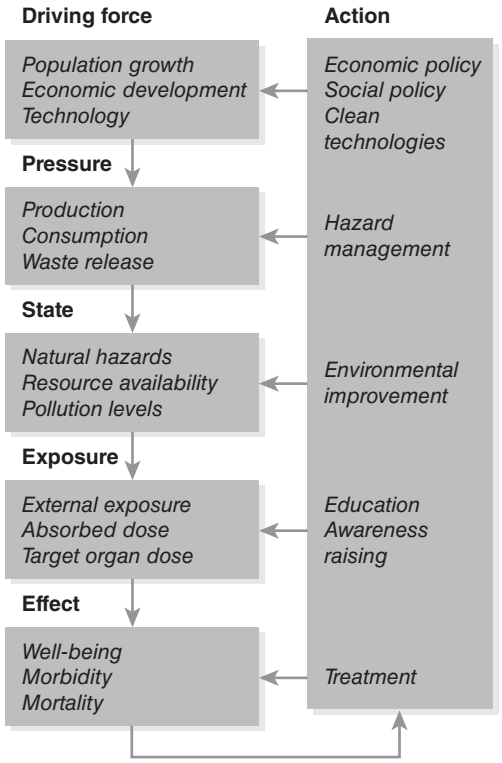


FIGURE 11-5 The DPSEEA framework to define environmental health indicators.

Reprinted from Kjellstrom, T., & Corvalán, C. (1995). Framework for the development of environmental health indicators. *World Health Statistics Quarterly, 48*, 144–154; Corvalán, C., Briggs, D., & Kjellstrom, T. (1996). Development of environmental health indicators. In D. Briggs, C. Corvalán, & M. Nurminen (Eds.), *Linkage methods for environment and health analysis*, (pp. 125–153). Geneva, Switzerland: World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/10665/62988/1/WHO_EHG_95.26_eng.pdf

- The amount of cadmium transported via river water to paddy fields
- The quantity of radionuclides released from a nuclear power station accident

The most common indicators are direct measurements of the environmental state—the concentration of a hazard in some environmental medium. For example:

- The enterococci concentration in beach water or drinking water
- The concentration of cadmium in rice paddy soil or rice
- The level of radioactive strontium in lichens or reindeer meat

Exposure indicators may be based on exposures calculated from state indicators, as described in the "Exposure and Dose: Assessment and Definitions" section earlier in this chapter. Biological indices add individual-based information and can be used to monitor both exposure (e.g., blood lead concentration, DNA-adduct level) and effects (e.g., enzyme assays for liver function, blood pressure). For example, such indicators were used to identify the marked decline in human milk concentrations of dichlorodiphenyltrichloroethane (DDT) (Smith, 1999).

Assessment of Environmental Health Impacts and Risks

Risk assessment is conventionally viewed as a stepwise sequence (WHO, 2002). It begins with the research-based identification of an environmental hazard. Subsequent studies then estimate, first, the exposure–response or dose–response relationship between the hazard and the specified health outcome, and, second, the distribution of exposure (doses) within the population of interest. Based on these two sets of data, the overall risk to the population is characterized, and risk management strategies are formulated.

Although there is now some agreement about how to calculate risks, there is no completely objective way to compare alternatives with different patterns of risk. This situation is illustrated in **EXHIBIT 11-3**, which presents a choice between two ways of producing electric power with different patterns of risk and explains how consumers with different worldviews could rationally choose either one or the other. This difference in risk perception explains why some people buy life insurance (risk aversion) and others do not, and why insurance companies run profitable businesses (based on expected-value calculations).

Many other considerations would be included in a full assessment of alternative power plants, including various outcomes of social importance other than those directly related to health. Thus, Exhibit 11-3 illustrates that a decision about how much or which kind of risk to take demands substantial scientific input but is ultimately a social and political choice. The extraordinary earthquake and tsunami damage to nuclear power stations in Japan in 2011, and the long-term challenges in rehabilitating and resettling Fukushima in its wake, illustrate just how difficult these decisions can be (Normile, 2011).

Probabilistic Risk Assessment

Because nuclear power plant operation is a relatively new enterprise, few accident statistics are available, particularly for large accidents in modern plants. How, then, can overall risk be determined? For this purpose, a technique called probabilistic risk assessment (PRA) is applied to understand the potential for adverse consequences. Developed originally by the National Aeronautics and Space Administration (NASA) for assessing the risks of manned space flight,

EXHIBIT 11-3 Risk Assessments: Additional Information, But No Substitute for Decision Making

Suppose a proposal has been made to build a large new power plant near your community. Everyone agrees on the need for more power and two options have been presented: coal and nuclear. These options are found to cost about the same. A full probabilistic risk assessment is done to compare the health implications of the two plants, yielding the results described here.

The coal plant, although of the best design, will have impacts in the form of air pollution and mining accidents. There is relatively little uncertainty about these factors—the probabilistic risk assessment finds a 90% chance that 30 persons will die prematurely because of the operation of the plant over its lifetime. Because there might be a large accident at the coal mine or a fire at the coal plant during meteorological conditions that lead to severe local air pollution, there is a 10% likelihood of killing 60 persons.

The nuclear plant is given a 99% chance of doing little damage (i.e., shortening only one person's life from the small amount of radiation released routinely). Unlike the coal plant, however, there is a small chance (here set at 1%) of a terrible accident in which 2,500 persons die from the radiation released.

Which plant is safer? Which would you rather live alongside? There are at least three ways to answer this question—none necessarily right or wrong, but all dependent on the set of values of the people making the decision:

- Maxi-min (maximize the chance of the minimum consequence): Because 99% of the time little damage is done and more may be learned in the future on how to reduce the chance of accidents, choose the nuclear plant.
- Mini-max (minimize the chance of the maximum consequence): Because a nuclear accident would be a tragedy (cause international headlines, go down in history, destroy the community, and so on), coal should be used so as not to have any chance of such an event.
- Expected value (calculate the odds): Because the expected number of deaths from the coal plant is 33 ($0.9 \times 30 + 0.1 \times 60$) and that from the nuclear plant is 26 ($0.99 \times 1 + 0.01 \times 2{,}500$), choose the nuclear plant.

Although the numbers used here are fictional and real probabilistic risk assessments for such facilities are much more complicated (involving hundreds of branches on the risk trees), the overall results are often similar. Nuclear plants generally have a lower expected value of damage but carry a small probability of terrible events (much less than 1%). Coal plants produce more damage on average but do not impose anxiety about large negative events (although growing concerns about possible climate change owing to the release of carbon dioxide from coal plants may change this perception).

The fact that groups of perfectly rational people may apply different decision rules does not mean that any group is right in an absolute sense, but rather indicates that groups start with different values. Indeed, although many other factors must be considered, the big difference in public acceptance of nuclear power between, say, France and Sweden can be partially accounted for by such differences in values.

this technique basically breaks down into subsystems the extremely complicated systems of large-scale technologies for which there is no overall accident information. These subsystems may then need to be broken down even further into even smaller components, until a level is reached for which failure data are available. The failure data for the various subsystems are combined to predict the performance of the total system.

For example, there is no information available that would enable one to predict directly how a new type of nuclear plant will operate over time, but such a facility is made up of thousands of components for which information is available. For example, information will be available about how often pumps of a certain size, switches of a certain voltage, or warning gauges of a particular brand are expected to fail under various conditions. Even if a new type of component is introduced, its failure rate can be experimentally determined without putting anyone at risk, unlike testing the entire power plant.

Such PRAs, of necessity, are extremely complex and difficult for even specialists to evaluate. In particular, it is difficult to tell whether all possible accident scenarios have been taken into account and whether unforeseen events, such as human sabotage, might potentially circumvent many subsystems at once. A PRA also assumes that the context in which each component was originally tested is not materially different in critical respects (e.g., temperature or vibration) from the working context of the assembled plant.

Where an assessment reveals a potential or actual exposure to a chemical for which there is no human exposure history, a similar dilemma results. It would be dangerous and unethical (not to mention time-consuming) to deliberately expose enough people to the chemical to discover its true risk. Instead, various techniques have been developed that can be used to estimate the risks of human exposures in this situation. The most commonly used method is to expose laboratory animals—usually rats and

mice—to high-enough doses of the chemical to observe effects within their relatively short lifespans. When deleterious effects, such as tumors or reproductive failure, are observed, then legitimate concern arises that such effects might also be seen in humans. To quantify that risk, however, requires extrapolating from high doses in animals to (usually) low doses in humans. To do so, researchers use a mathematical model that, ideally, is based on knowledge of metabolism, tumor induction, and other often poorly understood biological processes. Frequently, alternative animal models that are equally plausible biologically will predict different human risks for the same dose of the same chemical. Had the antinausea drug thalidomide been tested in rats, not rabbits, the tragedy of limbless human babies would probably have been averted. It is necessary, therefore, to establish standard conventions for which model will be used in which circumstances so that consistent risk estimates are made. At present, many of the model choices depend more on scientific intuition than on actual demonstrated knowledge. Finding more reliable and scientifically valid ways of performing such assessments is an active research area.

It might seem best to establish the convention of always using the most conservative model—that is, the model that predicts the largest risk. This strategy would seem to fit the classic public health dictum that, in cases of uncertainty, it is better to err on the side of caution than to underestimate potential risk. Indeed, for this reason, such a conservative approach is used in formulating official policy by some regulatory agencies. Unfortunately, and perhaps counterintuitively, if this approach is taken for each chemical independently, it can lead to the opposite effect—that is, exposing the public to unnecessary risks. The degree of conservatism (or the size of the safety factor) for specific chemicals can be quite different, depending on how well their observed effects on animals or humans actually compare with the predictions of the mathematical model employed. As a consequence, society might end up spending vast sums to control one chemical based on a risk assessment that uses a conservative safety factor of 1,000 because it appears to be more dangerous than another chemical that is actually more dangerous, but whose estimated risk is based on a safety factor only 10 times its true value.

Because it is rarely necessary to protect the population from only one hazard at a time, it is preferable to make judgments based on the best estimate of the actual risks rather than incorporating large, but varying, safety factors into the PRA. In the real world, which is characterized by limited resources and time to deal with many possible hazards, an overly cautionary approach might be described as "too safe is unsafe."

Measures of Ill Health for Risk Assessment Purposes

As discussed in the *Measures of Health and Disease in Populations* chapter, various measures of ill health have been proposed to take into account, separately or in combination, the degree of prematurity of death and the time lived with nonfatal disease. Several choices must be made in such calculations— for example, whether to use different life expectancies for men and women or for different regions, how to weigh the severity of different kinds of disease, and whether to discount the value of lost life years in the future (as is the practice in economics). Whatever choices are made, it is evident that measures of lost healthy life years, due to both death and disease, are useful as an indicator of environmental impacts.

As the environmental risk transition progresses, and as the time between creation of a risk factor and its expression as disease tends to lengthen, the need for formal risk assessment increases (Smith & Ezzati, 2005). When evaluating measures to control diarrhea, for example, it is reasonable to monitor diarrhea rates, because the diarrhea of today is due largely to environmental exposures that occurred within a few preceding days. A similar logic can be used for increasing heat exposure due to climate change: Heat exhaustion and heat stroke are due to the exposure during the most recent hours or days (Kjellstrom et al., 2009). In contrast, for controlling environmental carcinogens or greenhouse gases, which have decades-long latency periods, waiting until ill effects start to occur would be far too late. Thus, it is necessary to conduct risk assessments as best we can to make predictions well in advance of the actual events, so that appropriate measures can be taken in time.

Assessing Risk: Compared with What?

Risk assessments, when done in isolation, can be misleading. The following two examples are illustrative. Determining the risk imposed by an activity requires, explicitly or implicitly, choosing an appropriate baseline. In discussing the health effects of smoking, the appropriate baseline may be zero: It is possible for people to quit or not take up smoking. For air pollution, by comparison, the choice is not as clear. Natural sources of many air pollutants exist, and human sources are so difficult and expensive to control that a zero baseline is not feasible in many cases. Which level is then

appropriate: the national standard or the WHO guideline value, or perhaps the level of the cleanest city? It is necessary to choose something, if we are to calculate the risk of the incremental pollution above the baseline.

Most human endeavors that are subject to risk assessment (technologies, industries, chemicals, regulations, and others) actually result in a mixture of risk-lowering and risk-raising effects. A new factory in a low-income country, for example, might impose pollution on the public and accident risk on the workers, but it may provide jobs, housing, training, security, and other benefits that could lead to substantial improvements in health. Just because these effects are less direct than the accident risk does not mean they are small. Indeed, the overall impact of industrialization must be risk lowering—otherwise, the high-income countries would be unhealthy rather than the LMICs without industry.

Such risk lowering also occurs in high-income countries. Consider, for example, a pesticide residue on vegetables. Looked at in isolation, it may appear unacceptably risky. Viewed in terms of the overall impact on food cost and intake for low-income people, occupational risks to farm workers, and other factors, however, it may actually lower the overall risk compared with alternatives. This is not to say that all polluting activities will lower risks, but rather that all technologies should be judged on both their risk-lowering and risk-raising propensities. In addition, important equity and justice issues often arise in relation to who experiences the raised risk and who experiences the lowered risk. Unfortunately, our current risk assessment methods are not well developed enough to determine risk lowering, which is a source of potential bias in the results.

In this chapter, we have focused heavily on the methods for understanding hazards via environmental epidemiology and risk assessment. However, an expanding range of approaches and methods are also being developed in the disciplines of ecosystem health and planetary health. These approaches and methods are being applied to understand environments as habitats, the role of ecosystems and ecosystem services in human health, and the integrated positive and negative health impacts of environmental exposures; they also support decision making for health, equity, and environmental sustainability. Broad approaches include the following:

- Health impact assessment (HIA) of policies and projects outside the health sector, which aim to examine proposals for multiple health effects (Briggs, 2008)
- Health in All Policies, which aims to more deeply integrate health into policy making outside the health sector (including policies governing the natural and built environments) through partnership, co-funding, accountability, and management of vested interests (WHO, 2013a)
- Integrated catchment management, which is described in the "Freshwater" section of the chapter

Methods range from qualitative interviews and narrative methods (Charron, 2012) to community epidemiologic intervention studies examining multiple outcomes, geographic information systems (GIS) mapping and analysis (Delin & Shilong, 2016), and systems thinking and modeling (Brown, Proust, Spickett, & Capon, 2011; Macmillan et al., 2014).

Occupational Health

The workplace environment is generally more dangerous to human health than the ambient external environment. Machinery, chemicals, dusts, ergonomic hazards, and the fact that much work is carried out with the body at its peak performance are all factors that contribute to the overall level of risk. Nevertheless, many of the same specific hazards occur in the workplace and the general environment; hence there are many similarities in how their effects can be monitored and managed (Rosenstock, Cullen, Brodkin, & Redlich, 2005).

A variety of textbooks on occupational health and occupational medicine describe the many workplace hazards to health (e.g., Levy, Wegman, Karon, & Sokas, 2011; Rosenstock et al., 2005; Smedley, Dick, & Sadhra, 2013; Stellman & International Labor Office [ILO], 1998). The issues in LMICs have been highlighted by the web-based detailed handbook by Elgstrand and Petersson (2009), and a detailed treatise is given by Herzstein and colleagues (1998).

The convenient access via the Internet to information about different occupational health hazards has reduced the need for access to major textbooks giving details of each health hazard. One important source is the *Encyclopaedia of Occupational Health and Safety* (ILO, 2012), which is a web-based, updated version of the original report (Stellman & ILO, 1998). Another source of health information for a wide variety of industrial chemicals is the collection of International Chemical Safety Cards, also available via the ILO website (http://www.ilo.org/safework/info/publications/WCMS_113134/lang--en/index.htm). In addition,

the ILO website makes a substantive number of databases and statistics available. In this section, we have kept many of the references from previous editions of this text, and revised or added references only when some important new evidence has become available.

Important phenomena in the workplace environments of LMICs include the problems that occur during a time of rapid industrialization. Agricultural societies are transformed during this rapid transition, but often without the infrastructure for environmental and workplace health protection that was built up in high-income countries over their many decades of industrial development. Sometimes the new industries replace existing cottage industries, and conditions may improve. In other cases, industries with outmoded, dangerous technology are moved from high-income countries, creating new hazards in the receiving LMIC. The phrase "export of hazards" has been used to describe this problem (LaDou & Jeyaratnam, 1994; Moure-Eraso, Wilcox, Punnett, Copeland, & Levenstein, 1994). A new feature of this export of hazard is the transfer of much consumer product manufacturing (e.g., clothes, shoes, furniture, toys) to tropical countries with high ambient heat levels during much of each year (Kjellstrom, Freyberg, Lemke, Otto, & Briggs, 2017; Kjellstrom, Holmer, & Lemke, 2009). Few factories in such countries have efficient air cooling systems, and the resulting workplace heat stress is a major health and productivity threat, becoming worse as global climate change progresses. Another growing occupational health challenge is the increasing use of new migrants and "guest workers" in the economies of many high-income countries. The health and safety management of such workers is often less developed than for native workers (Arcury, Grzywacz, Sidebottom, & Wiggins, 2013).

Agriculture

Agriculture is the most common occupation in rural areas of LMICs, where most of the world's population lives. Most workers are engaged in subsistence agriculture, in which the boundaries between work and other aspects of daily life are fluid. Workplace hazards in this type of situation include generally poor and unhealthy living environments, which are often characterized by unsafe drinking water, poor sanitation, and inadequate shelter.

Specific hazards may also exist in these areas, such as injury hazards from tools used in tilling the soil, vector-borne diseases related to walking in water or mud, hazardous heat exposure under extreme conditions, bites from insects and animals, and falls or drowning from working on hillsides or riversides. The health risks are further increased because subsistence farm work involves the whole family, including children and the elderly. Epidemiologic evidence of these types of health risks is scanty.

As the planting and harvesting processes become more advanced and farmers develop cash crop production, new hazards emerge. The tools involved become more mechanized, which creates new types of injury risks, particularly because the use of these tools is unfamiliar to many of the workers. Pesticides are introduced, but often without the provision of full equipment and training (Thundiyil, Stober, Besbelli, & Pronczuk, 2008; WHO, 1990). As a consequence, an estimated 3 million cases of unintentional pesticide poisoning occur among agricultural workers in LMICs each year (WHO, 2004). In addition, pesticides have been estimated to be used as a deadly poison in 30% of global suicides, with 250,000 people taking their life in this way each year (Gunnell, Eddleston, Phillips, & Konradsen, 2007); the *Unintentional Injuries and Violence* chapter discusses self-directed violence in more detail.

Agricultural work hazards are not only related to planting and harvesting food crops. An important activity is the collection of fuel wood for cooking and heating. Wood and agricultural waste are often collected in local forests by women (Sims, 1994). An analysis of time use for these activities, along with water collection and cooking, in four low-income countries showed that women spent 9 to 12 hours per day engaged in such tasks, whereas men spent 5 to 8 hours. Firewood collection may be combined with harvesting of wood for local use in construction and small-scale cottage industry manufacturing. A number of health hazards are associated with the basic conditions of the forest, including insect bites, stings from poisonous plants, cuts, falls, and drowning.

In countries with tropical heat and humidity, great physiological strain is placed on the body, whereas the cold is a potential hazard in temperate countries and the Arctic area. Excessive heat exposure is not just a health hazard, but also affects work capacity and productivity (Kjellstrom et al., 2017; Sahu, Sett, & Kjellstrom, 2013), which can reduce incomes for already vulnerable low-income workers. In countries with a high sunshine level, ultraviolet radiation can be another health hazard, increasing the risk of skin cancer and cataracts (Gallagher & Lee, 2006).

Most forestry work is hard physical labor that is associated with a risk of ergonomic damage, such as painful backs and joints as well as fatigue, which increases the risk of injuries from falls, falling trees, or

equipment (Poschen, 1998). Women carrying heavy loads of firewood—and also at risk of ergonomic damage—are common in areas with subsistence forestry (Sims, 1994). Further, the living conditions and sanitary facilities of forestry workers are often poor, and workers may spend long periods in simple huts in the forest with limited protection against the weather.

Urbanization typically leads to the development of a commercial market for firewood and larger-scale production of firewood from logs or from smaller waste material left over after the logs have been harvested. Energy forestry then becomes more mechanized, and workers are exposed to additional hazards associated with commercial forestry (Poschen, 1998). Motorized hand tools (e.g., chainsaws) become more commonly used—a trend that leads to high injury risk, noise-induced hearing loss, and "white finger disease" caused by vibration of the hands (Griffin, Bovenzi, & Nelson, 2003). In addition, fertilizer and pesticides become a part of the production system, which results in the potential for pesticide poisoning among sprayers. As the development of forestry progresses, more of the logging becomes mechanized with large machinery, reducing the direct contact between workers and materials. Workers in highly mechanized forestry have only 15% of the injury risk of highly skilled forestry workers using chainsaws (Poschen, 1998). In contrast, firewood production continues to require manual handling and, therefore, may remain a hazardous operation.

The mortality and morbidity data for occupational hazard impacts on populations are in many countries severely underreported. However, based on the statistics reported to the ILO, an idea of the reported mortality can be gleaned (**TABLE 11-3**). Agriculture and fishing have relatively high injury mortality rates in the four countries assessed in Table 11-3, while mining, has even higher rates, particularly in Turkey and the Philippines.

Mining and Extraction

Mining is inherently dangerous to the mine workers, as highlighted by regular media reports of severe incidents in different types of mines. Most countries, including LMICs, have recognized this fact and developed specific legislation and systems to protect mine workers (Eisler, 2003; ILO, 2012; Ramani & Mutmansky, 1999; Stellman & ILO, 1998).

Two major types of mining exist, each with a somewhat different pattern of health hazards: underground mining and open-cast mining. Ergonomic hazards and

TABLE 11-3 Occupational Injury Mortality Rates in Different Economic Sectors, Five-Year Averages for the Latest Years of Reports, 2009–2016

Sector (ISIC Code)	Sweden	United States	Turkey	Philippines
All occupations	1.0	3.4	15	5.4
Agriculture and fishing (A + B)	6.2	24	9.7	29
Mining and quarrying (C)	7.0	16	56	44
Manufacturing (D)	0.8	2.2	13	4.1
Electricity, gas, water supply (E)	1.0	2.8	23	17
Construction (F)	2.2	9.5	24	9.2
Trade and restaurants (G + H)	0.4	1.7	5.5	5.9
Transport (I)	3.2	14	21	18
Financial and business (J + K)	0.0	0.4	0.4	0.0
Administration, education, health, and other (L–O)	0.4	2.1	5.9	1.8

Note: The unit here is the number of deaths per 100,000 economically active workers in the sector.
Data from International Labor Organization statistics (Kjellstrom & Hogstedt, 2009).

physical (accident-inducing) hazards occur in both, but underground work includes the added hazards of being crushed by falling rock, poisoned by gas or dust buildup, or affected by heat or radiation. Each type of mine entails specific hazards associated with the rock from which the ore is excavated. Most types of rock contain high levels of silica, leading to high levels of silica dust in the air of a mine and the risk of silicosis in workers. Certain types of rock (particularly uranium ore) contain radioactive compounds that are emanated as the gas radon, which increases the risk of lung cancer. Other types of rock contain metals that are inherently poisonous (e.g., lead and cadmium) and that in certain conditions can cause dangerous exposures.

According to the United Nations' Demographic Yearbooks, miners constitute a large occupational group on a global level. They represent as much as 2% of the economically active population in some countries. Although in recent years only 1% of the global workforce has been engaged in mining, this industry accounts for 8% of all fatal occupational accidents—approximately 15,000 deaths per year. A detailed review of occupational health and safety issues in mining was performed by Armstrong and Menon (1998).

Clearly, mining is a particularly dangerous occupation. Table 11-3 highlights the high overall occupational mortality rates in selected countries. In LMICs, coal mining employs millions of people. Coal is a major global energy source, contributing 28% of total energy consumption (International Energy Agency, 2017). It was the primary source of energy for the world between 1900 and 1960, but subsequently was overtaken by oil, much of it collected via more difficult extraction, such as "fracking" and tar sands extraction. Coal can be produced through surface mining (open cast) or underground mining. Both operations are inherently dangerous to the health of the workers.

Underground coal miners are exposed to the hazards of excavating and transporting materials underground. These risks include injuries from falling rocks and falls into mine shafts, as well as injuries from machinery used in the mine. There are no reliable global data on injuries of this type from LMICs (Jennings, 1998), but in high-income countries miners have some of the highest rates of compensation for injuries. In addition, much of the excavation involves drilling into silica-based rock, which creates high levels of silica dust inside the mine. As a consequence, silicosis is a common health problem in coal miners (Jennings, 1998). In addition, coal miners with silicosis have an increased risk of lung cancer.

Other health hazards specific to underground coal mining include the coal dust, which can cause "coal worker's pneumoconiosis," or anthracosis, often combined with silicosis. Coal dust is also explosive, so explosions in underground coal mines are an ever-present danger for coal miners. Fires in coal mines are not uncommon, and once started may be almost impossible to extinguish. Apart from the danger of burns, the production of smoke and toxic fumes creates great health risks for the miners. Even without fires, the coal material produces toxic gases when it is disturbed—namely, carbon monoxide, carbon dioxide, and methane. Carbon monoxide binds to hemoglobin in the blood, blocking oxygen transport and causing chemical suffocation. This colorless and odorless gas gives no warning before the symptoms of drowsiness, dizziness, headache, and unconsciousness occur. Carbon dioxide displaces oxygen in the underground air and can cause suffocation.

Mines where materials other than coal are extracted (e.g., iron, copper, lead, zinc, silver, gold) have similar risks of injuries, hot work environments, and exposure to silica dust. An additional health hazard is exhaust fumes from the diesel engines used in machinery or transport vehicles underground. These emissions contain fine particles, nitrogen oxides, and carbon monoxide, all of which can create serious health problems. Underground mining may also expose the workers to extreme heat stress due to high temperature or humidity (Wyndham, 1969). The climatic conditions underground need to be assessed when scheduling work shifts and breaks.

Surface mining avoids the hazards of working underground but still involves risks from machinery, falls, and falling rocks. In addition, mining is energy-intensive work, similar to forestry, and heat, humidity, and other weather factors can affect the workers' health. The machinery used is noisy, and hearing loss is a common effect in miners. Another health hazard is the often-squalid conditions in which many miners in LMICs live, which creates particular risks for the diseases of poverty.

A special type of mining is the extraction of oil and gas to supply the energy needs of societies in the midst of industrialization. Oil and gas exploration, drilling, extraction, processing, and transportation involve a number of the hazards mentioned previously: heavy workload, ergonomic hazards, injury risk, noise, vibration, and chemical exposures (Kraus, 1998). This type of work is often carried out in isolated geographic areas that are subject to inclement weather conditions. Long-distance commuting may also be involved, which increases fatigue, stress, and traffic accident risks. The ergonomic hazards have the potential to cause back pain and joint pain. Injuries may include

burns and those caused by explosions. Workers need well-designed protective clothing to avoid skin damage from exposure to the oil itself and from chemicals used in the drilling processes. In addition, many oil and gas installations have used asbestos for heat-insulating cladding of pipes and equipment, which leads to the hazard of inhalation of asbestos dust in the installation and repair of such equipment, which in turn increases the risk of lung cancer, asbestosis, and mesothelioma (Kamp, 2009; WHO, 1998, 2006).

Much exploration and drilling for oil and gas now occur offshore. This type of mining involves underwater diving work, which is inherently dangerous. In addition, the weather-related exposures can be extreme, particularly because the work often requires continuous, around-the-clock operations (Kraus, 1998).

Construction

Construction work is another dangerous occupation; workers may potentially be exposed to a variety of hazards (Weeks, 1998). Risks include injuries from falls and falling objects, from machinery, or related to excavation or underground work. Because much construction work is carried out in the open, weather conditions may create hazards related to heat, cold, ultraviolet radiation, and dust storms. Construction work also involves heavy lifting of materials and activities in awkward body positions, leading to ergonomic hazards. Injuries, strains, and sprains are common. Many injuries are severe, leading to the high ratio of construction workers in the occupational mortality statistics (see Table 11-3).

In many cases, construction work involves exposures to heat, cold, noise, chemicals, and biological hazards. Much of the machinery used is noisy, and this problem has only increased with the increasing mechanization of the industry. Demolition is a common aspect of construction work, and demolition activities are inherently noisy. As a consequence of these factors, noise-induced hearing loss is common among construction workers. Another aspect of the noisy environment is the increased safety problems caused by the masking of warning calls or other alarms.

Chemical and dust exposures are related to the composition of the building materials. Asbestos, which has long been used as insulation and as a component of asbestos-cement pipes and sheets, is a prime example of a hazardous material. Asbestosis (a form of pneumoconiosis), lung cancer, and mesothelioma (another fatal cancer strongly associated with occupational asbestos exposure) have been found in many construction workers (WHO, 1998). Indeed, mesothelioma is one of the most prominent occupational diseases in most countries where records are kept, while underreporting of asbestos-related lung cancer is common. In high-income countries, as many as 20,000 cases of lung cancer and 10,000 cases of mesothelioma occur each year due to workplace asbestos exposures that took place 20 to 40 years ago (Tossavainen, 2000).

The use of asbestos in LMICs remains widespread (Kameda et al., 2014; Kazan-Allen, 2003; Tossavainen, 2004) and, indeed, is continuing to increase (Le et al., 2010). The epidemic of serious health effects associated with exposure to asbestos takes decades to develop, so preventive actions need to be taken at an early stage. Some LMICs have instituted bans on use of asbestos in construction, which would be the most effective means of prevention. Alternatives to asbestos exist for almost every use, and their cost is often similar to that of asbestos. The epidemic of asbestos-related disease is still on the ascendancy in a number of high-income countries, even though virtual bans were instituted in the 1980s—for example, in New Zealand, where the full epidemic of asbestos cancer over 40 years may kill at least 2,000 to 3,000 workers (Kjellstrom, 2004).

Other chemical and dust exposures in the construction industry include cement dust among bricklayers and concrete workers (Meo, 2004); such dust causes lung disease and dermatitis. Sand-blasting or rock drilling creates silica dust in the air, which can lead to silicosis in exposed workers. Construction work often involves welding, which adds further health hazards, such as inhalation of welding fumes that leads to bronchitis. Paint fumes often contain organic solvents that may cause neurologic disorders. The hazards of dusts and fumes are increased inside confined spaces, where these concentrations can reach extremely high levels.

Because of climate conditions, work in the construction industry is often seasonal, so contract workers may rely on other work during parts of each year. This seasonality is influenced by climate. In tropical and subtropical countries, extreme heat exposure during the hot season may cause major health and productivity issues (Kjellstrom, Holmer, & Lemke, 2009; Nybo, Kjellstrom, Bogataj, & Flouris, 2017). The intermittent character of seasonal work also creates problems in maintaining efficient prevention programs to protect workers against these hazards (Weeks, 1998), especially in LMICs. Subcontracting or informal employment relations often reduce the responsibility taken by the main employer or contractor on a construction site. When the responsibility for health and safety is dispersed among many individuals, the protection against hazards may become insufficient. Thus, in construction work, it is important that

contracts include the necessary safety provisions and that systems be in place for monitoring and enforcing these provisions.

Transport

An important industry in modern society is motorized transport of goods and people. As a country develops economically, transport occupations typically account for an increasing proportion of the workforce. These jobs include drivers of trucks, buses, taxis, trains, airplanes, and other vehicles; people involved inside the vehicles (e.g., conductors, flight attendants); people involved in loading and offloading freight; and people involved in the management of the transport system (e.g., traffic police, air traffic controllers).

The motorized transport system creates several health hazards in communities as described in other sections of this chapter (e.g., air pollution, noise, injury risks). The health risks specific to the people working in the transport industry are also of importance. Table 11-3 shows the relatively high injury and mortality risks for transport workers in four countries. Most of these deaths are likely to be listed as "traffic accidents" in official statistics. Other occupational health hazards include exposures to noise, chemicals, and dusts, as well as ergonomic hazards associated with lifting heavy loads and sitting for hours in the same position while driving a vehicle (Kjellstrom, Ferguson, & Taylor, 2009).

Industry and Manufacturing

Industrialization brings many benefits in the form of income and jobs, but, unless regulated in some fashion, can lead to significant occupational and public health hazards. These health hazards may take the form of releases of toxic or potentially toxic materials—as in the notorious cases in postwar Japan (**EXHIBIT 11-4**).

EXHIBIT 11-4 Minamata and Itai-Itai Disease: Classic Environmental Health Disasters

Environmental health disasters have been important triggers for national and international action to prevent environmental pollution. The best-known such disaster may be Minamata disease, which struck the small coastal town of Minamata, Japan, in 1956 (ILO, WHO, International Program on Chemical Safety, & United Nations Environment Programme, 1990). Hundreds of people were seriously affected by methylmercury poisoning, and many victims died. This type of poisoning affects the nervous system, producing symptoms that range from slight numbness of the fingers to loss of the ability to talk and walk.

The source of the methylmercury in Minamata was a chemical production factory that used mercury as a catalyst in one of its processes. Surplus mercury was discharged via spill water into a nearby bay, and this mercury accumulated in bottom sediments. Microbes in the sediments converted the mercury to methylmercury, which eventually entered the food chain of fish and caused high methylmercury levels in the local fish. Minamata had a substantial population of small family fisheries, and these families were the most affected.

The outbreak of the disease developed over several months, and it was initially thought that a new type of infectious disease affecting the nervous system had appeared. It took months of detailed epidemiologic research to conclude that the cause of the disease was associated with the consumption of fish. Further toxicologic and epidemiologic research over many years eventually identified the specific chemical involved. A second outbreak of similar methylmercury poisoning, in Niigata, Japan, intensified the search for a definite cause. In 1968, the Japanese government committee responsible for elucidating the cause finally incriminated methylmercury—12 years after the disease was first reported.

A similar story can be told about Itai-Itai disease, a form of chronic cadmium poisoning that developed in farmers in Toyama, Japan, at about the same time as the first outbreak of Minamata disease (Friberg, Elinder, Kjellstrom, & Nordberg, 1986; WHO, 1992). Painstaking research identified that the consumption of cadmium-contaminated rice and drinking water was the cause. The cadmium, which came from a mining area and a lead/zinc ore concentration plant, had reached the affected community via a river that residents used to irrigate their rice fields. The farming families in the contaminated area had small subsistence farms, providing for all of their needs. Thus, if a family's farm was contaminated, the family members ended up with high daily cadmium intake. Cadmium is a cumulative poison that eventually damages the kidneys, indirectly leading to bone deformities and fractures because of severe osteomalacia and osteoporosis.

At the time when these two disease outbreaks occurred, Japan was a low-income country trying to recover from the disastrous economic effects of World War II. The living conditions were not dissimilar to those in today's low-income countries undergoing rapid industrialization. Rural populations consumed mainly locally produced food. Only basic healthcare services were available, and the environmental pollution situation was not closely monitored or managed. Local industry discharged wastes into the air, rivers, or sea without much pollution control equipment.

A country that currently has similar conditions to erstwhile Japan is China, which also has lead, zinc, and copper mining; rice farming; and high local food content in the diet. Indeed, cadmium-polluted areas have been found in China (Cai, Yue, Shang, & Nordberg, 1995), and environmental epidemiologic studies have discovered exposures and effects similar to those observed in the polluted areas of Japan.

Some industrial facilities carry the risk of large-scale accidental releases of toxic materials. The biggest such release in world history occurred in Bhopal, India, where an explosion at a pesticide manufacturing plant resulted in some 3,000 deaths caused by the chemical methyl isocyanate and significant health impairment in many tens of thousands of people. The impact of this accident at the facility was exacerbated by the lack of urban zoning controls, with hundreds of households having been built directly adjacent to the plant. There was also inadequate planning for alerting and evacuating the public once the accident had occurred.

Manufacturing workplaces can be the sites of any of a long list of occupational hazards, some of which were mentioned in earlier sections of this chapter. A review of the various hazards is included in the *Encyclopaedia of Occupational Health and Safety*, published by the International Labor Organization (ILO, 2012; Stellman & ILO, 1998). Manufacturing involves ergonomic injury hazards from improper work positions, heavy lifts, and dangerous machinery; physical hazards, such as noise, heat, poor lighting, and, occasionally, radiation; and chemical exposures of many kinds. However, as in the earlier history of today's high-income countries, the major chemical exposure problems in LMICs include lead, cadmium, chromium, mercury, other metals, organic solvents, and welding fumes. Increasingly, the most hazardous industries and processes are being exported from high-income countries to LMICs (LaDou & Jeyaratnam, 1994), but without the technological improvements that have reduced workers' exposures to these risks in the high-income countries. In addition, stress and other psychosocial hazards of long work hours and shift work are common in industries that have made large investments in machinery, from which economic benefits accrue only when the equipment is in operation. Notably, the transfer of consumer products manufacturing (e.g., clothes, shoes, furniture) to low-income countries has created health risks due to limited applications of occupational health and safety management precepts.

Of the various manufacturing industries, some with particular health risks are worth highlighting. In electrical appliance manufacturing (Stellman & ILO, 1998), major risks are found in lead-acid battery manufacturing operations, which produce batteries for vehicles. Because such batteries are too heavy to transport over long distances, local production is usually established at an early stage of the "motor car society." The operation of these factories typically involves many workers who receive unacceptably high lead exposures. The usual approach is to monitor workers' blood lead levels; if the levels exceed the national standard, the worker is taken off lead exposure work for a few weeks. Indeed, this type of risk management is enshrined in occupational health law in many countries. This approach displaces the exposure problem to the individual rather than analyzing the workplace environment as a whole.

Cadmium is another toxic metal of increased importance; it is used in rechargeable batteries in modern electronic equipment such as mobile phones. Unfortunately, the manufacturing of cadmium batteries involves even more risks than lead battery production. Vigilant exposure monitoring and effective protection are essential to avoid chronic poisoning in the form of kidney and lung diseases (Friberg et al., 1986).

Another manufacturing industry commonly found in high-income countries is metal processing and metal working. Smelting and refining of any metal create a major potential for occupational exposures to many types of hazards, especially exposure to toxic metal dusts, sulfur dioxide, and other fumes. Because these industries often involve large-scale operations, even small concentrations of toxic compounds in the processes can yield substantial emissions into the workplace and the surrounding environment. The experience with lead smelters in many countries has produced similar outcomes: high lead exposures for workers and contamination of the local environment. Often the workers and their families live in the vicinity of the industry, and high lead exposures from dust emissions are found in children who live in the area (Baghurst et al., 1992). These exposure situations have been studied in detail in the United States and Australia, and epidemiologic research there has produced some of the most valuable quantitative data on the health risks of lead in children and workers (see Exhibit 11-2).

Service Occupations

Many service industries involve important occupational hazards. The services reviewed by the International Labor Organization include those characterized by specialized and sometimes severe hazards, such as emergency and security services (e.g., firefighting and law enforcement), public and government services (e.g., garbage collection and hazardous waste disposal), and healthcare services (Stellman & ILO, 1998). Occupations that are considered likely to have less severe hazards include retail trades, banking, administrative services, telecommunications, restaurants, education, and entertainment. Nevertheless, stress due to increased mental workload and various

forms of interpersonal conflicts in such settings can become major occupational health hazards.

Firefighting involves exposure to carbon monoxide and toxic fumes, as well as the heat from the fire itself. Injuries from falling debris, falls, or working in awkward positions are also of concern. Protection of workers depends on the availability and use of protective equipment, which may be in short supply in LMICs. Law enforcement is another high-risk occupation, as it involves hostile contacts with persons who may be armed.

Garbage collection exposes workers to risks of cuts and other injuries from the garbage itself, as well as ergonomic hazards (e.g., heavy lifting) and chemical hazards. A notable at-risk group in LMICs is the people who scavenge on garbage dumps for recyclable materials from which to glean a meager existence. Sometimes these scavengers actually live on the garbage dump, where they are subjected to risks of infectious disease, bites from rats and dogs, and other dangers. Because hazardous wastes are not always separated out from general wastes, their presence adds to the risks encountered by the garbage collectors and scavengers. In areas where hazardous wastes are separated, the storage and handling of these materials requires sophisticated protective equipment, detailed information about the hazards, and efficient management systems; these factors are often missing even in high-income countries.

Healthcare workers face hazards such as infections from patients, transmission of HIV or hepatitis from needlesticks, and allergies to medications or to cleaning and disinfection chemicals. More generally, the most common problem faced by these workers is ergonomic hazards from lifting or moving patients, which may lead to back injuries. This type of injury creates great problems for nurses and nurse aides, and in many cases curtails their careers. Another service occupation with particular occupational health hazards is the tourism industry.

One special hazard for people in service trades that involve continuous work at computer keyboards is the development of repetitive strain injury, also known as occupational overuse syndrome (Rosenstock et al., 2005). Repetitive keyboard finger work, or repetitive fine movement of the computer mouse, may create a wear-and-tear reaction in tendons and muscles that can lead to chronic pain. As with back problems, these painful conditions are not always accompanied by measurable anatomic or pathological changes, which has caused substantial arguments among medical practitioners as to the genuineness of the disease. However, as many people using modern computers attest, the short-term pain after intensive use of a keyboard or a mouse is real. Ergonomic design of computer workstations and the provision of regular work breaks are essential for preventing this occupational hazard. The height of the keyboard should be adjusted to the individual user.

Other Occupations

Among the other occupational exposure situations of particular importance, especially in LMICs, are cottage industries of various types. At an early stage of industrialization, small-scale operations based on family members may be the mainstay of certain industries. They may take the form of work contracted out from a larger enterprise, or the businesses may arise directly in relation to the local market. The production of handicrafts, clothing, and consumer items for local households may be the starting point, although more hazardous activities, such as recycling car batteries, may also develop initially as cottage industries. Such work may entail extreme exposures to toxic chemicals, with little or no protection either for the workers or for other family members. Ergonomic hazards, injuries, noise damage, and all other occupational hazards are likely to be a greater danger in these cottage industries than in larger, more organized enterprises.

Major Issues in Environmental Health

The next sections describe five current major issues in environmental and occupational health. These issues illustrate well the measurement, research and practice methods, boundary, and policy challenges that have been outlined so far, as well as including aspects of environmental and occupational health. All combine multiple environmental hazards and multiple sources for those hazards, with complex sources and interactions, acting at a range of scales, from the household to global. These five issues are indoor and outdoor air pollution, freshwater, food production, cities and the built environment, and climate change. The last is the most challenging global environmental health issue of our time, and it potentiates many of the environmental hazards discussed in the remainder of the chapter.

Air Pollution

Air pollution, a complex mixture of gases and particles, is one of the most important risk factors for disease. While thousands of compounds are found in air

pollution, governments tend to focus on a specific subset as indicators of the overall air quality. In this section, we primarily address small particles, one of the most commonly measured and best-understood air pollutants. According to the Institute for Health Metrics and Evaluation's Global Burden of Disease (GBD) study, exposure to small particles contributed to approximately 6.1 million deaths (11% of the total) and 163 million DALYs (6.8% of the total) in 2016 (IHME, 2016).

The discovery and "taming" of fire is an event described as the transition between the prehuman and human periods. Since this development, people have polluted the air both inside and outside of their homes, first for cooking and heating, and later for transportation and as by-products of industrial activity and waste disposal, among other sources. Air pollution now occurs at many scales:

- In households, where biomass fuels are still burned for cooking, heating, lighting, or for other tasks
- At local levels, where numerous sources (automobiles, industry, and natural events) degrade ambient air quality
- Regionally and globally, as transport of pollutants across boundaries results in impacts far from the original sources

The burden of ill health resulting from exposure to ambient and household air pollution estimated in the GBD study focuses on five endpoints: heart disease, stroke, lung cancer, and chronic obstructive respiratory disease in adults and lower respiratory infection in children under 5. The majority of ill health arises from exposure to $PM_{2.5}$—that is, particulate matter, or soot, with an aerodynamic diameter of less than 2.5 μm (depicted next to a human hair and a grain of beach sand for comparison in **FIGURE 11-6**).

Ambient and household $PM_{2.5}$ exposure rank among the top global risk factors for mortality and account for more deaths than many other better-recognized risk factors, including alcohol use, high body mass index, and high salt intake. These estimates most likely underestimate the true global health impact of air pollution, as new evidence points to associations with neurologic and cognitive deficits (Dix-Cooper, Eskenazi, Romero, Balmes, & Smith, 2012; Morales et al., 2009; Perera et al., 2006; Suades-Gonzalez, Gascon, Guxens, & Sunyer, 2015; Suglia, Gryparis, Wright, Schwartz, & Wright, 2008), diabetes (Eze et al., 2015), asthma (Oluwole, Arinola, Huo, & Olopade, 2017a, 2017b; Wong et al., 2013), other cancers (Josyula et al., 2015), and pregnancy and birth outcomes (Amegah, Quansah, & Jaakkola, 2014).

The remainder of this section provides overviews of household and ambient air pollution and briefly discusses regional dispersion of air pollution.

Household Air Pollution

Use of the oldest of human energy technologies—the home cooking and heating fire fueled by wood—began approximately 1.8 million years ago (Wrangham, 2009). As cultivation and farming took hold, so, too, did burning of agricultural waste, including animal dung and crop residues, for heating and cooking. Approximately 1,000 years ago, coal use peaked in several European countries, where it was easily obtained; coal is still used widely in China today (Smil, 1994). The three broad fuel types—coal, wood, and crop residues/dung—are still used as a primary source of energy for cooking and heating by approximately 40% of households globally. They are commonly burned in simple cookstoves, resulting in large amounts of smoke termed "household air pollution."

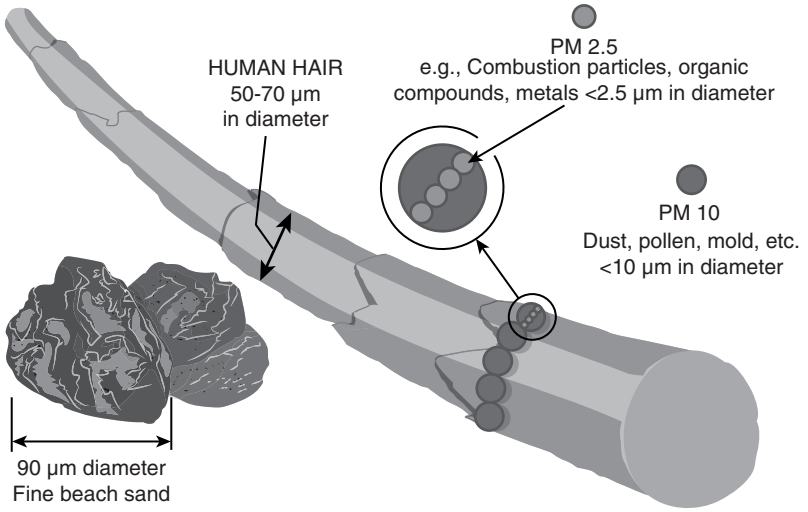


FIGURE 11-6 The size of $PM_{2.5}$ in comparison with a human hair, beach sand, and PM_{10}.

Reprinted from The Lancet, 383(9928), Guarnieri, M. and J. R. Balmes, Outdoor air pollution and asthma, 1581–1592, Copyright 2014, with permission from Elsevier.

A useful framework for examining the trends and impacts of household fuel use is the "energy ladder." It ranks household fuels along a spectrum running from the simple biomass fuels (dung, crop residues, and wood), through the fossil fuels (kerosene and gas), to the most modern form (electricity). In moving up the ladder, the fuel and stove combinations that represent the higher rungs increase the desirable characteristics of cleanliness, efficiency, storability, and controllability. At the same time, capital cost and dependence on centralized fuel cycles tend to increase with upward movement on the ladder. Although local exceptions are possible, history has generally shown that when alternatives are affordable and available, populations tend to naturally move up the ladder to use of higher-quality fuel and stove combinations. Indeed, 60% of the world's population uses gas or electricity for nearly all of their cooking. The remaining 40% of the world's population is either still using wood or has been forced down the ladder by local wood shortages to crop residues, animal dung, or, in some severe situations, the poorest-quality fuels, such as shrubs and grass.

Exposure to household air pollution arising from cooking resulted in approximately 2.8 million deaths in 2016 (IHME, 2016). **TABLE 11-4** lists some of the many hundreds of health-damaging pollutants (HDPs) emitted as products of incomplete combustion from a range of household solid-fuel stoves. Even though biomass fuels might produce few wastes other than carbon dioxide and water when combusted completely, in practice as much as one-fifth of the fuel carbon is diverted to products of incomplete combustion, many of which are important HDPs (Naeher et al., 2007). Unfortunately, not only do solid-fuel stoves produce substantial HDPs, but a large fraction lack chimneys for moving the emissions out of the living area. Consequently, household air pollution concentrations can reach high levels for many HDPs, including $PM_{2.5}$ and carbon monoxide. For example, fine particulate levels often reach 20 to 40 times the WHO guideline level of 10 $\mu g/m^3$ set to protect health. Furthermore, in areas in which households have chimneys, heavily polluting stoves can produce significant local or neighborhood outdoor pollution. This is particularly true in densely populated urban slums, where such neighborhood pollution can greatly exceed urban average levels. Globally, it is estimated that approximately 12% of total ambient air pollution comes from household sources; in India, this contribution is approximately 26%, while in China it is approximately 10% (Chafe et al., 2014).

Coal, by comparison, not only is difficult to burn completely because of its solid form, but also can contain significant intrinsic contaminants. Most prominent among such emissions from coal are sulfur oxides, but in many areas coal smoke also contains arsenic, fluorine, lead, mercury, or other toxic elements that lead to serious HDPs. Recent meta-analyses provide strong evidence that exposure to hazardous air pollutants from coal smoke increases the risk of lung cancer. In fact, the International Agency for Research on Cancer (IARC) classified coal emissions as a Group 1 carcinogen (Hosgood et al., 2011; Straif et al., 2006; Zhang & Smith, 2007).

China provides a practical example of the dangers of coal smoke exposure and an effective initiative to reduce this exposure. Since the 1980s, this country has introduced more than 180 million improved stoves with chimneys—more than in all other LMICs combined. In Yunnan province, this strategy has led to a 50% reduction in the risk of lung cancer and chronic obstructive pulmonary disease (COPD) (Zhang & Smith, 2007).

Petroleum-based liquid and gaseous fuels, such as liquid petroleum gas and natural gas, may also contain

TABLE 11-4 Major Air Pollutants, Their Physical State, and Their Sources

Pollutant	State	Sources
Ozone	Gas	Generated from other pollutants and natural processes
Sulfur dioxide	Gas	Fossil fuel combustion, natural emissions
Carbon monoxide	Gas	Fossil fuel combustion
Nitrogen oxides	Gas	Combustion processes
Particulate matter	Particulate	Combustion, fertilizer use, dust storms
Hazardous air pollutants, including benzene, formaldehyde, dioxin, and others	Gas	Incomplete combustion, chemical processing, solvent use

Data from *IARC Monographs Volume 109: Outdoor Air Pollution* (2013).

sulfur and other contaminants, albeit in much smaller amounts than in many coals. Further, their physical forms allow for much better premixing with air in simple devices, thereby assuring substantially higher combustion efficiencies and lower HDP emissions. In addition, gas stoves tend to be much more energy efficient. Hence, the HDP emissions per meal from these fuels are at least an order of magnitude less than those from solid fuels.

One of the significant challenges to reducing the burden of disease associated with household air pollution is choosing a set of solutions that both decreases emissions of HDPs and is suitable to the household to the extent that they nearly entirely stop use of the older, polluting traditional stove. Modern, clean fuels—such as liquefied petroleum gas (LPG) or natural gas—or electric cooking are an ideal solution, but may remain out of reach of many households who either cannot afford or access these technologies. A number of improved biomass stoves have been created and have shown promising reductions in HDP emissions in the laboratory, only to perform poorly in the field or to be used only occasionally for specific tasks, negating their benefit. Much emphasis is currently placed on finding ways to make clean, modern fuels available to the poor through innovative policy programs, targeted social welfare investments, and other policy paradigms.

Outdoor/Ambient Air Pollution

Ambient air pollution has, in recent decades, become recognized as a global health problem (Health Effects Institute, 2017; Landrigan, 2016; Pope & Dockery, 2006; WHO, 2016), in large part due to a combination of strong epidemiologic evidence and recent and public air pollution episodes in India and China. Between 1990 and 2015, global, population-weighted $PM_{2.5}$ exposures increased approximately 11% (Health Effects Institute, 2017). This change reflected rapid increases in ambient air pollution in heavily populated areas, including large megacities of India and China, but ambient pollution levels also exceed guidelines in large parts of rural Asia as well. For example, approximately 90% of the global population live in areas where the ambient air pollution exceeds WHO's exposure guideline of 10 $\mu g/m^3$. Approximately 85% of those living with the most extreme average concentrations (greater than 75 $\mu g/m^3$) reside in China, India, Pakistan, and Bangladesh (Health Effects Institute, 2017). Note that the mortality impacts of air pollution are increased by high environmental heat exposures, as happened during the European heat wave in August 2003 (Dear, Ranmuthugala, Kjellstrom, Skinner, & Hanigan, 2005).

Unsurprisingly, there are differences in the patterns and levels of exposure to air pollution globally. In high-income countries, industrial and hazardous air pollutants from coal burning that proliferated in earlier eras has been largely replaced by pollutants from motorized transport that form photochemical smog, including ozone (a strong irritant that affects the eyes, upper airways, and lungs) in summer and a heavy haze of particulates and nitrogen oxides in winter (see Table 11-4 for common pollutants of concern and their major sources). Although many industrialized cities do not yet meet annual standards for every pollutant, conditions are generally much better than in the past, with cleaner cities becoming even cleaner, primarily in high-income countries.

In contrast, in LMICs, urban air pollution has reached alarming levels in many cities. In New Delhi, Beijing, and several other Indian and Chinese cities, for example, the annual average concentrations of particulates have been 5 to 15 times greater than the WHO guideline. In China, the main source of pollution is combustion of coal, though the transportation sector and household biomass combustion make only slightly smaller contributions to the deaths attributable to $PM_{2.5}$ pollution (GBD MAPS Working Group, 2016). Both the morbidity and the mortality in Chinese cities due to air pollution are estimated by some observers to be increasing (Chen et al., 2017).

Studies relating ambient air pollution levels to health risks were, until the 1970s, largely confined to examining the health impacts of particular extreme episodes of high outdoor air pollution levels. Subsequent studies, based on daily mortality time series, have elucidated the role of respirable particulates, including black carbon, ozone, sulfates, and nitrogen oxides in acute mortality (Schwartz, 1994). The advantages and attractions of daily time-series statistical analysis were discussed in the section "Study Design Options" earlier in this chapter. Long-term follow-up studies of populations exposed to different levels of air pollution, especially $PM_{2.5}$, indicate that the higher the levels of exposure, the greater the mortality risk.

For example, a 16-year follow-up of 550,000 people in the United States was used to link urban air pollution exposures to mortality with appropriate control of confounding factors (Pope et al., 2002) as part of the American Cancer Society cohort study. An increased mortality in cardiopulmonary diseases was identified, and the strongest relation to air pollution (indicated by exposure to $PM_{2.5}$) was found for lung cancer. Further quantification of the exposure–response relationships based on the same cohort and other studies is now available (Pope, 2009; Pope et al., 2004; Pope, Ezzati, & Dockery, 2009). Since the original study, numerous updates and continued follow-up have occurred (WHO, 2013b), with all showing

remarkably consistent findings. Additional cohort studies from Asia, Canada, and Europe have provided further evidence of the impact of long-term exposure to $PM_{2.5}$ on mortality.

Asthma, whose prevalence has been increasing in high-income countries for three decades, has a still unresolved relationship to external air pollution (Guarnieri & Balmes, 2014). A 2009 meta-analysis found that traffic-related air pollution (TRAP) contributed to the development of asthma symptoms in healthy children (Bråbäck & Forsberg, 2009; Pollock, Shi, & Gimbel, 2017), as did a 2017 systematic review and meta-analysis (Khreis et al., 2017), but other studies and reviews have yielded less conclusive results for ambient air pollution writ large. The apparent increase in the susceptibility of modern generations of children to asthma may derive from changes in human ecology that have altered early-life immunologic experience, such as reduced exposure to childhood infections (due to smaller family sizes) or increased allergenic exposures (e.g., house-dust mites or fungal spores).

Epidemiologists have developed a diverse, sophisticated set of methods for assessing the health impacts of air pollution. Nevertheless, the measurement issue remains bedeviled by difficulties in exposure assessment, the uncertain differentiation of acute and chronic effects, the need to sort out independent and interactive effects between air pollutants that are often highly correlated, and the fact that the air pollution profile keeps evolving as human activity patterns change. Although many uncertainties remain, there is agreement that significant health effects occur at pollutant concentrations that were previously considered benign. In the case of small particulates, which can penetrate deeply into the respiratory system, there is no evidence of a threshold exposure below which no effect occurs. As a consequence, policy makers are forced to decide which level of health risk is acceptable to determine an appropriate exposure standard.

Traffic and Transport

As cities grow in size, their urban transport systems expand and evolve. In particular, private car ownership has increased spectacularly over the past half-century, creating new opportunities and freedoms—and a new raft of social and public health problems (London School of & Tropical Medicine Public Health, 1997; Nieuwenhuijsen, 2016; Nieuwenhuijsen, Khreis, Verlinghieri, & Rojas-Rueda, 2016).

Transportation is one of the key polluters in the process of economic development, urbanization, and industrialization. In traditional subsistence agricultural societies, the community's basic needs could be met within a relatively localized distance. Increases in population size and density meant that specialized resources for the community, such as firewood, had to be acquired from increasingly distant sources, which creates transportation needs. Modern economic development has accelerated this process through further specialization of economic tasks and dependence on resources from distant areas. Energy sources, such as coal, must be transported from afar to sustain local cottage industries; such is also the case with food items to sustain people in places where little can be grown or gathered for much of the year.

Today, the automobile has become the dominant source of air pollution in many cities. The World Bank and IHME estimate that pollution from vehicles is responsible for 184,000 deaths globally (Bhalla et al., 2014). Current technical solutions to this problem include making car engines more energy efficient and using pollution-control devices, such as catalytic converters. The obvious, more radical solution is to reduce dependence on car traffic by encouraging people to walk and to travel by trains, buses, or bicycles. Such measures, carried out as part of a clean air implementation plan, can significantly reduce automotive air pollution, as has been shown in certain towns in the United States and Germany (WHO, 1997). The unprecedented action to reduce emissions that took place during the 2008 Beijing Olympics provides a profound example of the potential efficacy of clean-air implementation plans. Pollution-control measures instituted during that period included limiting automobile usage, closing heavy industrial facilities, and pausing construction work. Collectively, these strategies resulted in a 54% reduction in particulate matter of 10 μm or less and subsequently a 46% reduction in outpatient visits for asthma during the two weeks of the Olympics (Li, Wang, Kan, Xu, & Chen, 2010).

In cities in LMICs, transportation problems have grown even faster. Whereas in Europe a local transport infrastructure based on railways, trams, and buses was already in place when the car boom emerged in the 1960s, LMICs often have negligible transport infrastructure. Hence, with increasing affluence and urbanization, the automobile is seen as the best solution for individual families. As a consequence, cities such as Mexico City experienced a rapidly deteriorating air pollution situation during the 1980s and 1990s (Davis, 2017). However, the recent experience of Mexico City also provides a case study of attempts to reduce car-generated air pollution (Davis, 2017).

Since the first oil crisis in the 1970s, the average fuel efficiency of the world's automobiles has increased substantially, largely through major improvements in North American car manufacturing through the mid-1990s that brought these vehicles' efficiencies nearly

to European levels. Emissions of air pollutants have also been greatly reduced, partly through combustion modifications and partly through extensive application of end-of-pipe controls, in the form of catalytic converters. Nevertheless, if the number of automobiles in LMICs continues to grow as it has in recent years, unacceptable air pollution levels will undoubtedly persist for many decades despite the application of even the best current automotive pollution-control technologies.

Regional Dispersal of Air Pollutants

Following the surge of industrial growth in high-income countries following World War II, the transboundary problem of acid deposition (commonly referred to as "acid rain") became increasingly problematic. It was featured prominently at the 1972 United Nations Conference on the Human Environment in Stockholm. By that time, eastern Canada was experiencing problems from acidic emissions traveling northeast from the United States, and Scandinavia was exposed to emissions from the United Kingdom and highly industrialized areas of West and East Germany. More recently, the strong increase in China's industrial production has subjected Japan to acid deposition from that source. Acid levels are also rising in other parts of Asia as energy use grows across the entire region.

The hazards to human health from this source are neither extreme nor direct. Nevertheless, the acidification of waterways and soils has demonstrably increased the mobilization of various elements—particularly heavy metals and aluminum—that enables them to enter the drinking water and the food chain. Human exposures to these elements have, therefore, increased in several regions.

The El Niño event of 1997–1998 created unusually strong drought conditions in Southeast Asia, which exacerbated the size and duration of forest fires originally lit to clear land. The result, in addition to extensive damage to the forests, was regional air pollution in the form of wood smoke plumes that extended over thousands of miles and could readily be seen by satellites. The forest fires raised outdoor particulate levels to several times the acceptable upper limit in a number of large cities in the region for a period of days to weeks, apparently with significant accompanying adverse health effects. Although it was larger in 1997–1998 than in previous years, the same phenomenon has actually been occurring for decades in the region, and is likely to be repeated in the future unless land-use practices are drastically modified.

Another issue of importance is the development of atmospheric "brown clouds" caused by the confluence of urban air pollution from the numerous cities and major industries in Southeast Asia, India, China, and elsewhere in the world (Ramanathan et al., 2005; Ramanathan & United Nations Environment Programme [UNEP], 2008). These high-level "clouds" result from ground-level emissions, which also cause outdoor air pollution levels to exceed air quality guidelines in large areas of Asia, even those regions located far from cities.

Freshwater

Ever since hunter-gatherers turned to cultivation and settled living, sanitation and unsafe drinking water have been health problems for human societies. Further, although nearly two-thirds of the Earth's surface consists of water, only 2.5% of this is freshwater, and only 11% of this freshwater is available for human use. Political and institutional factors operate on top of this scarcity to create a global water shortage. As the global population grows, direct pressures on local sources of fresh drinking water increase, as do indirect pressures, particularly from the increases in agricultural production required to support the ever-growing population. Despite efforts to mitigate unsafe sanitation and drinking water having been associated with the largest reductions in deaths and DALYs over the past decade, these problems remain among the top 30 risk factors for early deaths and disease, with unsafe water ranked number 16 and unsafe sanitation ranked number 21 on that list (Gakidou et al., 2017).

The management of freshwater perhaps demonstrates most clearly the addition of habitat approaches to existing and ongoing efforts dealing directly with environmental hazards. Clean water and sanitation is one of the 17 Sustainable Development Goals, as well as having a role in Goal 1: No poverty (United Nations, 2015b). Specific targets by 2030 under Goal 6 reflect an ongoing need to improve sanitation (e.g., Target 6.2: Achieve access to adequate and equitable sanitation and hygiene for all and end open defecation), reduce hazardous pollution (e.g., Target 6.3: Improve water quality by reducing pollution, eliminating dumping, and minimizing release of hazardous chemicals and materials), and ensure integrated management of water ecosystems (e.g., Target 6.5: Implement integrated water resource management at all levels, and Target 6.6: Protect and restore water-related ecosystems).

Hazards from Poor Sanitation and Contaminated Drinking Water

Unsafe sanitation and drinking water remain widespread health hazards in the world, particularly in LMICs in semiarid regions. More than 840 million people, almost all in LMICs, have inadequate access to water, and 2.3 billion people still lack basic sanitation (WHO &

UNICEF, 2017). In 2016, a lack of access to clean drinking water, appropriate sanitation, and hygiene facilities resulted in more than 1.6 million deaths from diarrheal diseases around the world (Gakidou et al., 2017) (see also the *Infectious Diseases* chapter).

Two factors need to be taken into account when evaluating the relationship between health outcomes and access to clean drinking water. Most prominent are the linkages between household water availability and the health burden attributable to diarrheal diseases. In addition, intense rainfall and drought play crucial roles in facilitating waterborne outbreaks of diseases through both surface water and piped water supplies (Confalonieri et al., 2007).

Diseases related to water supply can be classified based on their route of transmission. Waterborne diseases, such as cholera and campylobacteriosis, are a result of the ingestion of contaminated water. Water-washed diseases, such as *Chlamydia* infection and scabies, are caused by a lack of hygiene secondary to a lack of water. Some classifications also include categories for water-based diseases, which include schistosomiasis, and diseases transmitted by water-related vectors, with malaria being the typical example.

A lack of sanitation may lead to the contamination of drinking water, highlighting the intertwined nature of these two problems. The waterborne illnesses associated with poor sanitation make the greatest contribution to the overall economic costs of sanitation. In 2015, India suffered the greatest costs associated with health spending, productivity losses, and labor diversions due to inadequate sanitation and clean drinking water, with these costs exceeding $100 billion, more than 5% of India's gross domestic product (GDP) (Lixil, WaterAid Japan, & Oxford Economics, 2016). This figure outstripped the total aid flows and debt relief provided to the country in that year.

The main barrier to improving sanitation facilities is a lack of political will. By ensuring access to clean water and adequate sanitation facilities, the fecal–oral transmission pathway is interrupted, resulting in an enormous flow of improvements in health, poverty reduction, and economic development.

Fecal contamination of drinking water is not isolated to LMICs. Rising income inequalities and poverty in the United States have meant that some low-income populations in southern states continue to have poor sanitation and hygiene. A recent prevalence study of infection with the intestinal nematode *Necator americanus* (American hookworm), focusing on one of the poorest counties in Alabama, found that more than 30% of those people tested were infected with hookworm or other gastrointestinal parasites (McKenna et al., 2017). The implications for anemia, childhood stunting, and poor education make this a significant environmental justice issue.

Although most short-term illness from contaminated drinking water comes from microbial contaminants, chemical pollution of freshwater is another significant issue, often having longer-term impacts on victims following prolonged exposure. These impacts are difficult to quantify because of the long lag between exposure and outcomes, multiple sources of exposure to the same chemicals, and multiple chemicals with similar effects being present in water sources. In addition, only a minority of the more than 140,000 chemicals that have been manufactured since 1950 have been tested for their effects on human and ecological health, either singly or in combination (Landrigan et al., 2017). Chemical contaminants can occur naturally as a result of geological factors (arsenic in the groundwater of Bangladesh, Nepal, and West Bengal is a good example; Argos et al., 2010), or they can be anthropogenic, resulting from industrial or agricultural pollution or from the water treatment and distribution system itself. The WHO (2017) guidelines for drinking water quality include specific guideline values for 8 naturally occurring chemicals, 21 industrial chemicals (of which mercury has the greatest health burden), and agricultural source chemicals (nitrate, nitrite, and 25 agricultural pesticides).

Land Use and Water Engineering

Regional tensions over freshwater supplies are increasing in many locations as population pressures increase and an increasing proportion of agricultural production comes to depend on irrigation (Grimm et al., 2006). Water is essential for household hygiene, communal sanitation, and economic vitality.

During the third quarter of the 20th century, it became clear that large-scale human interventions in the natural environment—dams, intensification of livestock agriculture, irrigation schemes, land reclamation, road construction, and population resettlement programs—often affected infectious disease patterns. In particular, the composition of vector species generally changes following alterations in environmental conditions. Such large-scale developments in the eastern Mediterranean, Africa, South America, and Asia have been consistently associated with increases in vector-borne diseases, especially schistosomiasis and filariasis. In the Sudan in the 1970s, for example, schistosomiasis appeared soon after the start of the Gezira scheme, a large irrigated cotton project; the prevalence of malaria also increased markedly in this region (Fenwick, Cheesmond, & Amin, 1981; Gruenbaum, 1983). In Africa, the building of large dams in the Sudan, Egypt (the Aswan High Dam),

Ghana, and Senegal caused the prevalence of schistosomiasis in the surrounding populace to increase from very low levels to more than 90% (WHO, 1997).

Again, these interactions between land use and infectious disease patterns are not limited to LMICs. As a salutary example, in 2015 New Zealand experienced the largest waterborne campylobacteriosis outbreak in its history, when more than 5,000 people had a notified gastrointestinal illness in the town of Havelock North (population 13,000). The outbreak involved a complex interplay of factors: An extended period of drought was followed by a heavy rainfall event; poor groundwater management left an insecure water bore; and the area had seen increasingly intense development of sheep and dairy farming, with agricultural run-off as the source of contamination (Government Inquiry into Havelock North Drinking Water, 2017).

Integrated Catchment Management

While there has appropriately been a heavy focus on improved sanitation, and on the regulation, monitoring, and treatment of individual hazards in drinking water, it is increasingly being recognized that protecting source waters from contamination is both a crucial and cost-effective step to improve human and ecological health (WHO, 2017). Increasingly, this involves considering and planning the management of water catchments and surrounding land uses together.

Developing integrated catchment and land-use management plans requires a significant shift from top-down, centralized, regulatory approaches to significant intersectoral and *transdisciplinary* effort, bringing policy, community, and scientific citizens together. Behind these more integrative approaches is the understanding that watersheds have a range of influences on health that are wider than the physical health impacts of drinking water contamination, encompassing the social, ecological, economic, and cultural building blocks of health (Parkes, 2016). Often these implications for health are not explicit in integrated catchment planning, but instead remain hidden among discussions about "ecosystem services" (Stosch, Quilliam, Bunnefeld, & Oliver, 2017), "social outcomes" (Bunch et al., 2014), or ecological "conservation," which also must encompass protecting ecological integrity for human health (Horwitz & Finlayson, 2011). These interconnected ways of considering integrated catchment management have been described as facets on a prism (Bunch et al., 2014), as shown in **FIGURE 11-7**. When only a limited number of these perspectives are included in planning, then crucial aspects are likely to

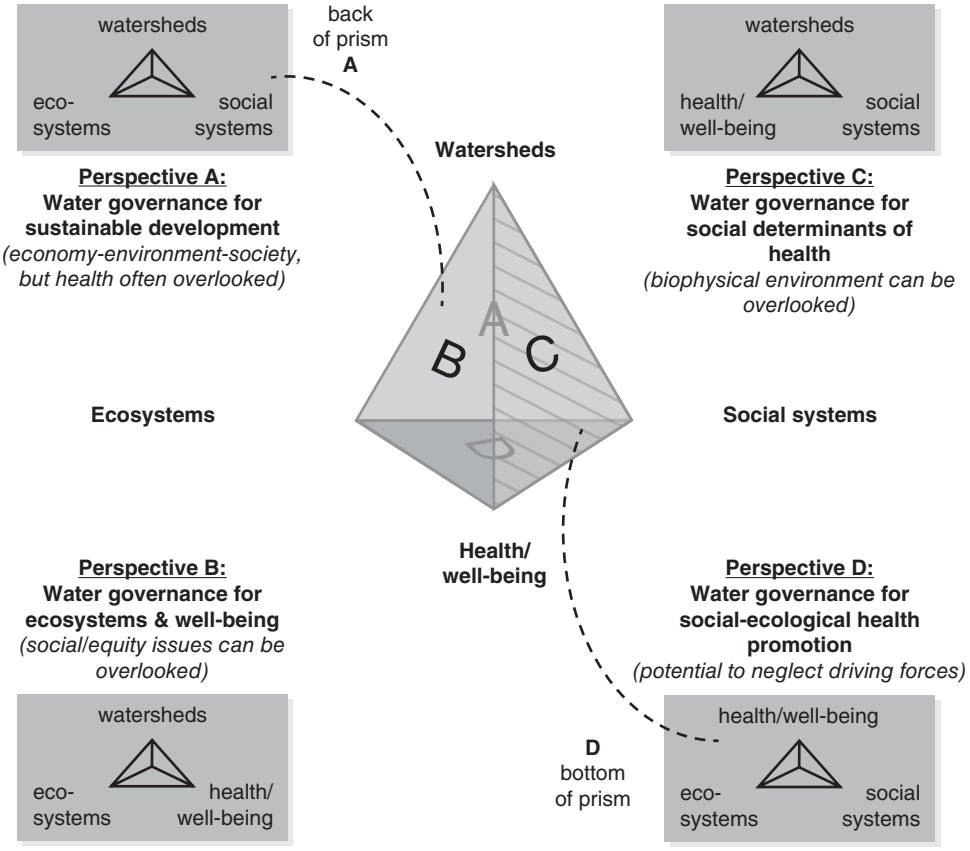


FIGURE 11-7 Watershed governance prism for integrated catchment management.

be overlooked. Unfortunately, examples where health systems, social systems, and ecological systems come together in catchment and land-use planning are generally lacking, despite ongoing recommendations and the development of tools to facilitate such collaboration (Horwitz & Finlayson, 2011; Stosch et al., 2017).

One successful example of integrated catchment and land-use management with the objective of protecting health through drinking water is New York City's Watershed Protection Program (Smith & Porter, 2010). Unfiltered and untreated water from reservoirs in three major catchments supplies New York City's 9 million residents. The water supply has been able to avoid filtration and disinfection requirements through a Watershed Agricultural Program (Watershed Agricultural Council, 2017) voluntarily involving 90% of farmers in the catchments (with funding to minimize agricultural sources of pollution and create run-off barriers), land acquisition, and wider community-based water protection through partnerships between community representatives, local and regional policy makers, and scientists to manage local septic systems, storm water, agriculture and forestry, and riparian planting. As a result of the program, two catchments supplying New York City's drinking water have been able to meet national quality standards without needing to be treated (NYC Environmental Protection, 2017).

Food Production

The world's population continues to grow, projected to reach 9.8 billion people in 2050 and 11.2 billion in 2100 (United Nations, Department of Economic and Social Affairs Population Division, 2017). The overall production of food has so far kept pace with levels of population growth, though it is distributed unevenly among countries. In addition to population growth, economic development, per capita food consumption (particularly of meat and dairy), and the increasing use of food crops for biofuel are currently putting pressure on global food production systems. Environmental factors such as soil, climate, and water supply are all central to the success of food production. Agriculture and fisheries workers are exposed to a wide variety of occupational hazards, many of which were described in the "Occupational Health" section of this chapter.

The "Green Revolution," which has to date allowed food production to increase with population growth, has involved increased mechanization, chemical inputs, monoculture cropping, and livestock intensification. All of these advances have had complex implications for environmental health as a by-product of their inarguable and measurable improvements in overall global nutrition. While industrialization and commodification of agriculture and livestock production have occurred in high-income countries, much of the world's population continues to rely on small-scale subsistence farming. The Green Revolution depended on the introduction of laboratory-bred, high-yield cereal grains, along with fertilizers, groundwater, and arable soils. In retrospect, the productivity gains associated with this agricultural trend appear to have come substantially from using up exhaustible ecological capital—especially topsoil and groundwater. Even as greater food yields to feed ever more people are pursued, almost one-tenth of the world's population remains malnourished in ways that impair health.

Meanwhile, at sea, since the 1970s an increasing proportion of wild-caught marine fish stocks have been overfished. The Food and Agriculture Organization (FAO) of the United Nations estimates that approximately 90% of marine fish stocks are now either overfished or fully fished, with nearly one-third of stocks currently fished at unstainable levels (FAO, Fisheries and Aquaculture Department, 2016).

The increasing length of food supply chains and the greater movement of food between countries have also meant food contamination can be spread rapidly across borders, while investigating and addressing the contamination source has become more difficult. Climate change will both impact global food production (see the *Nutrition* chapter) and require a transition to net zero-carbon food production systems as part of the needed low-carbon society, while continuing to accommodate increasing demand from growing populations. At the same time, it is increasingly being recognized that the intake side of the obesity equation is in major part the result of a complex interplay among physical, social, and policy environments shaping the availability, affordability, attractiveness, advertising, and hidden content of unhealthy and healthy foods (Allen, Prosperi, Cogill, & Flichman, 2014).

In synthesizing these issues, the FAO in 2010 published a scientific consensus definition of sustainable diets:

> [T]hose diets with low environmental impacts which contribute to food and nutrition security and to healthy life for present and future generations. Sustainable diets are protective and respectful of biodiversity and ecosystems, culturally acceptable, accessible, economically fair and affordable; nutritionally adequate, safe and healthy; while optimizing natural and human resources.

Land-Use Change

Increases in the amount of land available for growing crops and livestock in response to population and

consumption pressures have most often come at the expense of natural habitats, inevitably reducing indigenous biodiversity. We currently use half of all global vegetated land for food production (Searchinger et al., 2013). On a worldwide scale, the increased proportion of land used for crops and livestock is the main driver of deforestation, especially in tropical and subtropical regions (FAO, 2017). This trend has had complex and often significant negative effects on local disease patterns, though surprising positive effects have occasionally been noted.

For example, deforestation has had variable effects on malaria mosquito vectors. In some parts of Southeast Asia, deforestation has enabled malaria-transmitting *Anopheles punctulatus* species to become established. In contrast, several *Anopheles* species, including *A. dirus* in Thailand and *A. darlingi* in South America, have disappeared following deforestation that removed the flora and fauna on which they depended for feeding.

In South America, forest clearance during recent decades to extend agricultural land has also mobilized various viral hemorrhagic fevers that previously circulated quietly in wild animal hosts. For example, the Junin virus, which causes Argentine hemorrhagic fever, naturally infects wild rodents. Extensive conversion of grassland to maize cultivation stimulated a proliferation of the reservoir rodent species, which in turn exposed human farmworkers to this "new" virus. In the past 35 years, the land area where this new human disease is endemic has expanded many-fold, with several hundred cases of infection occurring annually, as many as one-third of whom die (Chomel, Belotto, & Meslin, 2007).

Deforestation also disrupts hydrological cycles, most markedly those governing tropical rainfall. At a more local scale, reductions in rainfall and increased probability of drought can undermine the productivity of the newly available farmland and threaten drinking water supplies (Lawrence & Vandecar, 2015). At a much larger scale, the Amazon rainforest, which accounts for 40% of the total global tropical forest area, has undergone rapid land-use change, with more than 20% of its original area lost by 2017. This has had an aggregate regional impact by creating greater divergence in rainfall patterns, with some areas becoming wetter while others become drier (Khanna, Medvigy, Fueglistaler, & Walko, 2017).

Fertilizer Use and the Nitrogen Cycle

Nitrogen is a vital input in food production. Until the beginning of the 20th century, farmers relied on crop rotation, nitrogen-fixing crops, and animal manures to increase the amount of nitrogen in their soil and maintain crop production. The Green Revolution was largely a result of development of the Haber-Bosch nitrogen fixation technology, which allowed ammonia to be produced so that bioactive nitrogen could be applied to soils in the form of synthetic urea fertilizers. Human production and application of nitrogen fertilizers now roughly equals the total natural amount fixed by all non-agricultural organisms on Earth, fundamentally altering the global nitrogen cycle, together with fossil fuel use more generally (Bouwman, Beusen, & Billen, 2009).

Although this has largely been a positive development for human health, with as much as half of global dietary protein now reliant on synthetic fertilizers, there are also significant negative impacts as nitrogen (and phosphorus) is washed into ground and surface waters, including drinking water. Groundwater contamination with nitrates exceeding the WHO guidelines is common as a result of increasing agricultural intensity, which requires concomitant increases in fertilizer use. The most well-established direct health effect is methemoglobinemia, particularly in infants, although further research is needed to understand the long-term health effects of exposure (Villanueva et al., 2014; Ward et al., 2005). Further, multiple indirect pathways between nitrate contamination of freshwater ecosystems and health exist. Increased levels of nitrogen in freshwater increase toxic cyanobacteria blooms (O'Neil, Davis, Burford, & Gobler, 2012), which can also facilitate the rapid growth of *Vibrio cholerae*, precipitating cholera outbreaks (Constantin de Magny et al., 2008). These ecosystem impacts are potentiated by climate change influences on freshwater, including rising water temperatures and changing rainfall patterns.

Microbiological Contamination of Food

Diarrheal disease from contaminated water and food remains one of the world's great global health problems, as discussed earlier in the section on household-related risks. Although this hazard is, in a sense, generated at the household level, failure to initiate community controls can lead to large-scale outbreaks.

An important example is cholera, whose incidence spread worldwide during the 1990s. During that decade, the largest ever cholera pandemic caused cases throughout Asia, the Middle East, Europe (occasionally), Africa (where the disease became endemic for the first time), and Latin America (where it spread widely during the 1990s), affecting more than 1 million people and causing 10,000 deaths. Meanwhile, an apparently new epidemic of cholera was detected in the early 1990s, appearing first in southern India and caused by a new strain (number 0139) of *V. cholerae*. The spread of cholera has been greatly enhanced by the increasing number of slum dwellers in LMICs, the speed and

distance of modern tourism, and an apparent increase in extreme weather events, such as the massive El Niño–associated floods in Kenya in 1997 that caused epidemics of cholera in two regions of the country. In addition, the aftermaths of earthquakes, such as the one in Haiti in 2010, have led to local epidemics of cholera because of poor water and sanitation facilities.

In the villages and slums of LMICs, poor household water quality and sanitation often lead to food contamination. The widespread and unregulated commercial street-food sector in cities offers additional opportunities for exposure. Food contamination remains a concern even in high-income countries, where food is supplied to most of the population via long agriculture, processing, and distribution chains.

The reported rates of food poisoning have increased in high-income countries during the past two decades. The spread of the potentially lethal toxin-producing *Escherichia coli* in North America and Europe in the mid-1990s appears to have accompanied beef imported from infected cattle in Argentina. As long-distance trade expands, as commercial supply lines lengthen in large cities, and as people more frequently opt for convenience or fast foods, the opportunities for foodborne illness increase.

The intensification of meat production is another hotspot for infectious disease problems, as evidenced by three examples: the development of "mad cow disease" in Great Britain and its human counterpart (a variant of Creutzfeldt-Jakob disease); the outbreak of a new strain of influenza in chickens in Hong Kong in 1997; and the surprise appearance in 1999 of the newly named (and often fatal) Nipah virus in intensively produced Malaysian pigs and in several hundred human contacts, many of whom died (Weiss & McMichael, 2004).

Chemical Contamination of Food

Chemical contamination of food derives from three sources: (1) industrial wastes and emissions and household and agricultural chemicals (**EXHIBIT 11-5**); (2) chemicals that form during the storage and handling of food, such as the biotoxin aflatoxin; and (3) natural chemical contaminants. Because changes in land-use patterns are typically accompanied by changes in population density, population mobility, pesticide usage, and regional climate, it is difficult to state with authority any specific causal explanations for chemical contamination of food. Indeed, many of the health outcomes result from interactions between these various change processes.

Several major food contamination episodes in which chemicals were implicated have occurred in Europe. Most recently, in 2008 in China, baby milk formula supplied by a company owned by the New Zealand dairy giant Fonterra was contaminated with melamine, resulting in the deaths of 6 children from kidney disease, and the hospitalization of more than 50,000 children (Ingelfinger, 2008). In 2017, eggs from the Netherlands were contaminated with the insecticide fipronil and then shipped to 15 European Union countries, prompting a recall. These two examples demonstrate the potential for the rapid globalization of such environmental problems in the modern free-trading world.

EXHIBIT 11-5 Learning from Bitter Experience: The Banning of Methylmercury Fungicides

The infamous methylmercury poisoning catastrophe in Minamata, Japan, in 1956 highlighted the severe health problems that could occur after environmental pollution with organic chemicals (ILO, 1990). The same chemical was used in the 1950s and 1960s to prevent fungal growth, which has the potential to impair crop-seed viability and cause mold damage in paper pulp. The use of this fungicide in the paper industry increased the concentrations of mercury in paper manufacturing factories' wastewater, which in turn led to methylmercury contamination of fish. In the late 1960s, Swedish ornithologists had identified both this source and the use of fungicide-treated seeds as a potential cause of infertility in wild birds. A major environmental pollution debate followed, and eventually these fungicides were banned—not only in Sweden, but also in most countries with functioning regulatory systems.

In 1971, Iraq received a large consignment of fungicide-treated wheat seeds from USAID. The shipments arrived when much of rural Iraq was suffering from severe drought, and farming families had little food. The seeds for planting had been dyed red to indicate that they should not be eaten, but the farmers soon found out that washing the seeds in water eliminated the dye. Unfortunately, the fungicide stayed in the seeds. The farmers then used the seeds to make bread, and about 2 months later a major epidemic of serious neurologic diseases began. Eventually 500 people died and approximately 5,000 were hospitalized from eating the tainted seeds.

The Iraqi disaster was the largest known epidemic of this type. Previously some smaller outbreaks had occurred in Africa from the same seed fungicide. After the Iraqi epidemic, a new level of awareness arose about such problems, and the use of this fungicide was subsequently banned in most countries. This intervention represents an example of successful chemical safety management, based on the scare caused by a major epidemic of toxicity.

In the future, addressing the environmental health challenges of the food system while continuing to increase production will require more integrated thinking about land use and farming practices, as well as a global redefinition of healthy diets to incorporate sustainability. Such a dietary shift will need to involve reductions in the consumption of resource-intensive calories and animal-based proteins (especially beef), While realistic targets to increase food production are needed, these must be accompanied by a range of ecological limits and targets—for example, for water pollution, land degradation, deforestation, nitrates use, and greenhouse gas emissions (Hunter, Smith, Schipanski, Atwood, & Mortensen, 2017). Environmental health researchers and practitioners, together with ecologists, have a role in identifying the human health elements of such targets, integrating environmental and health impacts in agricultural and nutrition research, and ensuring the findings are centralized in agriculture and food policies.

Cities and the Built Environment

Across the world, most people spend most of their time inside artificial structures or in other environments that have been deliberately shaped by human design. The importance of the built environment is likely to grow in the future. All around the globe countries are experiencing a spectacular shift of their populations into cities, an environment par excellence that is shaped by human design: The urban population has roughly doubled since 1990, and now includes more than half the people on the planet (UN Habitat, 2016). The damage to health caused by hazards in the built environment is enormous, but so are the gains that can be achieved by wise action. In other words, this aspect of environmental health deserves attention because the stakes are high and the causes of illness and injury are amenable to change.

In the space of half a century, the city has become the predominant human habitat. Indeed, urbanization is one of the defining features of the great acceleration of production, consumption, drawdown on natural resources, and pollution on a global scale (Steffen, Broadgate, Deutsch, Gaffney, & Ludwig, 2015). In official terms, cities are "large settlements," with the size of the settlement that qualifies for city status varying from one jurisdiction to another. In New Zealand, the minimum number is 50,000; the largest Chinese city by the end of the century will contain more people than the present population of Japan.

What defines a city more precisely is the social dynamic. According to Giovanni Botero (1544–1617), the city is "an assembly of people, a congregation drawn together to the end that they may thereby the better live at their ease in wealth and plenty" (Botero & Symcox, 2012). In other words, in cities the whole is greater than the sum of the parts. These are venues of opportunity. Cities thrive according to the density, variety, and connectedness of their populations, and the novel and creative outcomes that emerge when these elements are combined. "Great cities," wrote Jane Jacobs (1916–2006), "are not like towns, only larger. They are not like suburbs, only denser. They differ from towns and suburbs in basic ways, and one of these is that cities are, by definition, full of strangers" (Jacobs, 1961). Cities benefit human health and well-being in many ways, as is evident in the widening gap in the United States and other countries between the life expectancies of urban and rural populations; in the United States in 2005–2009, these were 79.1 and 76.7 years, respectively (Singh & Siahpush, 2014). Better access to education, employment, and health care are part of the explanation for this phenomenon, as are the chances available in cities for social mobility and economic advancement.

However, the features of cities that draw people in and help explain their astonishing productivity (in 2025, the largest 600 cities are projected to generate more than 60% of global GDP) also warrant health warnings. More people closer together, for instance, offers opportunities for the emergence and spread of infectious disease. What is important in this context is not only the population density (from the pathogen's point of view, this means the distance to the next host), but also the population size, the turnover, and the proportion of susceptible individuals: Ultimately, what matters most to the pathogen is the distance to the next *non-immune* host.

The history of cities includes the arrival of pathogens, their establishment, and eventually, in some instances, coexistence with human hosts and declining virulence. For instance, measles and rubella require large, close-living populations (in the range of 100,000 to 500,000 people) to be sustained. For this reason, it is unlikely they were endemic (i.e., established without the need for reinfection from outside) before about 5000 BP, when human settlements reached critical size. The first cities enabled economic specialization and innovation, but at the same time "opened up new niches for nimble microbes" (Woodward, 2014). Modern cities, particularly those distinguished by poverty, conflict, and pollution, continue to provide fertile environments for modern diseases such as HIV/AIDS, cholera, dengue, and tuberculosis (TB).

Other products of dense urban environments include noise, air pollution, and road traffic crash injuries; all of these health risk factors are described in detail elsewhere in this chapter. Heat waves are often more threatening in city centers, where temperatures may be higher than in the leafier suburbs and the surrounding countryside, and where the relief of nighttime cooling is lessened (Hajat, O'Connor, & Kosatsky, 2010). The "heat island" effect is not a necessary consequence of cities: It is typically associated with large heat-retaining structures and treeless asphalt expanses. Cities that feature abundant large street trees, green roofs, urban wetlands, and widespread use of reflective building materials are less likely to be affected.

Basic requirements of healthful living such as provision of shelter, clean water, and sanitation are still a challenge in many cities. UN Habitat (2016) estimates that the proportion of the urban population living in slums in LMICs fell from about 39% in 2000 to 30% in 2014, but the absolute number of people in these environments increased from 791 million to approximately 881 million over the same period. Improvements in sanitation and water supplies have been uneven, with lower coverage in sub-Saharan African cities and south Asia. Conflict and war still cause enormous health problems: The fastest-spreading outbreak of cholera in modern history occurred in 2017 in Yemen, due to population displacement, destruction of homes and infrastructure, and failure of health services. Farther afield, the surge in involuntary migration has placed extraordinary pressures on cities in many countries (UN Habitat, 2016).

The health risks of modern cities apply not just to local populations, but worldwide. Cities generate most of the pressures that today jeopardize the sustainability of current human ecology (Rees, 1996). Urbanism provides opportunities to improve environmental efficiency, including economies of scale in heating, lighting, and waste disposal; shared use of resources; and potential for reuse and recycling. Unfortunately, these opportunities have not matched the rate of growth in people and consumption. The highly urbanized Netherlands, for example, consumes resources from a total surface area approximately 20 times larger than its own geographical footprint. In the 1990s, the estimated consumption of resources—wood, paper, fibers, and food (including seafood)—by 29 cities of the Baltic Sea region, and the absorption of their wastes, depended on a total area 200 times greater than the combined area of those cities (Folke, Larsson, & Sweitzer, 1996). In 2016, UN Habitat estimated that cities account for 60% to 80% of energy consumption and approximately 70% of greenhouse emissions worldwide. The scale of these externalities of urbanism is growing, and includes massive contributions to global greenhouse gas accumulation, stratospheric ozone depletion, land degradation, and coastal zone destruction. The air pollution externalities were described earlier in this section.

A good deal of thought is now applied to building healthier places with an eye toward co-benefits. This means moderating urban form and function to achieve local health gains (by reducing particulate pollution, for instance), while simultaneously shrinking the ecological footprint. Urban transport systems offer many opportunities for win–win interventions of this kind (Shaw, Hales, Howden-Chapman, & Edwards, 2014). City planning and land-use regulations may be deployed to cut vehicle traffic, promote positive social interactions, improve access to social services, and increase opportunities for physical activity. Cities that have to some extent achieved these goals include Bogota, New York, and Seoul (Montgomery, 2013). Recent investments in Copenhagen in blue-green infrastructure along the coast are expected to take the edge off heat waves, improve air quality in the city, promote biodiversity, reduce storm run-off, increase local amenities for residents and visitors, and boost tourism (Institute for European Environmental Policy et al., 2016).

Housing

Housing is a basic human necessity, and its lack, or deficiency, affects environmental health in complex, interactive ways (Macmillan et al., 2016). These impacts may be direct, resulting, for instance, from exposures to physical hazards such as smoke from unvented stoves or allergens released from damp, moldy walls. There may also be less direct, but equally important, effects of inadequate housing on mental health, social functioning, and educational opportunities. In many cities, the high cost of housing is one of the most common causes of financial hardship, with negative consequences for health.

At a basic level, healthy housing depends on ventilation, drainage, lack of crowding, materials that resist pests, design that limits injuries, and insulation from sun, wind, cold, and heat. All of these factors have a significant influence on health. Indeed, much of the improvement in health that occurred in western Europe and North America during the latter part of the 19th century and early part of the 20th century may be attributed to improvements in what have been called "dreadful enclosures"—the urban slums built early in the industrial era to provide for the factory workforce (Preston, 1991).

Because the home is where much human activity takes place, the potential for damaging exposures in households is high if pollutants are present. Two of the most fundamental and mundane human household activities—defecation and cooking—produce significant volumes of potentially health-damaging waste products. When human waste is not removed completely, there is a risk of contamination and outbreaks of diarrhea and other waterborne diseases. As described in the "Air Pollution" section, when the smoke from cooking fires fueled with wood, coal, and other low-quality fuels is released indoors, this pollution may lead to respiratory diseases and other adverse health effects. Indeed, collectively these two environmental health factors account for a substantial fraction of the total environmental burden of ill health globally (Prüss-Üstün et al., 2016).

In some parts of the world, housing conditions substantially influence the risk of infectious diseases spread by vectors. In tropical countries, residences should be situated so that they avoid breeding sites for insect vectors of disease. Good drainage around the home and elimination of any sites where water can stagnate and mosquitoes can breed are of great importance in preventing malaria and dengue fever. In parts of Latin America, the prevention of Chagas disease (South American sleeping sickness) and the elimination of its vector, the triatomine bug, depend on the use of solid ceiling and wall materials without cracks. In the Pacific islands, leptospirosis has become a recurrent health problem among populations in informal settlements where rats are prevalent.

A poorly designed or maintained home may not provide sufficient shelter from extreme weather, protection from neighborhood environmental hazards, or safety in the case of natural disasters. A large literature describes the links between poor housing and ill health, but fewer reports have discussed successful interventions. Nevertheless, a small number of community trials have shown that housing improvements lead to a demonstrable gain in the health of inhabitants. These interventions include removal of injury hazards, installation of clean heating and cooking systems, insulation of walls and ceilings, removal of disease vector breeding sites, and establishment of secure safe water supplies and effective sanitation (Thomson, Thomas, Sellstrom, & Petticrew, 2013).

Transport

The transport system is an important element of the built environment because it has direct impacts on health, arrangements for transport dominate the urban landscape in health-damaging ways, and these systems are heavy users of energy, mostly in the form of fossil fuels. As a result, transport systems contribute substantially to global perturbations such as climate change.

The direct health risks of transport systems have been detailed in a review for Sweden by Kjellstrom, Ferguson, and Taylor (2009). Most obvious, on a global scale, is the large, and increasing, burden of road traffic injuries—which is why halving road traffic injury deaths is a Sustainable Development Goal target (Target 3.6) (United Nations, 2015b). According to the latest Global Burden of Disease analyses, almost 1.5 million people die each year from road traffic injuries, and such injuries are a leading cause of death among young adults (15–29 years) (Naghavi et al., 2017).

Approximately 90% of the road traffic injury fatalities occur in LMICs, even though these countries have only half the world's vehicles (WHO, 2015). Rates per 1,000 vehicles are approximately 30 times higher in Africa than in Norway, the United Kingdom, and the United States. These differences highlight the injury risks at an early stage of motorization, when roads are undeveloped, a safety culture has not taken hold in government, vehicles are frequently old and relatively poorly maintained, vulnerable road users such as pedestrians and cyclists are more numerous and not protected from fast-moving motor vehicle traffic, laws such as compulsory seat-belt use and drink driving are seldom enforced, and emergency services are rudimentary (see the *Unintentional Injuries and Violence* chapter).

Noise is a significant environmental health issue in many parts of the world, and motor vehicle traffic is the most common cause of noise pollution. Approximately 100 million people in Europe are exposed to noise above 55 dB, according to an EU report (European Environment Agency, 2017). Harmful effects arise from the stress of loud and frequent noise, occurring at any hour of the day, and links have been established between noise pollution and sleep disturbances, high blood pressure, weight gain, cardiovascular disease, and premature mortality.

The proliferation of roads and highways can also disrupt social interactions within communities. Unless town planning attends to the needs of pedestrians, cars and roadways tend to dominate the built-up landscape. In Great Britain, for example, the proportion of primary school children walking to school has declined dramatically over the past two decades from a clear majority to a shrinking minority, as traffic has become more intense and walking along roads less safe.

Transport systems that favor the private motor vehicle reduce the opportunities for physical activity. This is an important aspect of the "obesogenic environment," contributing to the global epidemic of obesity and associated disease problems (Kjellstrom, van Kerkhoff, Bammer, & McMichael, 2003). Changing the built environment to promote active transport (walking and cycling) may reduce vehicle use, increase healthy levels of physical activity, and reduce emissions of greenhouse gases (Kelly et al., 2017).

It is challenging to identify effective policies to achieve increases in active transport, and to demonstrate such effects in cities that are dominated entirely by cars and trucks and roads. Nevertheless, interventions have sought to modify road spaces to protect walkers and cyclists, provide alternative routes for active transport away from cars, and discourage use of private motor vehicles (for instance, through congestion charges). In turn, these programs have been linked with favorable changes in modes of transport and risk of disease and injury (Panter, Heinen, Mackett, & Ogilvie, 2016; Smith, M. et al., 2017). Based on existing evidence, integrated health, social, and environmental effects of specific policies have also been modeled using ecosystem health principles (as outlined earlier in the chapter) and complex systems modeling (Macmillan et al., 2014).

Road building and car parking lead in many cities to large land areas covered in tar-seal, which constitute environmental hazards in two respects: They add to the urban "heat island" effect, and they greatly increase storm water run-off during heavy rains. This latter effect was illustrated by the devastating floods that occurred in Houston, Texas, in August 2017 associated with Hurricane Harvey. Harvey brought heavier rains than seen previously in the United States, but the effects were compounded by lack of permeable surfaces across the city.

Climate Change

From an environmental and occupational health viewpoint, climate factors (temperature, humidity, air movement, and heat radiation) are important physical factors affecting individual and population health (Rosenstock et al., 2005), although it remains difficult to link a particular shift in health parameters to long-term climate change. A major review of the health impacts of climate change in the journal *The Lancet* (Watts et al., 2015) concluded that climate change is a very significant and growing global health risk.

The overall links between underlying factors, specific climate exposures, and the health impacts are shown in the DPSEEA framework (**FIGURE 11-8**) developed by WHO in the 1990s (Corvalán et al., 1995). This framework demonstrates that a broad range of climate factor exposures and health effects are likely to occur, and the overall picture has not changed much since the first detailed report on the health impacts of climate change was published in 1996 (McMichael, Haines, Sloof, & Kovats, 1996).

The United Nations' Intergovernmental Panel on Climate Change (IPCC) concluded in 2013 that there is a significant human influence on global climate (Collins et al., 2013). Further, trends in greenhouse gas emissions will, in the IPCC's estimation, cause an increase in average world temperature of approximately 1.8°C to 4°C over the coming century. Rainfall patterns will also likely change, as will the variability of weather patterns. All of these changes are expected to vary considerably by region. The health effects of climate change are expected to encompass both direct and indirect effects, as well as immediate and delayed effects (Smith, K. R. et al., 2014), and have been organized into primary, secondary, and tertiary effects (Butler & Harley, 2010).

The impact of regional climatic variations on human well-being and health has attracted great research interest in recent years. As a consequence of these investigations, there is new understanding about the regional pattern of storms, floods, cyclones, and droughts in response to these quasi-periodic cycles. Indeed, advance warning of interannual variations in rainfall and drought conditions (with their implications for regional food production and outbreaks of certain infectious diseases) is now becoming possible. The distribution of heat on Earth is indicated in **FIGURE 11-9**. The tropical and subtropical areas are most affected, as would be expected, and climate change will make the heat exposure situation worse.

Figure 11-9 indicates that heat stroke and heat exhaustion are important direct effects of increasing heat exposure. These heat stress impacts are boosted by physical activity, because exercise creates surplus heat inside the body. Thus, working populations are vulnerable to climate change health effects (Kjellstrom, Holmer, & Lemke, 2009). The negative effects of daily heat on daily life activities have not been sufficiently analyzed to date, but given that they depend on the physiological limits for coping with heat (Parsons, 2014), one can classify these impacts as health effects even though the outcomes occur at the social level. In tropical and subtropical countries, heat stress effects are not only a problem during "heat waves," but are also related to daily levels of typical heat.

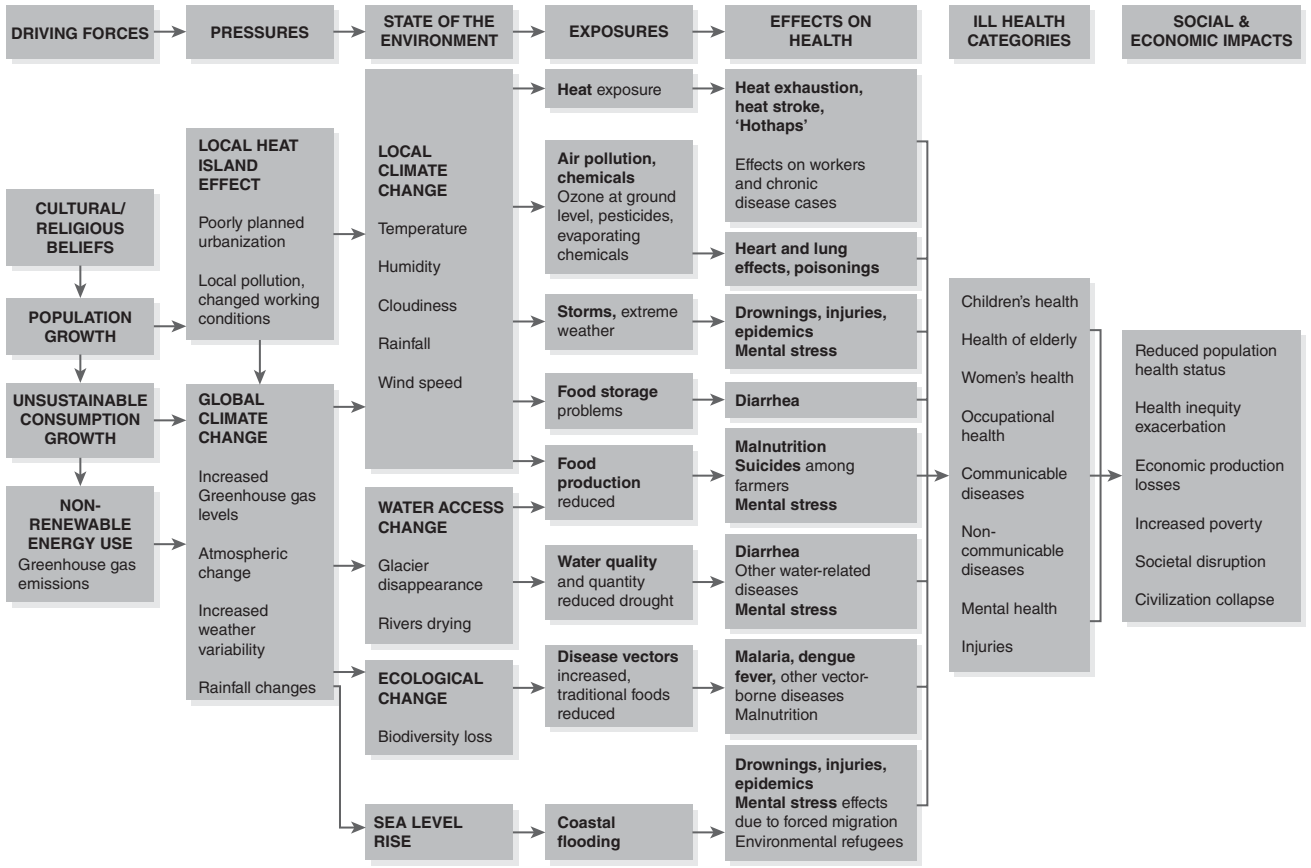


FIGURE 11-8 Using the DPSEEA framework to understand climate change impacts on global health.

Modified from Kjellstrom, T., & McMichael, A. J. (2013). *Climate change threats to population health and well-being: The imperative of protective solutions that will last. Global Health Action;* Corvalán, C., Briggs, D., & Kjellstrom, T. Development of environmental health indicators. In D. Briggs, C. Corvalán, & M. Nurminen (Eds.), *Linkage methods for environment and health analysis* (pp. 125–153). Geneva, Switzerland: World Health Organization. Retrieved from: http://apps.who.int/iris/bitstream/10665/62988/1/WHO_EHG_95.26_eng.pdf

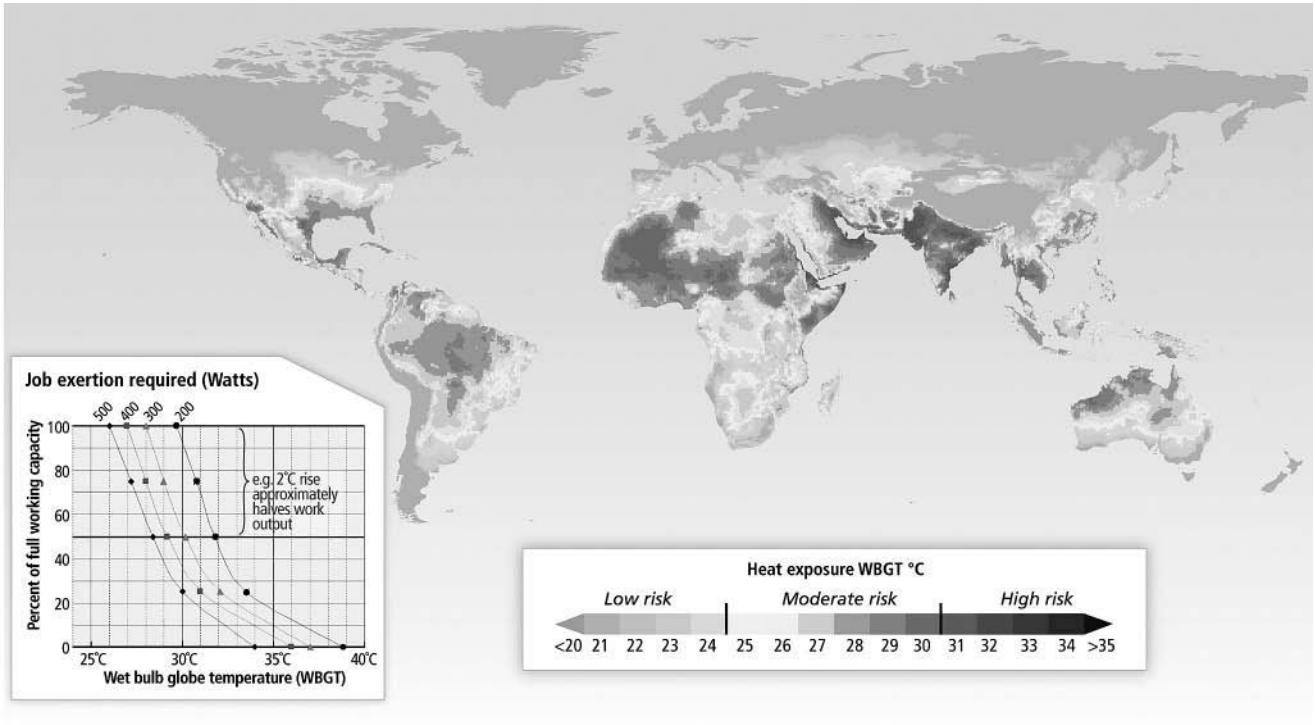


FIGURE 11-9 Hottest month, afternoon heat levels, 1995.

Reproduced from Smith, K. R., et al. (2014). *Human health: Impacts, adaptation, and co-benefits*. In Field , C. B., et al. (Eds.), *Climate change 2014: Impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. (Chapter 11). Cambridge, UK: Cambridge University Press. pp. 709–754.

The predicted global annual mortality due to climate change in 2030 and 2050 (WHO, 2014) is expected to increase in a major way owing to the direct effects of heat on elderly people. By comparison, the other mortality risks (e.g., malnutrition, vector-borne diseases, diarrheal diseases) may be stable or even decrease due to adaptation possibilities. The dramatic effects on mortality during the severe heat wave in Europe in 2003 and subsequent heat waves bear witness to the severe consequences for the health of older people (WHO, 2004).

The increasing temperatures will boost the local production of ozone from nitrogen oxides emerging from motor vehicle exhausts. This may lead to additional health effects of air pollution, as well as the combined effects of heat and air pollutants (Dear et al., 2005). Proposed interventions to mitigate climate change include replacing fossil fuel–burning vehicle engines with cleaner electric engines, in a move that should have important health co-benefits (Haines et al., 2009).

Many indirect health threats emerge from problems with food storage, food production, drinking water quality, disease vectors, sea level rise, impacts on livelihoods, conflict, and migration. The health outcomes of these events in relation to climate change will depend on the application of adaptation methods. As pointed out in the IPCC's assessment of the health threats (Smith, K. R. et al., 2014), the future strengthening of general public health programs and facilities will reduce these threats. This is one reason why the WHO mortality impact assessment predicted reductions of climate change–induced mortality via indirect health effects between 2030 and 2050 (WHO, 2014). The links between climate conditions and vector-borne diseases have attracted substantial attention, but demonstrating the pathways from exposure to effects has proved to be quite complicated (see **EXHIBIT 11-6**). Although certain health outcomes in selected areas would be beneficial—some tropical regions may become too hot for mosquitoes, for example, and winter cold snaps would become milder in temperate-zone countries where death rates typically peak in wintertime—most of the anticipated health effects would be adverse (Smith, K. R. et al., 2014).

Direct health effects from climate change include changes in mortality and morbidity from an altered pattern of exposure to thermal extremes; the respiratory health consequences of increased exposures to photochemical pollutants and aeroallergens; the physical hazards of the increased occurrence of storms, floods, or droughts, in at least some regions; and the increased direct heat exposure on people working in situations without cooling systems (Kjellstrom, Holmer, & Lemke, 2009). Intensified rainfall, with flooding, can overwhelm urban wastewater and sewer systems, leading to contamination of drinking water supplies. This outcome is most likely in large crowded cities where infrastructure is old or inadequate, as illustrated by an outbreak of typhoid in Tajikistan in 1994, when the city's wastewater system flooded during torrential rains.

Indirect health effects are likely to have a greater aggregate impact over time. They may include alterations in the range and activity of vector-borne infectious diseases (e.g., malaria, dengue fever, and leishmaniasis). These diseases are spread by vectors (e.g., mosquitoes) that are sensitive to climatic

EXHIBIT 11-6 Modeling the Future Impacts of Climate Change on Malaria Transmissibility

Mosquitoes—the vector organisms for malaria—are very sensitive to temperature and humidity. Further, patterns of rainfall, river flow, and surface water affect their opportunity to breed. Given these facts, scientists anticipate that a change in world climatic conditions during the 21st century will affect the pattern of potential transmission of malaria (Martens et al., 1999).

Both temperature and humidity affect the growth, biting rate, reproductive cycle, and longevity of the mosquito. Mosquitoes are most comfortable in a temperature range from 20°C to 30°C and at approximately 60% humidity. The malarial parasite, a single-celled sexually reproducing plasmodium, is also affected by temperature during the extrinsic phase of its complex life cycle, when sporozoites are forming within the mosquito.

From studies done during the 1980s and 1990s, it is clear that malaria outbreaks are closely related to interannual climatic variations in many countries. Studies in India, Pakistan, and Sri Lanka all reveal strong associations between malaria outbreaks and El Niño events and the associated monsoonal changes. A correspondingly strong correlation with interannual variations in temperature and rainfall has been reported for outbreaks of dengue fever (a mosquito-borne infectious disease) in Pacific island populations (Hales, Weinstein, & Woodward, 1996).

Because other determinants of malaria occurrence will also inevitably change over coming decades (e.g., vaccine efficacy, human demography, land-use patterns, or parasite resistance to drugs), the modeling of climate change impacts typically indicates only changes in potential transmission.

conditions, as is the parasite's development while incubating in the vector. Scenario-based mathematical modeling has suggested that the geographic zone and seasonality of potential transmission of malaria or dengue fever will increase in many parts of the world in the coming century (McMichael et al., 2003). In temperate Europe and North America, climate-sensitive, vector-borne infections include tick-borne encephalitis and Lyme disease.

Other indirect effects may include altered transmission of person-to-person infections (especially summertime foodborne and waterborne pathogens). Of great potential importance to population health would be the adverse nutritional consequences of the likely regional declines in agricultural productivity (McMichael, Powles, Butler, & Uauy, 2007), which are forecast to reach 10% to 20% by the latter half of the 21st century in many already food-insecure populations in low-latitude regions. Livestock production, particularly of ruminant, digastric animals (sheep, cattle, buffalo, goats, and others), contributes substantial emissions of methane, a potent greenhouse gas, which are adding significantly to ongoing climate change (Friel, Marmot, McMichael, Kjellstrom, & Vagero, 2008; McMichael et al., 2007). Finally, climate change will inevitably result in adverse physical and psychological health consequences from population displacement and economic disruption due to rising sea levels (e.g., small island states, coastal Bangladesh, and the Nile Delta), declines in agroecosystems, and freshwater shortages. Most of these negative health impacts, along with the heat-induced reductions of the work capacity of working people (Kjellstrom, Holmer, & Lemke, 2009), are expected to have their greatest effects in low-income populations (Friel et al., 2008; Patz, Gibbs, Foley, Rogers, & Smith, 2007).

An analysis of different potential health effects enables the identification of gaps in the current presentations in major international reports. **TABLE 11-5** uses heat impacts to demonstrate that a number of effects are suggested by focused studies, but have not yet been quantified at global level. This table could be expanded to all climate effects, but further research is clearly needed to do so.

On the positive side, a growing body of research has identified potentially significant benefits for health and health equity of well-designed actions to

TABLE 11-5 Climate Change–Related Health Impacts of Heat According to IPCC, WHO, and Other Sources

Hazard Exposure	Health Impact	Confidence* of This Impact	Specific Effects at Organ Level	Source
Intense heat	Heat stroke death	Very high	Heart strain; central nervous system (CNS) malfunction; dehydration	IPCC (Smith K. R. et al., 2014) WHO (2014)**
	Heat stroke morbidity	Very high	Heart strain; CNS malfunction; dehydration	IPCC (Smith K. R. et al., 2014)
	Heat exhaustion Work capacity loss	High	Heart strain; mental fatigue	IPCC (Smith K. R. et al., 2014)
Forced migration	Undernutrition; infections; mental stress; injuries	High	Work capacity loss, heart disease, fatigue	IPCC (Smith K. R. et al., 2014)
Health Concerns Not Mentioned in Detail in the IPCC and WHO Reports				
Intense heat	Chronic kidney disease linked to dehydration		Wesseling et al. (2013)	
	Increased incidence of violent crimes		Gamble & Hess (2012)	
	Increased incidence of suicides		Berry et al. (2010); Kim et al. (2016)	

	Teratogenic effects of high body temperature in pregnant women; damage to development of brain	Edwards et al. (1995)
	Interactions with prescription drugs	Vanakoski & Seppala (1998)
	Deteriorated clinical status in chronic noncommunicable diseases	Parsons (2014)
	Increased damage due to head trauma	Titus et al. (2015)

* IPCC assessment judgment of the confidence for climate change impact.
** The WHO (2014) report deals with climate change–related mortality.
Modified from De Blois, J., Kjellstrom, T., Agewall, S., Ezekowitz, J. A., Armstrong, P. W., & Atar, D. The effects of climate change on cardiac health. *Cardiology, 131*(4), 209–217.

TABLE 11-6 Health and Health Equity Co-Benefits of Sector Actions to Mitigate Climate Change

Emissions Reduction Measure (Mitigation)	Health and Equity Co-benefits
Investments in low-carbon public and active transport	Increased physical activity; reduced air pollution; reduced road traffic injuries; improved equity of access to employment, education, and health-promoting goods and services
Housing insulation and clean, dry, low-carbon, energy-efficient heating and cooking	Reduced respiratory infections, cardiovascular disease, and asthma; improved housing affordability for low-income households; reduced lung cancer
Lowering livestock production and consumption of animal products	Reduced cardiovascular disease and cancer (especially bowel cancer); improved freshwater quality; reduced deforestation; improved availability of grains for direct consumption
Phasing out fossil-fuel energy generation, particularly coal	Reduced air pollution; reduced mining and oil-drilling occupational deaths, injuries, and chronic diseases

Modified from Haines, A., McMichael, A. J., Smith, K. R., Roberts, J., Woodcock, J., & Markandya, A., et al. (2009). Public health benefits of strategies to reduce greenhouse-gas emissions: Overview and implications for policy makers. *The Lancet 374:* 2104–2114; Bennett, H., Jones, R., Keating, G., Woodward, A., Hales, S., & Metcalfe, S. (2014). Health and equity impacts of climate change in Aotearoa-New Zealand, and health gains from climate action. *The New Zealand Medical Journal (Online), 127*(1406), 16–31.

mitigate climate change (Haines et al., 2009; Watts et al., 2015). These co-benefits can be achieved across all the major sectors producing greenhouse gas emissions—agriculture and food, energy supply, transport, and housing. Reducing air pollution–related deaths by phasing out coal burning has drawn the largest amount of attention in recent years, because of the potential to save a significant proportion of the 9 million lives lost from this cause each year in LMICs, particularly China and India. **TABLE 11-6** summarizes these co-benefits by sector.

In the last few years, research on the global, regional, and national impacts of climate change on health has intensified. A major effort to monitor the emerging health trends has been initiated by the highly respected medical journal *The Lancet*. Its Countdown project (Watts et al., 2017), which aims to show climate change–related trends from 2000 to the most recent year, includes a large number of indicators (**TABLE 11-7**) covering not only a variety of health impacts, but also adaptation and mitigation actions, as well as economic and policy issues.

The first Countdown report with data has a number of gaps (Watts et al., 2017), but additional material is expected to be included as the project progresses. It is clear from the report that the risks of health impacts are increasing, and students of climate change and health need to look for the latest impact assessments

TABLE 11-7 Indicators for Climate Change and Health from *The Lancet* Countdown

Section 1: Climate change impacts, exposures, and vulnerability
1.1. Health effects of temperature change
1.2. Health effects of heat waves
1.3. Change in labor capacity
1.4. Lethality of weather-related disasters
1.5. Global health trends in climate-sensitive diseases
1.6. Climate-sensitive infectious diseases
1.7. Food security and undernutrition
1.7.1. Vulnerability to undernutrition
1.7.2. Marine primary productivity
1.8. Migration and population displacement

Section 2: Adaptation planning and resilience for health
2.1. National adaptation plans for health
2.2. City-level climate change risk assessments
2.3. Detection and early warning of, preparedness for, and response to health emergencies
2.4. Climate information services for health
2.5. National assessment of vulnerability, impacts, and adaptation for health
2.6. Climate-resilient health infrastructure

Section 3: Mitigation actions and health co-benefits
3.1. Carbon intensity of the energy system
3.2. Coal phase-out
3.3. Zero-carbon emission electricity
3.4. Access to clean energy
3.5. Exposure to ambient air pollution
3.5.1. Exposure to air pollution in cities
3.5.2. Sectoral contributions to air pollution
3.5.3. Premature mortality from ambient air pollution by sector
3.6. Clean fuel use for transport
3.7. Sustainable travel infrastructure and uptake
3.8. Ruminant meat for human consumption
3.9. Healthcare-sector emissions

Section 4: Economics and finance
4.1. Investments in zero-carbon energy and energy efficiency
4.2. Investment in coal capacity
4.3. Funds divested from fossil fuels
4.4. Economic losses due to climate-related extreme events
4.5. Employment in low-carbon and high-carbon industries
4.6. Fossil fuel subsidies
4.7. Coverage and strength of carbon pricing
4.8. Use of carbon pricing revenues
4.9. Spending on adaptation for health and health-related activities
4.10. Health adaptation funding from global climate financing mechanisms

Section 5: Public and political engagement
5.1. Media coverage of health and climate change
5.1.1. Global newspaper reporting on health and climate change
5.1.2. In-depth analysis of newspaper coverage on health and climate change
5.2. Health and climate change in scientific journals
5.3. Health and climate change in the United Nations General Assembly

from reputable sources to learn about the latest evidence.

The Future of Environmental Health in an Unequal World

Sometimes LMICs are criticized for not imposing stricter environmental and occupational standards in tandem with efforts to pursue economic development. There are indeed many situations where minimal efforts could achieve great reductions in health risks. In addition, given the likely increase in public willingness to pay for stricter standards applied in conjunction with economic development, the early imposition of regulations and standards can often be cost-effective using standard economic criteria. Nevertheless, most large LMICs today have standards that are much stricter than those found in high-income countries when they were at the equivalent level of development. Indeed, although cases involving the export of environmental hazard and the degradation of local environments by exploitative and destructive commercial practice can be cited (see "The Export of Hazard," later in this chapter), in the right circumstances the overall impact of globalization can be a net export of positive environmental health. That is, through education, political and public pressures, technology transfer, trade agreements, and other initiatives, LMICs may have, on average, better environmental quality than would otherwise be the case.

Setting Standards: Too Safe Can Be Unsafe

Setting and enforcing environmental standards for protection of the public and workers is a difficult exercise, even in high-income countries that have substantial resources. In LMICs, which face a number of critical needs with few resources, this process is even more challenging. Although official rhetoric about setting standards often states that the only concern is protection of health, in reality there are always economic and other tradeoffs involved. It is too expensive and there are too many other demands on resources to bring every pollutant under the maximum control quickly. Politically, however, it is often difficult for governments in LMICs to set standards that are significantly less stringent than those in high-income countries. As a result, sometimes standards are unrealistically strict and cannot be met or enforced. This disparity between the standards and the reality leads to graft, cynicism, apathy, and, too often, levels of exposures and ill health greater than what could be achieved by adopting more realistic approaches.

One way out of this dilemma is to emphasize environmental health protection and standards setting as dynamic, evolutionary processes rather than as static, one-time-forever efforts. For example, a country might set 20-year goals for its standards that are as strict as any in the world, while establishing interim objectives that become progressively more strict. Thus, pollutants and industries that pose the most risk can be emphasized first, while control over other pollutants and industries (of the kind that now attract attention in high-income countries) can be postponed until a later period. No hazard is ignored with this approach, but merely put into a rational order of priority. Industrial interests are likely to be more willing to accept stricter standards under such a scheme because they tend to have shorter-term financial goals, which are facilitated by clarity of the regulatory environment, whereas they are unsettled by longer-term uncertainty. A recurring problem, however, is attaining sufficient stability within governments to implement such a long-term approach.

The Export of Hazard

In recent years, concern has risen about the potential export of environmental hazards that may be occurring as part of the globalization of the world economy (LaDou & Jeyaratnam, 1994; Leichenko, 2008). Because environmental and occupational standards and enforcement tend to be less strict in LMICs, one might expect that polluting activities would tend to migrate to those countries, thereby imposing excess risk on their workers and public. On the whole, labor costs and tax regimes often appear to be a stronger driver of industry shifts from higher- to lower-income countries. Further, as noted earlier, nations at different stages of development may be juggling multiple tradeoffs between environmental health and occupational standards, job creation, and infrastructure development. The most outrageous examples of injustice arise where decisions about such tradeoffs are made by a small oligarchy without considering the needs and wishes of the population as a whole. In these cases, perhaps the only way to protect the interests of all workers and residents would be to establish international norms.

Some stark examples of problems created by the exporting of hazard between high-income and LMICs can be cited, however. On the Mexican side of the long U.S.–Mexican border, for example, one can find many highly polluting industries with poor occupational safety and health conditions relative to standards on the U.S. side of the border (Moure-Eraso et al., 1994). Because of the proximity of these operations, both the pollution and health problems tend to cross the border, frustrating attempts by U.S. border communities to maintain acceptable conditions. Attempts to impose U.S. standards on Mexican facilities understandably create friction. There have been some encouraging successes in joint efforts by neighboring Mexican–U.S. communities to address these issues in a way that takes into account the need for jobs and for clean working and living conditions on both sides.

The export of plastics and electronics recycling is another example of the export of hazard. China has for many years been the largest importer of wastes for recycling, including from the European Union, the United States, and Japan, with multiple attendant environmental and occupational health risks (Velis, 2014). As a result of the negative environmental and health impacts, in 2017 China notified the World Trade Organization that it would no longer be accepting 24 classes of plastics for recycling.

The political and moral dilemmas associated with industrialization of LMICs are widely recognized. On the one hand, today's high-income world followed a path to development and wealth that put economic gains ahead of human welfare and environmental conservation for most of the 19th century. On the other hand, today's LMICs are dissuaded, if not formally barred, from following the same pathway because (1) their populations have access to more information and hence have higher expectations; (2) the high-income world has a moral obligation to assist them via the transfer of knowledge, wealth, and technology; (3) it is now clear that the integrity of the biosphere at large is jeopardized by the prospect of huge populations in these countries accruing wealth via environmentally damaging behaviors; and (4) there is pressure from public opinion in industrialized countries to achieve higher standards.

Environmental Hazards Resulting from Forms of Globalization

A dominant trend in the global economic environment over the past 50 years has been the rapid growth in international trade in goods, services, and human resources. The globalizing processes of the past quarter century have transformed patterns of connectedness around the world and have created new power relations among countries, international and national governance, and the public and private sectors (see the *Global Health Governance and Diplomacy* chapter).

In traditional agrarian-based societies that produce, consume, and trade on a local basis and with relatively low-impact technologies, the effects of those activities on the environment are predominantly local. Few such societies remain today, in the face of the strong and pervasive economic, technologic, and cultural influences now spanning the globe. The industrialization of the past century and the more recent globalizing processes have altered the scale of contact between societies, intensified environmental impacts, and extended the health impacts of one society on another. In the name of economic development and free markets, LMICs have come under pressure to grant unrestricted access to their resources, workforces, and consumer markets. This process has been associated with increasing poverty in many parts of the world, exacerbating inequalities between and within countries, expanding pressures to reduce the power of the state, and subordinating national programs of social welfare and environmental protection to the agenda of economic growth.

Following the international debt crisis of the early 1980s, many struggling LMICs were obliged to accept the economic stringencies mandated by the World Bank's structural adjustment program. These dictates included a reduction in spending on health throughout the 1980s and into the 1990s. Growing evidence suggests that these structural adjustment policies adversely affected environmental health, including through degradation of freshwater, a shift from mixed agriculture to cash cropping, and rapid, poorly controlled urbanization (Kentikelenis, 2017). In its current form, the world's globalizing economy operates to the general disadvantage of LMICs. The exacerbation of land degradation, rural unemployment, food shortages, and urban crowding all contribute to health deficits for the rural dispossessed, the underfed, and the slum dweller. The health and safety conditions of working people can create particular threats to occupational health (Kjellstrom & Hogstedt, 2009).

In addition, many features of today's globalizing world contribute to the spread of infectious diseases (Weiss & McMichael, 2004). Human mobility has escalated dramatically, in volume and speed, between and within countries in recent decades. Long-distance trade facilitates the geographic redistribution of pests and pathogens—a phenomenon well illustrated in

recent years by the HIV pandemic, the worldwide dispersal of rat-borne hantaviruses, and the rapid dissemination of a new epidemic strain of bacterial meningitis along routes of travel and trade. Likewise, transport on ships has facilitated the introduction of the Asian tiger mosquito, *Aedes albopictus*—a vector for yellow fever and dengue—into South America, North America, and Africa (Morse, 1995) and of the cholera bacterium into South American coastal waters (Colwell, 1996). In an analysis of cholera outbreaks since 1817, Lee and Dodgson (2000) argue that the current (seventh) pandemic is clearly different from earlier ones, reflecting the unprecedented scale of social and environmental change in the world over recent years, the exacerbation of urban poverty, and the rapidity and intensity of intercontinental contacts.

One aspect of globalization that may have negative effects on environmental health is the harmonization of trade-related rules and legislation via the World Trade Organization (WTO). Recent multilateral trade and investment agreements, such as the Trans-Pacific Partnership Agreement, contain clauses that allow multinational corporations to sue governments for loss of profits, including those losses incurred because of governments' new regulations to protect environmental health (Keating, Freeman, Macmillan, Neuwelt, & Monasterio, 2016).

Similar situations may develop with the banning of hazardous products such as asbestos. If, for example, there is no international health guideline banning the use of asbestos, then any country taking a unilateral decision to ban asbestos use would risk trade sanctions from other countries wanting to export asbestos. If the trend in international environmental and occupational health guidelines went toward stricter prevention, this harmonization via trade rules could be good for health and the environment. In reality, the intense lobbying from those commercial groups and countries that would benefit from lax rules makes it likely that the opposite will happen—meaning that compromises veering toward less protection will be made (LaDou & Jeyaratnam, 1994).

Population Health: Index of Social and Environmental Sustainability

Current models of government reflect the compartmentalization of knowledge and policy that grew out of the classic development of scientific disciplines in the 19th century. To deal with a multifaceted world, our predecessors defined various sectors of knowledge, policy, and social action: environment, industry, agriculture, transport, health, social welfare, and education. Subsequently, one of the great lessons to emerge from 20th-century science, which had its origins in the realm of physics (the 1920s debate about quantum mechanics and uncertainty), was that the complexities of the real world require us to think in more integrative ways, across disciplines and topics, elevating holism (or ecological thinking) above mechanistic, reductionist thinking.

It is within this type of integrative framework that population health can be understood as part of the total social experience, a manifestation of how well the social and natural environments are being managed. Population health should be a primary criterion for all social policy making, but particularly in relation to the goal of achieving the "sustainability transition." Health is not just a type of sideshow in the policy arena, but rather is affected by the social, environmental, and (in the longer term) ecological consequences of policies in all sectors. For all these reasons, population health should be an integrating index of social policy across all sectors.

Conclusion

The perceived importance of environmental exposures as health hazards—at local, regional, and global levels—has increased steadily over the past several decades. Currently, scientists estimate that one-fourth to one-third of the global burden of disease and premature death is attributable to ambient (including household) environmental risk factors.

In high-income countries during much of the past four decades, the generally greater ease of measurement of specific exposures relating to individual lifestyle factors (eating, smoking, and sexual behaviors) and the workplace, compared with the more diffuse, lower-concentration exposures in the external environment, resulted in the latter topic area attracting less attention and having lower credibility. More recently, improvements in exposure assessment, the harnessing of time-series analyses, the advent of spatial analytic techniques, the recognition of the legitimacy of population-level analyses, and the extra leverage afforded by molecular biological indices of exposure, susceptibility, and biological damage have all helped reveal the range and extent of ambient environmental risks to health.

Meanwhile, in LMICs, the age-old scourges of diarrheal disease, acute respiratory infections, tuberculosis, and vector-borne infections have remained the dominant health problems. The ascendancy of specific health system interventions for those

problems—sanitation, household hygiene, vaccination, pesticides, and drug treatment— has led to their wider ecological dimensions being somewhat overlooked. Many problems of environmental contamination have their origins in poverty; deficient regulation of mining, industry, and agriculture; and mismanagement of surface water and groundwater supplies. Differences in household exposure to indoor air pollution reflect the traditional division of labor (women are mostly exposed), the ongoing use of low-grade technology, and the persistence of the biomass fuels of poverty. Infectious diseases are often spread by environmental encroachments—land clearing, water damming, irrigation, and expanded trade.

The environmental health agenda is widening. Today, the burden of the growing human population and its aggregate consumption and waste generation is beginning to overload various aspects of the planet's great natural systems. The resultant global environmental changes, signifying that the biosphere's human population carrying capacity is being exceeded, pose further risks to human health. Therefore, even as environmental health scientists strive to improve their research methods for characterizing the health risks associated with local physical, chemical, and microbiological hazards, they must also extend their ideas and methods to encompass larger-scale environmental hazards and the health consequences of disrupted ecosystems. Policy makers, in many sectors, must understand the tendency of human-wrought changes in the social, built, and natural environments to affect health—if not immediately, then in the longer term, and sometimes via pathways with which we yet have little familiarity.

Discussion Questions

1. What should be the scope of the term "environment"?
2. How do conceptualization and methodological approaches change with a shift from thinking about environmental hazards to thinking about human habitats?
3. Which methodological problems are particularly characteristic of environmental epidemiology?
4. How might differences be explained in the profile of environmental health problems between high-income countries and LMICs?
5. What are the characteristics of particular environmental health problems that render them more, or less, tractable to amelioration or elimination?

Acknowledgments

We wish to thank the four Master of Public Health students at the University of California, Berkeley, who kindly reviewed the previous edition of this chapter: Miriam Magaña Lopez, Maximo Prescott, Luciana Rocha, and Morgan Westfall. Their constructive comments assisted with improvements.

References

Agyeman, J., Schlosberg, D., Craven, L., & Matthews, C. (2016). Trends and directions in environmental justice: From inequity to everyday life, community, and just sustainabilities. *Annual Review of Environment and Resources, 41*(1), 321–340. doi:10.1146/annurev-environ-110615-090052

Allen, T., Prosperi, P., Cogill, B., & Flichman, G. (2014). Agricultural biodiversity, social–ecological systems and sustainable diets. *Proceedings of the Nutrition Society, 73*(4), 498–508. doi:10.1017/S002966511400069x

Amegah, A. K., Quansah, R., & Jaakkola, J. J. (2014). Household air pollution from solid fuel use and risk of adverse pregnancy outcomes: A systematic review and meta-analysis of the empirical evidence. *PLoS One, 9*(12), e113920. doi:10.1371/journal.pone.0113920

Arcury, T. A., Grzywacz, J. G., Sidebottom, J., & Wiggins, M. F. (2013). Overview of immigrant worker occupational health and safety for the agriculture, forestry, and fishing (AgFF) sector in the southeastern United States. *American Journal of Industrial Medicine, 56*(8), 911–924. doi:10.1002/ajim.22173

Argos, M., Kalra, T., Rathouz, P. J., Chen, Y., Pierce, B., Parvez, F., . . . Ahsan, H. (2010). Arsenic exposure from drinking water, and all-cause and chronic-disease mortalities in Bangladesh (HEALS): A prospective cohort study. *The Lancet, 376*(9737), 252–258. doi:10.1016/S0140-6736(10)60481-3

Armstrong, J., & Menon, R. (1998). Mining and quarrying. In J. M. Stellman (Ed.), *Encyclopaedia of occupational health and safety* (4th ed., Vol. III, pp. 74.2–74.55). Geneva, Switzerland: International Labor Office.

Baghurst, P. A., McMichael, A. J., Wigg, N. R., Vimpani, G. V., Robertson, E. F., Roberts, R. J., & Tong, S. L. (1992). Environmental exposure to lead and children's intelligence at the age of seven years: The Port Pirie Cohort Study. *New England Journal of Medicine, 327*(18), 1279–1284. doi: 10.1056/NEJM199210293271805

Baker, D., & Nieuwenhuijsen, M. (2008). *Environmental epidemiology: Study methods and application*. Oxford, UK: Oxford University Press.

Bennett, H., Jones, R., Keating, G., Woodward, A., Hales, S., & Metcalfe, S. (2014). Health and equity impacts of climate change in Aotearoa-New Zealand, and health gains from climate action. *The New Zealand Medical Journal (Online), 127*(1406), 16–31.

Berry, H. L., Bowen, K., & Kjellstrom, T. (2010). Climate change and mental health: A causal pathways framework. *International*

Journal of Public Health, 55(2), 123–132. doi:10.1007/s00038-009-0112-0

Bhalla, K., Shotten, M., Cohen, A., Brauer, M., Shahraz, S., Burnett, R., . . . Murray, C. J. L. (2014). *Transport for health: The global burden of disease from motorized road transport*. Washington, DC: World Bank Group.

Blakely, T. A., & Woodward, A. J. (2000). Ecological effects in multi-level studies. *Journal of Epidemiology and Community Health, 54*(5), 367. doi:10.1136/jech.54.5.367

Bonita, R., Beaglehole, R., & Kjellstrom, T. (2007). *Basic epidemiology* (2nd ed.). Geneva, Switzerland: World Health Organization.

Botero, G., & Symcox, G. T. (2012). *On the causes of the greatness and magnificence of cities*. Toronto, Canada: University of Toronto Press.

Bouwman, A. F., Beusen, A. H. W., & Billen, G. (2009). Human alteration of the global nitrogen and phosphorus soil balances for the period 1970–2050. *Global Biogeochemical Cycles, 23*(GB0A04), 1–6236. doi:10.1029/2009GB003576

Bråbäck, L., & Forsberg, B. (2009). Does traffic exhaust contribute to the development of asthma and allergic sensitization in children: Findings from recent cohort studies. *Environmental Health, 8*, 17. doi:10.1186/1476-069X-8-17

Briggs, D. J. (2008). A framework for integrated environmental health impact assessment of systemic risks. *Environmental Health, 7*(1), 61. doi:10.1186/1476-069x-7-61

Brown, H. L., Proust, K., Spickett, J., & Capon, A. (2011). The potential role of health impact assessment in tackling the complexity of climate change adaptation for health. *Health Promotion Journal of Australia, 22*, S48–S53.

Bunch, M., Parkes, M., Zubrycki, K., Venema, H., Hallstrom, L., Neudorffer, C., . . . Morrison, K. (2014). Watershed management and public health: An exploration of the intersection of two fields as reported in the literature from 2000 to 2010. *Environmental Management, 54*(2), 240–254. doi:10.1007/s00267-014-0301-3

Butler, C. D., & Harley, D. (2010). Primary, secondary and tertiary effects of eco-climatic change: The medical response. *Postgraduate Medicine Journal, 86*(1014), 230–234. doi:10.1136/pgmj.2009.082727

Cai, S., Yue, L., Shang, Q., & Nordberg, G. (1995). Cadmium exposure among residents in an area contaminated by irrigation water in China. *Bulletin of the World Health Organization, 73*(3), 359.

Ceballos, G., Ehrlich, P. R., & Dirzo, R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences USA, 114*(30), E6089–E6096. doi: 10.1073/pnas.1704949114

Chafe, Z. A., Brauer, M., Klimont, Z., Van Dingenen, R., Mehta, S., Rao, S., . . . Smith, K. R. (2014). Household cooking with solid fuels contributes to ambient $PM_{2.5}$ air pollution and the burden of disease. *Environmental Health Perspectives, 122*(12), 1314–1320. doi: 10.1289/ehp.1206340

Charron, D. F. (2012). Ecohealth: Origins and approach. In D. F. Charron (Ed.), *Ecohealth research in practice: Innovative applications* (pp. 1–32). Ottawa, Canada: International Development Research Centre.

Chen, R., Yin, P., Meng, X., Liu, C., Wang, L., Xu, X., . . . Zhou, M. (2017). Fine particulate air pollution and daily mortality: A nationwide analysis in 272 Chinese cities. *American Journal of Respiratory and Critical Care Medicine, 196*(1), 73–81. doi:10.1164/rccm.201609-1862OC

Chivian, E., & Bernstein, A. (2008). *Sustaining life*. Oxford, UK: Oxford University Press.

Chomel, B. B., Belotto, A., & Meslin, F.-X. (2007). Wildlife, exotic pets, and emerging zoonoses. *Emerging Infectious Diseases, 13*(1), 6. doi:10.3201/eid1301.060480

Cohen, S. (2017, 23 October). The human and financial cost of pollution. *Huffington Post*. Retrieved from https://www.huffingtonpost.com/entry/the-human-and-financial-cost-of-pollution_us_59edde97e4b02c6e3c609c80

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., . . . Tett, S. (2013). Long-term climate change: Projections, commitments and irreversibility. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, . . . P. M. Midgley (Eds.), *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 1029–1136). Cambridge, UK: Cambridge University Press.

Colwell, R. R. (1996). Global climate and infectious disease: The cholera paradigm. *Science, 274*(5295), 2025–2031.

Constantin de Magny, G., Murtugudde, R., Sapiano, M. R. P., Nizam, A., Brown, C. W., Busalacchi, A. J., . . . Colwell, R. R. (2008). Environmental signatures associated with cholera epidemics. *Proceedings of the National Academy of Sciences, 105*(46), 17676–17681. doi: 10.1073/pnas.0809654105

Corvalán, C., Briggs, D., & Kjellstrom, T. (1995). Development of environmental health indicators. In D. Briggs, C. Corvalán & M. Nurminen (Eds.), *Linkage methods for environment and health analysis* (pp. 125–153). Geneva, Switzerland: World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/10665/62988/1/WHO_EHG_95.26_eng.pdf

Confalonieri, U., Menne, B., Akhtar, R., Ebi, K. L., Hauengue, M., Kovats, S., . . . Woodward, A. (2007). *Human health. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press.

Davis, L. W. (2017). Saturday driving restrictions fail to improve air quality in Mexico City. *Scientific Reports, 7*, 41652. doi:10.1038/srep41652

De Blois, J., Kjellstrom, T., Agewall, S., Ezekowitz, J. A., Armstrong, P. W., & Atar, D. (2015). The effects of climate change on cardiac health. *Cardiology, 131*(4), 209–217. doi:10.1159/000398787

Dear, K., Ranmuthugala, G., Kjellstrom, T., Skinner, C., & Hanigan, I. (2005). Effects of temperature and ozone on daily mortality during the August 2003 heat wave in France. *Archives of Environmental and Occupational Health, 60*(4), 205–212. doi:10.3200/AEOH.60.4.205-212

Deichmann, W., Henschler, D., Holmstedt, B., & Keil, G. (1986). What is there that is not poison? A study of the Third Defense by Paracelsus. *Archives of Toxicology, 58*(4), 207–213. doi:10.1007/BF00297107

Delin, L., & Shilong, H. (2016). Ecosystem health assessment at county-scale using the pressure–state–response framework on the Loess Plateau, China. *International Journal of Environmental Research and Public Health, 14*(1), 2. doi:10.3390/ijerph14010002

Diamanti-Kandarakis, E., Bourguignon, J. P., Giudice, L. C., Hauser, R., Prins, G. S., Soto, A. M., . . . Gore, A. C. (2009). Endocrine-disrupting chemicals: An Endocrine Society scientific statement. *Endocrine Reviews, 30*(4), 293–342. doi:10.1210/er.2009-0002

Dix-Cooper, L., Eskenazi, B., Romero, C., Balmes, J., & Smith, K. R. (2012). Neurodevelopmental performance among school

age children in rural Guatemala is associated with prenatal and postnatal exposure to carbon monoxide, a marker for exposure to woodsmoke. *Neurotoxicology, 33*(2), 246–254. doi:10.1016/j.neuro.2011.09.004

Edwards, M. J., Shiota, K., Smith, M. S., & Walsh, D. A. (1995). Hyperthermia and birth defects. *Reproductive Toxicology, 9*(5), 411–425.

Eisler, R. (2003). Health risks of gold miners: A synoptic review. *Environmental Geochemistry and Health, 25*(3), 325–345.

Elgstrand, K., & Petersson, N. F. (2009). *OSH for development: Occupational safety and health for development*. Stockholm: Stockholm: KTH Royal Institute of Technology.

European Environment Agency. (2017). *Road traffic remains biggest source of noise pollution in Europe.* Copenhagen, Denmark: Author.

Eze, I. C., Hemkens, L. G., Bucher, H. C., Hoffmann, B., Schindler, C., Kunzli, N., . . . Probst-Hensch, N. M. (2015). Association between ambient air pollution and diabetes mellitus in Europe and North America: Systematic review and meta-analysis. *Environmental Health Perspectives, 123*(5), 381–389. doi:10.1289/ehp.1307823

Fawell, J., Bailey, K., Chilton, J., Dahi, E., Fewtrell, L., & Magara, Y. (2006). *Flouride in drinking-water*. London, UK: IWA Publishing.

Fenwick, A., Cheesmond, A. K., & Amin, M. A. (1981). The role of field irrigation canals in the transmission of Schistosoma mansoni in the Gezira Scheme, Sudan. *Bulletin of the World Health Organization, 59*(5), 777.

Fletcher, T., & McMichael, A. J. (Eds.). (1997). *Health at the crossroads: Transport policy and urban health*. London School of Hygiene and Tropical Medicine Fifth Annual Public Health Forum. Chichester, UK: John Wiley & Sons Ltd.

Folke, C., Larsson, J., & Sweitzer, J. (1996). Renewable resource appropriation by cities. *Getting Down to Earth*, 201–221.

Food and Agriculture Organization (FAO), United Nations. (2017). *The future of food and agriculture: Trends and challenges.* Rome, Italy: Author.

Food and Agriculture Organization (FAO), United Nations, Fisheries and Aquaculture Department. (2016). *State of world fisheries and aquaculture 2016: Contributing to food security and nutrition for all.* Rome, Italy: Author.

Friberg, L., Elinder, C.-G., Kjellstrom, T., & Nordberg, G. F. (Eds.). (1986). *Cadmium and health* (Vol. 2). Boca Raton, FL: CRC Press.

Friel, S., Marmot, M., McMichael, A. J., Kjellstrom, T., & Vagero, D. (2008). Global health equity and climate stabilisation: A common agenda. *The Lancet, 372*(9650), 1677–1683. doi:10.1016/S0140-6736(08)61692-X

Gakidou, E., Afshin, A., Abajobir, A. A., Abate, K. H., Abbafati, C., Abbas, K. M., . . . Murray, C. J. L. (2017). Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990; 2013; 2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet, 390*(10100), 1345–1422. doi:10.1016/S0140-6736(17)32366-8

Gallagher, R. P., & Lee, T. K. (2006). Adverse effects of ultraviolet radiation: A brief review. *Progress in Biophysics and Molecular Biology, 92*(1), 119–131. doi:10.1016/j.pbiomolbio.2006.02.011

Gamble, J. L., & Hess, J. J. (2012). Temperature and violent crime in Dallas, Texas: Relationships and implications of climate change. *Western Journal of Emergency Medicine, 13*(3), 239–246. doi:10.5811/westjem.2012.3.11746

GBD MAPS Working Group. (2016). *Burden of disease attributable to coal-burning and other major sources of air pollution in China. Special Report 20.* Boston, MA: Health Effects Institute.

Government Inquiry into Havelock North Drinking Water. (2017). Report of the Havelock North drinking water inquiry: Stage 1. Auckland, New Zealand: New Zealand Department of Internal Affairs.

Griffin, M. J., Bovenzi, M., & Nelson, C. M. (2003). Dose–response patterns for vibration-induced white finger. *Occupational and Environmental Medicine, 60*(1), 16. doi:10.1136/oem.60.1.16

Griggs, D. (2013). Sustainable development goals for people and planet: Planetary stability must be integrated with United Nations targets to fight poverty and secure human well-being [Comment]. *Nature, 495*(7441), 305.

Grimm, M., Harttgen, K., Klasen, S., & Misselhorn, M. (2006). *Human development report 2006*. New York: United Nations Development Programme.

Gruenbaum, E. (1983). Struggling with the mosquito: Malaria policy and agricultural development in Sudan. *Medical Anthropology, 7*(2), 51–62. doi:10.1080/01459740.1983.9987034

Guarnieri, M., & Balmes, J. R. (2014). Outdoor air pollution and asthma. *The Lancet, 383*(9928), 1581–1592. doi:10.1016/S0140-6736(14)60617-6

Gunnell, D., Eddleston, M., Phillips, M. R., & Konradsen, F. (2007). The global distribution of fatal pesticide self-poisoning: Systematic review. *BMC Public Health, 7*, 357. doi:10.1186/1471-2458-7-357

Haines, A., McMichael, A. J., Smith, K. R., Roberts, I., Woodcock, J., Markandya, A., . . . Wilkinson, P. (2009). Public health benefits of strategies to reduce greenhouse-gas emissions: Overview and implications for policy makers. *The Lancet, 374*, 2104–2114. doi:10.1016/S0140-6736(09)61759-1

Hajat, S., O'Connor, M., & Kosatsky, T. (2010). Health effects of hot weather: From awareness of risk factors to effective health protection. *The Lancet, 375*(9717), 856–863. doi:10.1016/S0140-6736(09)61711-6

Hales, S., Weinstein, P., & Woodward, A. (1996). Dengue fever epidemics in the South Pacific: Driven by El Nino Southern Oscillation? *The Lancet, 348*(9042), 1664–1665. doi:10.1016/S0140-6736(05)65737-6

Hanna-Attisha, M., LaChance, J., Sadler, R. C., & Champney Schnepp, A. (2016). Elevated blood lead levels in children associated with the Flint drinking water crisis: A spatial analysis of risk and public health response. *American Journal of Public Health, 106*(2), 283–290. doi:10.2105/AJPH.2015.303003

Health Effects Institute. (2017). State of global air 2017. *Global Burden of Disease Study 2015.* Institute for Health Metrics and Evaluation, 2016. Retrieved from https://www.stateofglobalair.org/

Herzstein, J. A., Bunn, W. B., Fleming, L. E., Harrington, J. M., Jeyaratnam, J., & Gardner, I. R. (1998). *International occupational and environmental medicine*. St. Louis, MO: Mosby.

Horton, R., & Lo, S. (2015). Planetary health: A new science for exceptional action. *The Lancet, 386*(10007), 1921–1922. doi:10.1016/S0140-6736(15)61038-8

Horwitz, P., & Finlayson, C. M. (2011). Wetlands as settings for human health: Incorporating ecosystem services and health impact assessment into water resource management. *BioScience, 61*(9), 678–688. doi:10.1525/bio.2011.61.9.6

Hosgood, H. D., 3rd, Wei, H., Sapkota, A., Choudhury, I., Bruce, N., Smith, K. R., . . . Lan, Q. (2011). Household coal use and

lung cancer: Systematic review and meta-analysis of case-control studies, with an emphasis on geographic variation. *International Journal of Epidemiology, 40*(3), 719–728. doi:10.1093/ije/dyq259

Hunter, M. C., Smith, R. G., Schipanski, M. E., Atwood, L. W., & Mortensen, D. A. (2017). Agriculture in 2050: Recalibrating targets for sustainable intensification. *BioScience, 67*(4), 386–391. doi:10.1093/biosci/bix010

Ingelfinger, J. R. (2008). Melamine and the global implications of food contamination. *New England Journal of Medicine, 359*(26), 2745–2748. doi:10.1056/NEJMp0808410

Institute for European Environmental Policy, Collingwood Environmental Planning, ESP, ICLEI Local Government for Sustainability, Natural Resources Institute Finland (Luke), Milieu, & Worldwide Fund for Nature (WWF). (2016). *The health and social benefits for nature and biodiversity protection: Background for a workshop by the European Commission and hosted by the Committee of the Regions (27–28 January 2016)*. London, UK/Brussels, Belgium: Authors.

Institute for Health Metrics and Evaluation (IHME). (2016). GBD Compare data visualization. Retrieved from https://vizhub.healthdata.org/gbd-compare/

International Energy Agency. (2017). *Key world energy statistics*. Paris, France: Author.

International Labor Organization (ILO). (2012). *ILO encyclopaedia of occupational health and safety*. Geneva, Switzerland: Author.

International Labor Organization (ILO), World Health Organization (WHO), International Program on Chemical Safety, & United Nations Environment Programme. (1990). *Methylmercury*. Geneva, Switzerland: WHO.

Jacobs, J. (1961). *The death and life of great American cities*. New York, NY: Random House.

Jatikusumah, A., Darmawan, R. D., Widyastuti, M. D. W., Basuno, S. E., Putra, A. A. G., Willyanto, I., . . . Gilbert, J. (2013). *Community-based model for rabies control in Bali: An EcoHealth approach*. Paper presented at the 10th Anniversary of Veterinary Public Health Centre for Asia Pacific 2–6 July 2013, Chiang Mai, Thailand.

Jennings, N. S. (1998). Mining: An overview. In J. M. Stellman (Ed.), *Encyclopaedia of occupational health and safety* (4th ed., Vol. III, pp. 74.72–74.74). Geneva, Switzerland: International Labor Office.

Josyula, S., Lin, J., Xue, X., Rothman, N., Lan, Q., Rohan, T. E., & Hosgood, H. D., 3rd. (2015). Household air pollution and cancers other than lung: A meta-analysis. *Environmental Health, 14*, 24. doi:10.1186/s12940-015-0001-3

Kameda, T., Takahashi, K., Kim, R., Jiang, Y., Movahed, M., Park, E.-K., & Rantanen, J. (2014). Asbestos: Use, bans and disease burden in Europe/Amiante: Utilisation, interdictions et charge de morbidite en Europe/Amianto: Uso, prohibiciones y carga de morbilidad en Europa [Research]. *Bulletin of the World Health Organization, 92*(11), 790. doi:10.2471/BLT.13.132118

Kamp, D. W. (2009). Asbestos-induced lung diseases: An update. *Translational Rsearch, 153*(4), 143–152. doi:10.1016/j.trsl.2009.01.004

Kazan-Allen, L. (2003). The asbestos war. *International Journal of Occupational and Environmental Health, 9*(3), 173–193. doi:10.1179/oeh.2003.9.3.173

Keating, G., Freeman, J., Macmillan, A., Neuwelt, P., & Monasterio, E. (2016). TPPA should not be adopted without a full, independent health assessment. *New Zealand Medical Journal, 129*, 713.

Kelly, M. C., Jeremy, J. H., John, M. B., Jonathan, J. B., David, A. C., Maggie, L. G., . . . Kristie, L. E. (2017). Ancillary health effects of climate mitigation scenarios as drivers of policy uptake: A review of air quality, transportation and diet co-benefits modeling studies. *Environmental Research Letters, 12*(11), 113001.

Kentikelenis, A. E. (2017). Structural adjustment and health: A conceptual framework and evidence on pathways. *Social Science and Medicine, 187*, 296–305. doi:10.1016/j.socscimed.2017.02.021

Khanna, J., Medvigy, D., Fueglistaler, S., & Walko, R. (2017). Regional dry-season climate changes due to three decades of Amazonian deforestation. *Nature Climate Change, 7*(3), 200–+. doi:10.1038/Nclimate3226

Khreis, H., Kelly, C., Tate, J., Parslow, R., Lucas, K., & Nieuwenhuijsen, M. (2017). Exposure to traffic-related air pollution and risk of development of childhood asthma: A systematic review and meta-analysis. *Environment International, 100*, 1–31. doi:10.1016/j.envint.2016.11.012

Kim, Y., Kim, H., Honda, Y., Guo, Y. L., Chen, B. Y., Woo, J. M., & Ebi, K. L. (2016). Suicide and ambient temperature in East Asian Countries: A time-stratified case-crossover analysis. *Environmental Health Perspectives, 124*(1), 75–80. doi:10.1289/ehp.1409392

Kjellstrom, T. (2004). The epidemic of asbestos-related diseases in New Zealand. *International Journal of Occupational and Environmental Health, 10*(2), 212–219. doi:10.1179/oeh.2004.10.2.212

Kjellstrom, T., Ferguson, R., & Taylor, A. (2009). *The total public health impact of road transport in Sweden: Technical report*. Borlange, Sweden: Swedish National Road Authority.

Kjellstrom, T., Freyberg, C., Lemke, B., Otto, M., & Briggs, D. (2017). Estimating population heat exposure and impacts on working people in conjunction with climate change. *International Journal of Biometeorology*. doi:10.1007/s00484-017-1407-0

Kjellstrom, T., & Hogstedt, C. (2009). Global situation concerning work-related injuries and diseases. In K. Elgstrand & I. Petterson (Eds.), *OSH for development* (pp. 741–761). Stockholm, Sweden: Royal Institute of Technology.

Kjellstrom, T., Holmer, I., & Lemke, B. (2009). Workplace heat stress, health and productivity: An increasing challenge for low and middle-income countries during climate change. *Global Health Action, 2*. doi:10.3402/gha.v2i0.2047

Kjellstrom, T., van Kerkhoff, L., Bammer, G., & McMichael, A. J. (2003). Comparative assessment of transport risks: How it can contribute to health impact assessment of transport policies. *Bulletin of the World Health Organization, 81*(6), 451–457.

Kraus, R. S. (1998). Oil exploration and drilling. In J. M. Stellman (Ed.), *Encyclopaedia of occupational health and safety* (4th ed., Vol. III, pp. 75.1–75.14). Geneva, Switzerland: International Labor Office.

LaDou, J., & Jeyaratnam, J. (1994). Transfer of hazardous industries: Issues and solutions. In J. Jeyaratnam & K. S. Chia (Eds.), *Occupational health in national development* (pp. 227–244). River Edge, NJ: World Scientific Publications.

Landrigan, P. J. (2016). Air pollution and health. *The Lancet Public Health, 2*(1), e4–e5. doi:10.1016/S2468-2667(16)30023-8

Landrigan, P. J., Fuller, R., Acosta, N. J. R., Adeyi, O., Arnold, R., Basu, N., . . . Zhong, M. (2017). *The Lancet* Commission on pollution and health. *The Lancet, 391*(10119), 462–512. doi:10.1016/S0140-6736(17)32345-0

Lawrence, D., & Vandecar, K. (2015). Effects of tropical deforestation on climate and agriculture. *Nature Climate Change, 5*(1), 27–36. doi:10.1038/nclimate2430

Le, G. V., Takahashi, K., Karjalainen, A., Delgermaa, V., Hoshuyama, T., Miyamura, Y., . . . Wagner, G. (2010). National use of asbestos in relation to economic development. *Environmental Health Perspectives, 118*(1), 116–119. doi:10.1289/ehp.0901196

Lee, K., & Dodgson, R. (2000). Globalization and cholera: Implications for global governance. *Global Governance, 6*(2), 213–236.

Leichenko, R. M. (2008). *Environmental change and globalization: Double exposures*. Oxford, UK: Oxford University Press.

Levy, B. S., Wegman, D. H., Karon, S. L., & Sokas, K. (Eds.). (2011). *Occupational health: Recognizing and preventing work-related disease and injury* (6th ed.). Philadelphia, PA: Lippincott, Williams and Wilkins.

Lewis, S. L., & Maslin, M. A. (2015). Defining the Anthropocene. *Nature, 519*, 171. doi:10.1038/nature14258

Li, Y., Wang, W., Kan, H., Xu, X., & Chen, B. (2010). Air quality and outpatient visits for asthma in adults during the 2008 Summer Olympic Games in Beijing. *Science of the Total Environment, 408*(5), 1226–1227. doi:10.1016/j.scitotenv.2009.11.035

Liu, L.-Y., He, K., Hites, R. A., & Salamova, A. (2016). Hair and nails as noninvasive biomarkers of human exposure to brominated and organophosphate flame retardants. *Environmental Science & Technology, 50*(6), 3065. doi:10.1021/acs.est.5b05073

Lixil, WaterAid Japan, & Oxford Economics. (2016). *The true cost of poor sanitation.* Tokyo, Japan: Authors.

Macmillan, A., Connor, J., Witten, K., Kearns, A., Rees, D., & Woodward, A. (2014). The societal costs and benefits of commuter bicycling: Simulating the effects of specific policies using system dynamics modeling. *Environmental Health Perspectives, 122*(4). doi:10.1289/ehp.1307250

Macmillan, A., Davies, M., Shrubsole, C., Luxford, N., May, N., Chiu, L. F., . . . Chalabi, Z. (2016). Integrated decision-making about housing, energy and wellbeing: A qualitative system dynamics model. *Environmental Health, 15*(1), 23–34. doi:10.1186/s12940-016-0098-z

Martens, P., Kovats, R. S., Nijhof, S., de Vries, P., Livermore, M. T. J., Bradley, D. J., . . . McMichael, A. J. (1999). Climate change and future populations at risk of malaria. *Global Environmental Change, 9*, S89–S107. doi:10.1016/S0959-3780(99)00020-5

McKenna, M. L., McAtee, S., Bryan, P. E., Jeun, R., Ward, T., Kraus, J., . . . Mejia, R. (2017). Human intestinal parasite burden and poor sanitation in rural Alabama. *American Journal of Tropical Medicine and Hygiene, 97*(5), 1623–1628. doi:10.4269/ajtmh.17-0396

McMichael, A. J. (1994). Invited commentary. "Molecular epidemiology": New pathway or new travelling companion? *American Journal of Epidemiology, 140*(1), 1–11.

McMichael, A. J. (2009). Human population health: Sentinel criterion of environmental sustainability. *Current Opinion in Environmental Sustainability, 1*(1), 101–106. doi:10.1016/j.cosust.2009.07.001

McMichael, A. J., & Bambrick, H. (2011). The global environment. In R. Detels, R. Beaglehole, M. A. Lansang, & R. Gulliford (Eds.), *Oxford textbook of public health* (5th ed., pp. 200–237). Oxford, UK: Oxford University Press.

McMichael, A. J., Campbell-Lendrum, D., Ebi, K. L., Githeko, A., Scheraga, J., & Woodward, A. (Eds.). (2003). *Climate change and human health: Risks and responses*. Geneva, Switzerland: World Health Organization.

McMichael, A. J., Haines, A., Sloof, R., & Kovats, S. (Eds.). (1996). *Climate change and human health*. Geneva, Switzerland: World Health Organization.

McMichael, A. J., Powles, J. W., Butler, C. D., & Uauy, R. (2007). Food, livestock production, energy, climate change, and health. *The Lancet, 370*(9594), 1253–1263. doi:10.1016/S0140-6736(07)61256-2

McMichael, A. J., Woodruff, R. E., & Hales, S. (2006). Climate change and human health: Present and future risks. *The Lancet, 367*(9513), 859–869.

Meo, S. A. (2004). Health hazards of cement dust. *Saudi Medical Journal, 25*(9), 1153–1159.

Millennium Ecosystem Assessment. (2005). *Living beyond our means: Natural assets and human well-being. Statement from the Board.* Washington, DC: Island Press.

Mock, C. N., Smith, K. R., Kobusingye, O., Nugent, R., Abdalla, S., Ahuja, R. B., . . . Watkins, D. A. (2017). Injury prevention and environmental health: Key messages from *Disease Control Priorities,* 3rd edition. In C. N. Mock, R. Nugent, O. Kobusingye & K. R. Smith (Eds.), *Disease control priorities: Injury prevention and environmental health* (3rd ed., pp. 1–24). Washington, DC: World Bank Group.

Montgomery, C. (2013). *Happy city: Transforming our lives through urban design.* New York, NY: Farrar, Straus and Giroux.

Moochhala, S. M., Shahi, G. S., & Cote, I. L. (1997). The role of epidemiology, controlled clinical studies, and toxicology in defining environmental risks. In G. S. Shahi, B. S. Levy, A. Binger, T. Kjellstrom, & R. Lawrence (Eds.), *International perspective on environment, development and health: Toward a sustainable world* (pp. 341–352). New York, NY: Springer.

Moore, C. F. (2003). *Silent scourge: Children, pollution and why scientists disagree*. Oxford, UK: Oxford University Press.

Morales, E., Julvez, J., Torrent, M., de Cid, R., Guxens, M., Bustamante, M., . . . Sunyer, J. (2009). Association of early-life exposure to household gas appliances and indoor nitrogen dioxide with cognition and attention behavior in preschoolers. *American Journal of Epidemiology, 169*(11), 1327–1336. doi:10.1093/aje/kwp067

Morgenstern, H., & Thomas, D. (1993). Principles of study design in environmental epidemiology. *Environmental Health Perspectives, 101*(suppl 4), 23–38.

Morse, S. S. (1995). Factors in the emergence of infectious diseases. *Emerging Infectious Diseases, 1*(1), 7–15. doi:10.3201/eid0101.950102

Moure-Eraso, R., Wilcox, M., Punnett, L., Copeland, L., & Levenstein, C. (1994). Back to the future: Sweatshop conditions on the Mexico–U.S. border. I. Community health impact of maquiladora industrial activity. *American Journal of Industrial Medicine, 25*(3), 311–324. doi:10.1002/ajim.4700250302

Naeher, L. P., Brauer, M., Lipsett, M., Zelikoff, J. T., Simpson, C. D., Koenig, J. Q., & Smith, K. R. (2007). Woodsmoke health effects: A review. *Inhalation Toxicology, 19*(1), 67–106. doi:10.1080/08958370600985875

Naghavi, M., Abajobir, A. A., Abbafati, C., Abbas, K. M., Abd-Allah, F., Abera, S. F., . . . Murray, C. J. L. (2017). Global, regional, and national age-sex specific mortality for 264 causes of death, 1980; 2013; 2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet, 390*(10100), 1151–1210. doi:10.1016/S0140-6736(17)32152-9

National Academy of Sciences. (2009). *Science and decisions: Advancing risk assessment.* Washington, DC: National Academy of Sciences, Division on Earth and Life Sciences.

National Research Council. (2012). *Exposure science in the 21st century: A vision and a strategy*. Washington, DC: National Academies Press.

Nesbit, J. (2017, 30 October). Climate change is bad for your health. *The New York Times*. Retrieved from https://www.nytimes.com/2017/10/30/opinion/climate-change-health-heat.html?_r=0

Nieuwenhuijsen, M. J. (2016). Urban and transport planning, environmental exposures and health-new concepts, methods and tools to improve health in cities. *Environmental Health, 15*(suppl 1), 38. doi:10.1186/s12940-016-0108-1

Nieuwenhuijsen, M. J., Khreis, H., Verlinghieri, E., & Rojas-Rueda, D. (2016). Transport and health: A marriage of convenience or an absolute necessity. *Environment International, 88*, 150–152. doi:10.1016/j.envint.2015.12.030

Normile, D. (2011). Tohoku-Oki earthquake: Fukushima revives the low-dose debate. *Science, 332*(6032), 908–910. doi:10.1126/science.332.6032.908

Nriagu, J. O., Blankson, M. L., & Ocran, K. (1996). Childhood lead poisoning in Africa: A growing public health problem. *Science of the Total Environment, 181*(2), 93–100.

Nybo, L., Kjellstrom, T., Bogataj, L. K., & Flouris, A. D. (2017). Global heating: Attention is not enough; we need acute and appropriate actions. *Temperature, 4*(3), 199–201. doi:10.1080/23328940.2017.1338930

NYC Environmental Protection. (2017). Filtration Avoidance Determination (FAD) reports. Retrieved from http://www.nyc.gov/html/dep/html/watershed_protection/fad.shtml

Oluwole, O., Arinola, G. O., Huo, D., & Olopade, C. O. (2017a). Biomass fuel exposure and asthma symptoms among rural school children in Nigeria. *Journal of Asthma, 54*(4), 347–356. doi:10.1080/02770903.2016.1227334

Oluwole, O., Arinola, G. O., Huo, D., & Olopade, C. O. (2017b). Household biomass fuel use, asthma symptoms severity, and asthma underdiagnosis in rural schoolchildren in Nigeria: A cross-sectional observational study. *BMC Pulmonary Medicine, 17*(1), 3. doi:10.1186/s12890-016-0352-8

One Health Commission. (2017). What is One Health? Retrieved from https://www.onehealthcommission.org/en/why_one_health/what_is_one_health/

O'Neil, J. M., Davis, T. W., Burford, M. A., & Gobler, C. J. (2012). The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change. *Harmful Algae, 14*, 313–334. doi:10.1016/j.hal.2011.10.027

Organisation for Economic Co-operation and Development (OECD). (1993). *OECD core set of indicators for environmental performance reviews.* Environmental Monographs No. 83. Paris, France: Author.

Panter, J., Heinen, E., Mackett, R., & Ogilvie, D. (2016). Impact of new transport infrastructure on walking, cycling, and physical activity. *American Journal of Preventive Medicine, 50*(2), e45–e53. doi:10.1016/j.amepre.2015.09.021

Parkes, M. W. (2016). Pacific connections for health, ecosystems and society: New approaches to the land–water–health nexus. *Reviews on Environmental Health, 31*(1), 125. doi:10.1515/reveh-2015-0067

Parsons, K. (2014). *Human thermal environment: The effects of hot, moderate and cold temperatures on human health, comfort and performance* (3rd ed.). New York, NY: CRC Press.

Patz, J. A., Gibbs, H. K., Foley, J. A., Rogers, J. V., & Smith, K. R. (2007). Climate change and global health: Quantifying a growing ethical crisis. *Ecohealth, 4*(4), 397–405. doi:10.1007/s10393-007-0141-1

Pell, J. P., Haw, S., Cobbe, S., Newby, D. E., Pell, A. C., Fischbacher, C., . . . Borland, W. (2008). Smoke-free legislation and hospitalizations for acute coronary syndrome. *New England Journal of Medicine, 359*(5), 482–491. doi:10.1056/NEJMsa0706740

Perera, F. P., Rauh, V., Whyatt, R. M., Tsai, W. Y., Tang, D., Diaz, D., . . . Kinney, P. (2006). Effect of prenatal exposure to airborne polycyclic aromatic hydrocarbons on neurodevelopment in the first 3 years of life among inner-city children. *Environmental Health Perspectives, 114*(8), 1287–1292.

Pocock, S. J., Smith, M., & Baghurst, P. (1994). Environmental lead and children's intelligence: A systematic review of the epidemiological evidence. *BMJ, 309*(6963), 1189–1197.

Pollock, J., Shi, L., & Gimbel, R. W. (2017). Outdoor environment and pediatric asthma: An update on the evidence from North America. *Canadian Respiratory Journal, 2017*, 8921917. doi:10.1155/2017/8921917

Pope, C. A., 3rd. (2009). The expanding role of air pollution in cardiovascular disease: Does air pollution contribute to risk of deep vein thrombosis? *Circulation, 119*(24), 3050–3052. doi:10.1161/CIRCULATIONAHA.109.870279

Pope, C. A., 3rd, Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., & Thurston, G. D. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Journal of the American Medical Association, 287*(9), 1132–1141.

Pope, C. A., 3rd, Burnett, R. T., Thurston, G. D., Thun, M. J., Calle, E. E., Krewski, D., & Godleski, J. J. (2004). Cardiovascular mortality and long-term exposure to particulate air pollution: Epidemiological evidence of general pathophysiological pathways of disease. *Circulation, 109*(1), 71–77. doi:10.1161/01.CIR.0000108927.80044.7F

Pope, C. A., 3rd, & Dockery, D. W. (2006). Health effects of fine particulate air pollution: Lines that connect. *Journal of the Air & Waste Management Association, 56*(6), 709–742.

Pope, C. A., 3rd, Ezzati, M., & Dockery, D. W. (2009). Fine-particulate air pollution and life expectancy in the United States. *New England Journal of Medicine, 360*(4), 376–386. doi:10.1056/NEJMsa0805646

Pope, C. A., 3rd, Hansen, M. L., Long, R. W., Nielsen, K. R., Eatough, N. L., Wilson, W. E., & Eatough, D. J. (2004). Ambient particulate air pollution, heart rate variability, and blood markers of inflammation in a panel of elderly subjects. *Environmental Health Perspectives, 112*(3), 339–345.

Poschen, P. (1998). General profile (forestry). In J. M. Stellman (Ed.), *Encyclopaedia of occupational health and safety* (4th ed., Vol. III, pp. 68.62–68.66). Geneva, Switzerland: International Labor Office.

Preston, S. H. (1991). *Fatal years: Child mortality in late nineteenth-century America*. Princeton, NJ: Princeton University Press.

Prüss-Üstün, A., & Corvalán, C. (2006). *Preventing disease through healthy environments: Towards an estimate of the environmental burden of disease.* Geneva, Switzerland: World Health Organization.

Prüss-Üstün, A., Wolf, J., Corvalán, C., Bos, R., & Neira, M. (2016). *Preventing disease through healthy environments: A global assessment of the burden of disease from environmental risks.* Geneva, Switzerland: World Health Organization.

Putra, A. A. G., Hampson, K., Girardi, J., Hiby, E., Knobel, D., Mardiana, I. W., . . . Scott-Orr, H. (2013). Response to a rabies epidemic, Bali, Indonesia, 2008–2011. *Emerging Infectious Diseases, 19*(4), 648. doi:10.3201/eid1904.120380

Qian, G. S., Ross, R. K., Yu, M. C., Yuan, J. M., Gao, Y. T., Henderson, B. E., . . . Groopman, J. D. (1994). A follow-up study of urinary markers of aflatoxin exposure and liver cancer risk in Shanghai, People's Republic of China. *Cancer Epidemiology, Biomarkers, and Prevention, 3*(1), 3–10.

Ramanathan, V., Chung, C., Kim, D., Bettge, T., Buja, L., Kiehl, J. T., . . . Wild, M. (2005). Atmospheric brown clouds: Impacts on South Asian climate and hydrological cycle. *Proceedings of the National Academy of Sciences USA, 102*(15), 5326–5333. doi:10.1073/pnas.0500656102

Ramanathan, V., & United Nations Environment Programme (UNEP). (2008). *Atmospheric brown clouds: Regional assessment report with focus on Asia: Summary.* Nairobi, Kenya: UNEP.

Ramani, R., & Mutmansky, J. (1999). Mine health and safety at the turn of the millennium. *Mining Engineering, 51*(9), 25.

Rees, W. E. (1996). Revisiting carrying capacity: Area-based indicators of sustainability. *Population and Environment, 17*(3), 195–215. doi:10.1007/Bf02208489

Rockstrom, J., Steffen, W., Noone, K., Persson, A., Chapin, F. S., Lambin, E. F., . . . Foley, J. A. (2009). A safe operating space for humanity. *Nature, 461*(7263), 472–475.

Rosenstock, L., Cullen, M. R., Brodkin, C. A., & Redlich, C. A. (Eds.). (2005). *Textbook of clinical occupational and environmental medicine* (2nd ed.). Philadelphia, PA: W. B. Saunders.

Rothman, K. J., Greenland, S., & Lash, T. L. (2012). *Modern Epidemiology* (3rd ed.). Philadelphia, PA: Wolters Kluwer.

Sahu, S., Sett, M., & Kjellstrom, T. (2013). Heat exposure, cardiovascular stress and work productivity in rice harvesters in India: Implications for a climate change future. *Industrial Health, 51*(4), 424–431.

Samet, J. M., Schnatter, R., & Gibb, H. (1998). Epidemiology and risk assessment. *American Journal of Epidemiology, 148*(10), 929–936.

Schwartz, J. (1994). Air pollution and daily mortality: A review and meta analysis. *Environmental Research, 64*(1), 36–52. doi:10.1006/enrs.1994.1005

Scutti, S. (2017, 20 October). Air pollution linked to 9 million deaths worldwide in 2015, study says. *CNN International Edition*. Retrieved from http://edition.cnn.com/2017/10/19/health/pollution-1-in-6-deaths-study/index.html

Searchinger, T., Hanson, C., Ranganathan, J., Lipinski, B., Waite, R., Winterbottom, R., & Dinshaw, A. (2013). *Creating a sustainable food future: Interim findings. A menu of solutions to sustainably feed more than 9 billion people by 2050.* Washington, DC: World Resources Institute.

Shaw, C., Hales, S., Howden-Chapman, P., & Edwards, R. (2014). Health co-benefits of climate change mitigation policies in the transport sector. *Nature Climate Change, 4*(6), 427–433. doi:10.1038/Nclimate2247

Sims, J. (1994). *Women, health and the environment.* Geneva, Switzerland: World Health Organization. Document WHO/EHG/94.11.

Singh, G. K., & Siahpush, M. (2014). Widening rural–urban disparities in life expectancy, U.S., 1969–2009. *American Journal of Preventive Medicine, 46*(2), e19–e29. doi:10.1016/j.amepre.2013.10.017

Smedley, J., Dick, F., & Sadhra, S. (2013). *Oxford handbook on occupational health* (2nd ed.). Oxford, UK: Oxford University Press.

Smil, V. (1994). *Energy in world history* (Vol. 36). Boulder, Colorado: Westview Press.

Smith, D. (1999). Worldwide trends in DDT levels in human breast milk. *International Journal of Epidemiology, 28*(2), 179–188.

Smith, K. R., & Ezzati. (2005). How environmental health risks change with development: The epidemiologic and environmental risk transitions revisited. *Annual Review of Environment and Resources, 30*, 291–333. doi:10.1146/annurev.energy.30.050504.144424

Smith, K. R., Woodward, A., Campbell-Lendrum, D., Chadee, D. D., Honda, Y., Liu, Q., . . . Sauerborn, R. (2014). Human health: Impacts, adaptation, and co-benefits. In C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, . . . L. L. White (Eds.), *Climate change 2014: Impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 709–754). Cambridge, UK: Cambridge University Press.

Smith, L. E. D., & Porter, K. S. (2010). Management of catchments for the protection of water resources: Drawing on the New York City watershed experience. *Regional Environmental Change, 10*(4), 311–326. doi:10.1007/s10113-009-0102-z

Smith, M., Hosking, J., Woodward, A., Witten, K., MacMillan, A., Field, A., . . . Mackie, H. (2017). Systematic literature review of built environment effects on physical activity and active transport: An update and new findings on health equity. *International Journal of Behavioral Nutrition and Physical Activity, 14*(1), 158. doi:10.1186/s12966-017-0613-9

Steffen, W., Broadgate, W., Deutsch, L., Gaffney, O., & Ludwig, C. (2015). The trajectory of the Anthropocene: The great acceleration. *Anthropocene Review, 2*(1), 81–98. doi:10.1177/2053019614564785

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., . . . Sörlin, S. (2015). Sustainability. Planetary boundaries: Guiding human development on a changing planet. *Science, 347*(6223), 1259855. doi:10.1126/science.1259855

Stellman, J. M., & International Labor Office (ILO). (1998). *Encyclopaedia of occupational health and safety* (4th ed.). Geneva, Switzerland: International Labor Office.

Stosch, K., Quilliam, R., Bunnefeld, N., & Oliver, D. (2017). *Managing multiple catchment demands for sustainable water use and ecosystem service provision* (Vol. 9, p. 677). Basel, Switzerland: MDPI AG.

Straif, K., Baan, R., Grosse, Y., Secretan, B., El Ghissassi, F., Cogliano, V., & WHO International Agency for Research on Cancer Monograph Working Group (2006). Carcinogenicity of household solid fuel combustion and of high-temperature frying. *The Lancet Oncology, 7*(12), 977–978.

Suades-Gonzalez, E., Gascon, M., Guxens, M., & Sunyer, J. (2015). Air pollution and neuropsychological development: A review of the latest evidence. *Endocrinology, 156*(10), 3473–3482. doi:10.1210/en.2015-1403

Suglia, S. F., Gryparis, A., Wright, R. O., Schwartz, J., & Wright, R. J. (2008). Association of black carbon with cognition among children in a prospective birth cohort study. *American Journal of Epidemiology, 167*(3), 280–286. doi:10.1093/aje/kwm308

Talbot, T. O., Haley, V. B., Dimmick, W. F., Paulu, C., Talbott, E. O., & Rager, J. (2009). Developing consistent data and methods to measure the public health impacts of ambient air quality for environmental public health tracking: Progress to date and future directions. *Air Quality, Atmosphere, and Health, 2*(4), 199–206. doi:10.1007/s11869-009-0043-1

Thomson, H., Thomas, S., Sellstrom, E., & Petticrew, M. (2013). Housing improvements for health and associated socio-economic outcomes. *Cochrane Database of Systematic Reviews, 2*, CD008657. doi:10.1002/14651858.CD008657.pub2

Thundiyil, J. G., Stober, J., Besbelli, N., & Pronczuk, J. (2008). Acute pesticide poisoning: A proposed classification tool/ Intoxication aigue par les pesticides: Proposition un outil de classification/Intoxication aguda por plaguicidas: Propuesta de instrumento de clasificacion. *Bulletin of the World Health Organization, 86*(3), 205.

Titus, D. J., Furones, C., Atkins, C. M., & Dietrich, W. D. (2015). Emergence of cognitive deficits after mild traumatic brain injury due to hyperthermia. *Experimental Neurology, 263*, 254–262. doi:10.1016/j.expneurol.2014.10.020

Tong, S. L., Baghurst, P. A., Sawyer, M. G., Burns, J., & McMichael, A. J. (1998). Declining blood lead levels and changes in cognitive function during childhood: The Port Pirie Cohort Study. *Journal of the American Medical Association, 280*(22), 1915–1919.

Tossavainen, A. (2000). International expert meeting on new advances in the radiology and screening of asbestos-related diseases. *Scandinavian Journal of Work, Environment, and Health, 26*(5), 449–454.

Tossavainen, A. (2004). Global use of asbestos and the incidence of mesothelioma. *International Journal of Occupational and Environmental Health, 10*(1), 22–25. doi:10.1179 /oeh.2004.10.1.22

United Nations. (2015a). *The Millennium Development Goals report 2015.* New York, NY: Author.

United Nations. (2015b). *Transforming our world: The 2030 agenda for sustainable development.* New York, NY: Author.

United Nations, Department of Economic and Social Affairs Population Division. (2017). *World population prospects: The 2017 revision, key findings and advance tables.* New York, NY: United Nations.

United Nations Framework Convention on Climate Change. (2015). The Paris Agreement. Paris, France. Retrieved from http://unfccc.int/paris_agreement/items/9485.php

United Nations Habitat. (2016). From Habitat II to Habitat III: Twenty years of urban development. In *World cities report 2016: Urbanization and development: Emerging futures* (pp. 1–24). Nairobi, Kenya: Author.

Vanakoski, J., & Seppala, T. (1998). Heat exposure and drugs:. A review of the effects of hyperthermia on pharmacokinetics. *Clinical Pharmacokinetics, 34*(4), 311–322. doi:10.2165/00003088-199834040-00004

Velis, C. A. (2014). *Global recycling markets: Plastic waste: A story for one player—China.* Report prepared by FUELogy and formatted by D-waste on behalf of International Solid Waste Association, Globalisation and Waste Management Task Force. Vienna, Austria.

Villanueva, C. M., Kogevinas, M., Cordier, S., Templeton, M. R., Vermeulen, R., Nuckols, J. R., . . . Levallois, P. (2014). Assessing exposure and health consequences of chemicals in drinking water: Current state of knowledge and research needs. *Environmental Health Perspectives, 122*(3), 213–221. doi:10.1289/ehp.1206229

Ward, M. H., deKok, T. M., Levallois, P., Brender, J., Gulis, G., Nolan, B. T., & VanDerslice, J. (2005). Workgroup report: Drinking-water nitrate and health—recent findings and research needs. *Environmental Health Perspectives, 113*(11), 1607–1614. doi:10.1289/ehp.8043

Watershed Agricultural Council. (2017). Retrieved from https:// www.nycwatershed.org/

Watts, N., Adger, W. N., Agnolucci, P., Blackstock, J., Byass, P., Cai, W., . . . Costello, A. (2015). Health and climate change: Policy responses to protect public health. *The Lancet*. doi:10.1016 /S0140-6736(15)60854-6

Watts, N., Amann, M., Ayeb-Karlsson, S., Belesova, K., Bouley, T., Boykoff, M., . . . Costello, A. (2017). *The Lancet* Countdown on health and climate change: From 25 years of inaction to a global transformation for public health. *The Lancet*, *391*(10120), P581–P630. doi:10.1016/S0140-6736(17)32464-9

Weeks, J. L. (1998). Health and safety hazards in the construction industry. In J. M. Stellman (Ed.), *Encyclopaedia of occupational health and safety* (4th ed., Vol. III, pp. 93.92–93.98). Geneva, Switzerland: International Labor Office.

Weiss, R. A., & McMichael, A. J. (2004). Social and environmental risk factors in the emergence of infectious diseases. *Nature Medicine, 10*(12s), S70. doi:10.1038/nm1150

Wesseling, C., Crowe, J., Hogstedt, C., Jakobsson, K., Lucas, R., & Wegman, D. H. (2013). The epidemic of chronic kidney disease of unknown etiology in Mesoamerica: A call for interdisciplinary research and action. *American Journal of Public Health, 103*(11), 1927–1930. doi:10.2105 /Ajph.2013.301594

Wong, G. W., Brunekreef, B., Ellwood, P., Anderson, H. R., Asher, M. I., Crane, J., . . . Group, I. P. T. S. (2013). Cooking fuels and prevalence of asthma: A global analysis of phase three of the International Study of Asthma and Allergies in Childhood (ISAAC). *The Lancet Respiratory Medicine, 1*(5), 386–394. doi:10.1016/S2213-2600(13)70073-0

Woodward, A. (2014). *The healthy country? A history of life and death in New Zealand.* Auckland, New Zealand: Auckland University Press.

World Health Organization (WHO). (1990). *Public health impacts of the use of pesticides in agriculture.* Geneva, Switzerland: Author.

World Health Organization (WHO). (1992). *Cadmium.* Geneva, Switzerland: Author.

World Health Organization (WHO). (1997). *Health and environment in sustainable development.* Geneva, Switzerland: Author.

World Health Organization (WHO). (1998). *Chrysotile asbestos.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2002). *World health report 2002: Reducing risks, promoting healthy life.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2004). *Heat waves: Risks and responses.* Copenhagen, Denmark: Author.

World Health Organization (WHO). (2006). *Elimination of asbestos related diseases.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2008). *Our cities, our health, our future: Acting on social determinants for health equity in urban settings.* Kobe, Japan: WHO, Centre for Health Development.

World Health Organization (WHO). (2013a). *The Helsinki Statement on Health in All Policies.* Helsinki, Finland: Autbor.

World Health Organization (WHO). (2013b). *Review of evidence on health aspects of air pollution: REVIHAAP Project: Technical report.* Copenhagen, Denmark: Author.

World Health Organization (WHO). (2013c). *WHO expert consulation on rabies: Second report. WHO Technical Report Series*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2014). *Quantitative risk assessment of the effects of climate change on selected causes of death, 2030s and 2050s.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2015). *Global status report on road safety.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2016). WHO Global Urban Ambient Air Pollution Database (update 2016). Retrieved from http://www.who.int/phe/health_topics/outdoorair/databases/cities/en/

World Health Organization (WHO). (2017). *Guidelines for drinking-water quality* (4th ed., incorporating the first addendum). Geneva, Switzerland: Author.

World Health Organization (WHO), Global Health Observatory. (2016). Life expectancy at birth (years), 2000–2015. Retrieved from http://gamapserver.who.int/gho/interactive_charts/mbd/life_expectancy/atlas.html

World Health Organization (WHO) & United Nations Children's Fund (UNICEF). (2017). *Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines.* Geneva, Switzerland: WHO.

Wrangham, R. W. (2009). *Catching fire: How cooking made us human*. New York, NY: Basic Books.

Wyndham, C. H. (1969). Adaptation to heat and cold. *Environmental Research, 2*(5), 442–469.

Zhang, J. J., & Smith, K. R. (2007). Household air pollution from coal and biomass fuels in China: measurements, health impacts, and interventions. *Environmental Health Perspectives, 115*(6), 848–855. doi:10.1289/ehp.9479

Zinsstag, J., Schelling, E., Waltner-Toews, D., & Tanner, M. (2011). From "one medicine" to "one health" and systemic approaches to health and well-being. *Preventive Veterinary Medicine, 101*(3–4), 148–156. doi:10.1016/j.prevetmed.2010.07.003

CHAPTER 12

Complex Emergencies

Michael J. Toole and Ronald J. Waldman

Introduction

This chapter focuses on public health emergencies that arise from complex political crises but draws, in places, on lessons learned from humanitarian responses to natural disasters. Terminology changes frequently in this area, and different definitions emphasize different aspects of a concept. The term *complex humanitarian emergencies* came into popular use following the Kurdish refugee exodus during the Gulf War in 1991. It was defined by the Centers for Disease Control and Prevention (CDC) as "a situation affecting large civilian populations which usually involves a combination of factors including war or civil strife, food shortages, and population displacement, resulting in significant excess mortality" (Burkholder & Toole, 1995, p. 1012). Goodhand and Hulme (1999) defined *complex political emergencies* as conflicts that combine a number of features: They often occur within but also across state boundaries; they have political antecedents, often relating to competition for power and resources; they are protracted in duration; they are embedded in and are expressions of existing social, political, economic, and cultural structures and cleavages; and they are often characterized by predatory social formations. This latter definition clearly locates the causes and effects in the political sphere—a point echoed by numerous other writers, and one that has considerable implications for those working in these settings with a primarily public health agenda. In the new millennium, these terms have merged into *complex emergencies* (CEs), the term that we use in this chapter to maintain simplicity and consistency.

This chapter grapples with current understanding of CEs and their political causes, and considers their impact on populations and health systems. It incorporates new knowledge and experience gained in the past six years, including the massive humanitarian crisis resulting from the civil conflict in Syria. It highlights current knowledge in humanitarian assistance and indicates that effective technical interventions are possible to help alleviate suffering and limit adverse effects on the health of populations. We draw attention to current efforts by the humanitarian community to improve the effectiveness, efficiency, and equity of humanitarian responses, and consider how the pattern of early responses may influence the longer-term survival of populations and systems and the nature of any post-conflict society established. We are acutely aware, however, that the solutions to CEs are political and not humanitarian, and that it is in the political sphere that both upstream and downstream responses to complex emergencies must receive priority.

Many of the public health interventions currently promoted as best practices were developed and tested during the 1970s and 1980s, when most conflict-affected populations were accessed in relatively secure refugee camps in low- and middle-income countries (LMICs) administered by the Office of the United Nations High Commissioner for Refugees (UNHCR). Recent changes in the nature of conflicts have introduced a great deal of complexity into these responses.

Most contemporary wars are of "protracted duration, intrastate, fought by irregular armed groups, and fueled by economic opportunities and ethnic rivalry" (Spiegel, Checchi, Colombo, & Paik, 2010, p. 341). Many recent displaced populations are in or fleeing from middle-income countries, such as Syria. The populations most in need are often trapped by the armed conflict, unable to flee, not residing in organized camps, and, therefore, need to be accessed by relief agencies "in situ." Such "trapped" populations can be found in northeast Nigeria, Central African Republic, Somalia, Syria, and Yemen.

The conditions experienced by those who manage to flee their country and obtain refugee status have also changed significantly. Since 2011, most Syrian and Iraqi refugees have been housed outside camps. For example, in 2017, the majority of the 3 million Syrian refugees in Turkey were living in urban areas, with only 260,000 accommodated in the 21 government-run refugee camps (UNHCR, 2017a). In Jordan, more than 655,000 Syrian men, women, and children were trapped in exile. Approximately 75% of them lived outside camps, while more than 140,000 found sanctuary at the camps of Za'atari and Azraq. There were no formal refugee camps in Lebanon; as a result, more than 1 million registered Syrians were scattered throughout more than 2,100 urban and rural communities and locations in Lebanon, often sharing small basic lodgings with other refugee families in overcrowded conditions.

Although the impact of natural disasters is not the subject of this chapter, we have made two important exceptions: the destructive tsunami that originated off the coast of Indonesia on December 26, 2004, and the earthquake that devastated Port-au-Prince, Haiti, and its environs on January 12, 2010. We include these catastrophic events because they had many elements of a complex emergency.

The tsunami caused more than 228,000 deaths in 14 countries and displaced more than 1.5 million people from their homes (ReliefWeb, 2009). Two of the areas most severely affected—the Aceh province of Indonesia and northeast Sri Lanka—had experienced several decades of armed conflict. More than 127,000 Burmese refugees were living in the 5 Thai provinces affected by the tsunami: At least 1,000 perished (United States Committee for Refugees, 2005). The massive displacement and disruption to food and water supplies, shelter, and sanitation created conditions similar to those experienced by refugees and internally displaced persons (IDPs) fleeing armed conflict.

The Haiti earthquake killed 230,000 people, injured an approximately equal number, left 604,000 people homeless in Port-au-Prince, and caused the flight of an additional 598,000 residents from the capital city to the countryside. Its impact on the only major city in the poorest country of the Western Hemisphere, including the government paralysis that followed in the earthquake's wake, the massive international relief effort that included interventions by armed forces from many different nations, and the difficulty that is foreseen in returning people to their homes is, again, similar to what might be encountered in conflict settings.

In the period from the end of World War II to the end of the Cold War, most conflicts took place in the developing regions of the world, primarily in Africa, the Middle East, Asia, and Latin America (**FIGURE 12-1**). The end of the Cold War, the break-up of the Soviet Union, and the pace and intensity of globalization led, in the 1990s, to major conflicts in Europe and the former Soviet Union—most notably in Tajikistan, Chechnya, Georgia, Abkhazia, and the former Yugoslavia. The number of armed conflicts globally reached a peak of 52 in 1991. This point was followed by a steady decline until 2012, when an upswing began in part due to conflicts following the so-called Arab Spring, reaching 50 again in 2015 (Uppsala Conflict Data Program, 2016).

Modern-day conflicts are increasingly internal, rather than between states, and often have as a prime objective, alongside the quest for economic and political power, the undermining of the lives and livelihoods of civilian populations associated with opposing factions. As many as 90% of those persons affected in recent conflicts have been civilians, with all ages and both sexes suffering the effects. The distribution of impact and health outcomes vary substantially, however, and depend on the nature of the conflict and its history, its extent and form, and the prior health and health system status.

Many CEs and major natural disasters have attracted considerable media attention and have caused people to seek to promote availability of at least a basic degree of humanitarian assistance, even if fundamental political solutions are not sought. Other ongoing crises, despite causing massive loss of life, population displacement, and infrastructure destruction, are not necessarily explicitly recognized as CEs; as a result, they have attracted few resources and attention. These "hidden emergencies" nevertheless pose fundamental challenges to the health and well-being of affected populations. For example, the ongoing conflicts in Central African Republic, northeast Nigeria, and the Democratic Republic of the Congo (DRC) have drawn little attention and responses from the rest of the world: Such discrepancies are likely to

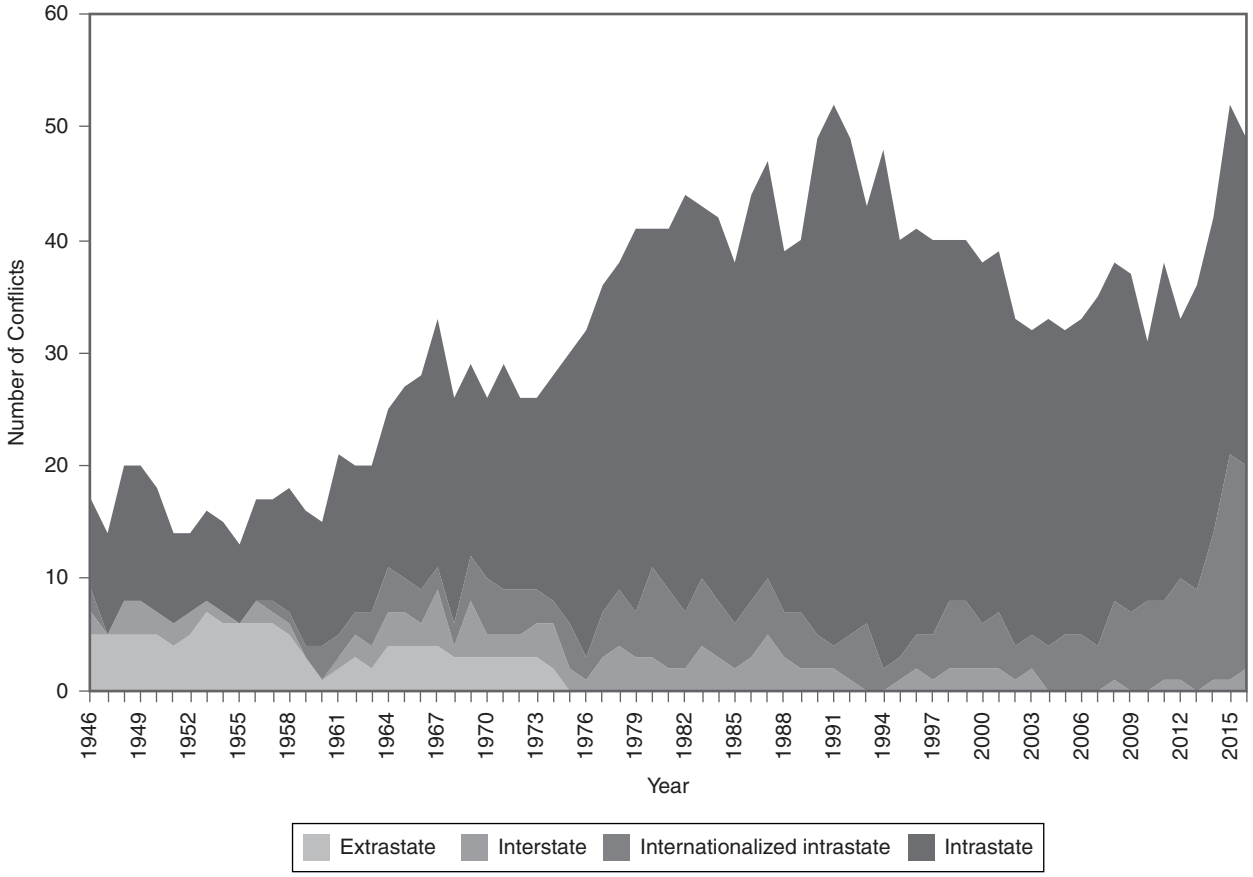


FIGURE 12-1 Armed conflict by type, 1946–2015.

Reproduced from Uppsala Conflict Data Program (2016). Retrieved from http://ucdp.uu.se/.

result from geopolitical concerns, media interest, and economic factors.

In the aftermath of the September 2001 terrorist attacks on New York City and Washington, D.C., United States–led coalitions have used military force to change the regimes in Afghanistan and Iraq. These two conflicts have been "internationalized" and represent a new trend in modern wars, given the huge disparities in military resources between the warring parties. These two conflicts were ongoing in 2018 with intensification of fighting in both countries and, in Iraq, the additional brutality of the radical jihadist movement Islamic State of Iraq and the Levant (ISIL).

During the decades since the Holocaust, numerous episodes of massive human rights atrocities and genocide have been committed against particular groups: Pol Pot's killing fields in Cambodia, the genocide against Tutsis in Rwanda, and "ethnic cleansing" in the former Yugoslavia. Other conflicts, such as that in West Darfur, Sudan, which began in 2004, highlight the brutal nature of internal wars, including the use of repressive techniques to evict people from their homes and to undermine their sense of security and safety, accompanied by the targeted use of force to destroy social, political, and economic structures.

A particularly insidious development is the targeting of violence toward individuals and groups on the basis of their ethnicity or religion. Such conflicts have occurred frequently enough that the term "ethnic cleansing" has entered the language. However, the reality has been that opportunistic politicians have often inflamed the perceived differences between groups, especially during times of economic and political uncertainty, resulting in open armed conflict.

In Syria, Iraq, and other countries, ISIL has taken violence against civilians to a new level. Targeting Shia Muslims, Christians, and Sunni Muslims who oppose them, members of ISIL have engaged in brutal torture and killings and led to massive movements of internally displaced persons and refugees. Other radical jihadist groups in Somalia (Al Shabab) and Nigeria (Boko Haram) have employed similar violent tactics.

In the same conflicts, we see evidence of other key features of modern-day CEs:

- The willingness of powerful segments of the international community to intervene in internal conflicts, and to do so in a way that minimizes their exposure to risk
- The changing nature of humanitarian assistance, which increasingly forms only one dimension of the management of conflicts, alongside political, economic, and military responses
- The changing role of the private sector, as well as the increasingly important role of local and global media
- Trends in globalization that simultaneously integrate peripheral areas within the global economy, yet contribute to their fragmentation as elites compete for access to the economic and political resources associated with integration in the global political economy

One consequence of the targeting of entire communities and their livelihoods has been a dramatic rise in the number of forcibly displaced people. By the end of 2016, 65.6 million individuals were forcibly displaced worldwide as a result of persecution, conflict, violence, or human rights violations (UNHCR, 2017a). That number represented an increase of 300,000 people over the previous year, and the world's forcibly displaced population remains at a record high. Globally, in December 2016, there were an estimated 22.5 million refugees, 2.8 million asylum-seekers (those seeking refugee status), and 40.3 million IDPs, the vast majority of whom were fleeing conflict zones.

Although refugee numbers are typically assessed to plan and provide relief, relatively little attention has been devoted to developing the most appropriate methods for establishing the precise composition of refugee and IDP populations, whether in terms of age, sex, religion, local geographic origin, or ethnicity. This imposes constraints given the differing needs and roles of groups within populations and may make it easier for the more complex issues of dealing with gender, equity, and ongoing intergroup rivalry to be overlooked. Particular groups, such as older adults, refugees not in camps, and IDPs, may neither be identified nor receive the required attention for their differing needs.

The changing pattern of conflict has been accompanied by significant changes in the delivery of humanitarian assistance. On the one hand, the number of agencies operating in these complex settings has increased dramatically; for example, more than 240 nongovernmental organizations (NGOs) were working in and around Rwanda in the aftermath of the genocide and more than 100 agencies were working with the Kosovar Albanians during 1999. On the other hand, lower-profile CEs may be as severe and life threatening to large populations, such as that in Sierra Leone in 1999, yet attract much less media attention and intergovernmental and humanitarian responses. New NGOs established in response to a specific conflict may be short-lived, inexperienced, and unable to cope with the challenges they face in providing services in complex political environments. The principle that such an organization must do more good than harm must underlie all interventions.

Every conflict has winners and losers. Predators identify opportunities amidst the turmoil to further enrich themselves and entrench their political position; thus, these players may have an interest in perpetuating the conflict. Humanitarian aid itself may become a resource over which groups compete, and such assistance and resources may directly or indirectly stoke the conflict. In some distressing circumstances, humanitarian aid has been used unwittingly to attract populations that were subsequently targeted by combatants, as in the DRC and Sudan. Humanitarian workers have increasingly been directly targeted in latter-day conflicts, leading to increased efforts to work closely with the military and security sector. Despite some benefits, such as improved logistics support, this trend may bring negative consequences and additional dangers and may threaten the neutrality and impartiality to which many agencies aspire.

In recent conflicts, health facilities, including those supported by international NGOs, have been targeted by combatants. For example, between 2015 and 2017, nearly 100 medical facilities supported by Médecins sans Frontières (MSF) were bombed (ReliefWeb, 2017a). The vast majority of these sites were in Syria, but facilities in Yemen, Afghanistan, Ukraine, and South Sudan were also hit.

NGOs are not a homogeneous community. Although some are highly professional and have given considerable thought to the development of humanitarian and technical policies and programs, the way in which humanitarian assistance has been provided has brought some negative consequences. Recognition by the humanitarian community of these problems has led to a great deal of evaluation and introspection, along with measures to improve practice, including the development of codes of conduct for humanitarian agencies, the promotion of minimum standards for service provision, and debate regarding enhancing accountability to affected populations. The most widely used minimum standards have been compiled

by the Sphere Project, which is in the process of producing its fourth edition (Sphere Project, 2011).

The first-ever World Humanitarian Summit took place in Istanbul in May 2016. As part of the lead-up to the summit, a number of donors and NGOs negotiated what was known as the "Grand Bargain"—a commitment from donors and aid organizations to provide 25% of global humanitarian funding to local and national responders by 2020, along with more non-earmarked money, and increased multiyear funding to ensure greater predictability and continuity in humanitarian response, among other commitments. The Grand Bargain was endorsed at the World Humanitarian Summit.

Promoting the derivation and uptake of good practices is particularly difficult in humanitarian agencies given their rapid staff turnover, unwillingness to publicly acknowledge failures and limitations because of the possible funding consequences, and a culture of doing rather than reflecting. Interventions are sometimes not evidence-based, and, despite most agencies valuing the concept of coordination, few wish to be coordinated. Poor-quality services have significant adverse consequences: increased morbidity, mortality, and disability; further spread of communicable diseases; community dissatisfaction and breakdown; and psychosocial distress. Clear policy objectives for interventions are often lacking, and mechanisms for working with new players such as the military and the private sector remain inadequately developed. Despite recognition that the accountability of relief efforts to affected populations should be enhanced, mechanisms to assure this relationship is honored are in their infancy.

Ongoing humanitarian challenges include understanding how best to upgrade host population health services alongside efforts to improve those services made available specifically to refugees, how best to ensure that urban refugees can access effective and affordable health care, how to most humanely and efficiently provide good-quality services, and how to maintain the role of communities in structuring both the determination of priorities and the pattern of service provision. A key issue relates to how and whether to bolster and support resilient health and social systems and individual adaptations to conflict. Our level of knowledge regarding these responses, and the potential to further support them, is weak. A persistent challenge for humanitarian workers is the quest to institutionalize a sensitive and inclusive evidence-based culture and to build sustainable mechanisms for crystallizing policy advice from the vast and valuable foundation of field experience.

Direct Public Health Impact of War

Measuring the impact and hidden costs of conflict is complex for a variety of reasons—methodological and theoretical shortcomings, inconsistencies in definitions and terms, restricted access to areas of conflict and sources of information, the rapid evolution of many emergencies, political manipulation of data, resource constraints, and the hidden or indirect nature of the impact. One of the consequences of the data-related limitations is difficulty in identifying more precisely which sections of the population are at greatest risk so as to develop more appropriate responses. Most poor countries lack reliable health information and vital registration systems, the absence of which increases the difficulties of determining the conflict-associated costs in terms of morbidity, mortality, and disability. Furthermore, CEs may seriously disrupt any surveillance and information systems that do exist.

Lack of consistency in definitions used makes it difficult to compare data within and across populations. Different agencies may define refugees and IDPs in different ways, case definitions for particular conditions vary, and techniques for estimating nutritional deficiencies, for example, may vary among different agencies working with the same population. Data are at times incomplete because impartial observers who attempt to provide more accurate figures may not have access to witnesses or other reliable sources of information.

During the last 20 years, a number of expert meetings have been held to develop more reliable methods of measuring both direct and indirect mortality resulting from CEs. Some of their proceedings have been published (Reed & Keeley, 2001; Working Group for Mortality Estimation in Emergencies, 2007). Other publications have sought to help non-epidemiologists analyze and use mortality data (Checchi & Roberts, 2005). Cross-sectional retrospective mortality surveys have been increasingly employed to measure conflict-related mortality. In 2004 alone, large mortality surveys were conducted in three conflict zones: Iraq, the DRC, and the West Darfur region of Sudan. All of these surveys used cluster sampling methodology.

Between 2000 and 2007, the International Rescue Committee (IRC), an international NGO, conducted a series of mortality surveys to evaluate the humanitarian impact of the conflict in DRC. The third nationwide survey in 2007 found that the crude mortality rate (CMR) remained elevated at a level almost 70% higher than that reported in the 1984 national census

and more than 55% higher than the reported baseline for sub-Saharan Africa (Coghlan et al., 2009). Fewer than 10% of deaths documented in the studies were directly attributable to violence. Instead, most were due to the indirect public health effects of conflict, including higher rates of infectious diseases, increased prevalence of malnutrition, and neonatal conditions. The authors of these studies concluded that security interventions, rather than health, would have the greatest impact on reducing mortality rates.

Although innovative techniques may be used to try to build up a picture of what transpired during a particular CE, the needs that arose during the CE, and the nature of the response, the sources of data may be biased, as may the ways in which information has been collected. Despite these challenges, innovative groups—often NGOs without a political agenda linked to any of the key players in the conflict—may be able to play a valuable role in documenting precisely what occurred and what the nature of present needs is. MSF has taken an innovative approach to documenting the impact of CEs through its MSF Speaking Out series, in which extracts from all relevant internal and external MSF documents are collated and annotated to form a compelling chronicle of a crisis and published both in hard copy and on the "MSF-CRASH" website (http://www.msf-crash.org/en). One example is the book published in 2017: *Humanitarian Aid, Genocide and Mass Killings: The Rwandan Experience, 1982–97* (Bradol & Le Pape, 2017).

Even where huge numbers of people are involved in a CE, agreement on the magnitude of impact varies. For example, the reported numbers of deaths in Syria since the beginning of the civil war have varied between government and opposition sources. The most reliable estimate is probably the December 2016 estimate of the Syrian Centre for Policy Research, which suggested that 470,000 direct or indirect deaths had occurred due to the Syrian conflict since it began in 2011, and 11.5% of the population had been killed or injured. Approximately 10% of these deaths occurred among IDPs. The annual CMR in Syria increased from 4.4 deaths per 1,000 in 2010 to 10.9 deaths per 1,000 in 2014, accounting directly and indirectly for the death of approximately 1.4% of the total population.

Data on numbers of IDPs and refugees may be manipulated by states and organizations in an attempt to make a political point or to maximize access to resources. It has been alleged that some refugee camp administrators and refugees report fewer deaths than actually occur so as to maintain levels of international assistance—having fewer beneficiaries could result in less assistance. In Nepal, in an attempt to encourage reporting of deaths among Bhutanese refugees, free funeral shrouds were offered to relatives of the deceased together with assurances that the reporting would not result in decreases of rations; this approach has also been used elsewhere.

Physical Impact

Political conflicts in the latter half of the 20th century were mostly waged between armies and trained combatants. The main direct results, in the form of deaths, morbidity, and disability, reflected the nature of the conflict, the level of technology and the nature of weapons, the prior preparation and protective clothing available to military personnel, and the quality of emergency medical care and evacuation facilities.

In the vast majority of latter-day conflicts, however, the entire population is often targeted, in part directly, but also with massive and sustained efforts at reducing the viability and integrity of the affected community. Injuries and disabilities may follow the use of firearms, but technology levels need not be high to achieve terrible levels of destruction. For example, the Rwandan genocide was largely committed with a combination of guns and machetes.

Antipersonnel land mines have also been responsible for significant population burdens, especially in a small number of heavily infested countries such as Angola, Cambodia, and Afghanistan. Despite a global ban, land mines have been deployed in recent conflicts. One of the most notorious examples was the Syrian town of Madaya in 2016. Syrian government forces had laid siege to the town, in the Rural Damascus Governorate, depriving roughly 40,000 residents of food and medicine for almost a year. The town was surrounded by an estimated 12,000 land mines. Between November 2015 and May 2016, 86 people died from siege-related causes, including 14 from land mines and 6 from snipers (Physicians for Human Rights [PHR] & Syrian American Medical Society, 2016).

Sexual Violence

Rape is increasingly recognized as a feature of internal wars, but has been present in many different types of conflicts. In some conflicts, rape has been used systematically as an attempt to undermine opposing groups. For example, an estimated 215,000 to 257,000 women were victims of sexual violence during the civil war in Sierra Leone (Project Ploughshares, 2002). The effects of gender-based violence (GBV) in conflict and post-conflict areas are numerous and severe. Sexually transmitted infections are a lasting

consequence of GBV and are a major health concern for women in conflict areas. Physical harms such as injury to reproductive organs, traumatic fistulas, and infertility often accompany brutal or repeated rapes. Attempts at abortion following an unwanted pregnancy from rape may also have severe medical complications.

The widespread sexual violence associated with armed conflict in eastern DRC has led it to be named the "rape capital of the world." A 2014 survey conducted in North Kivu Province showed that 22% of women were victims of sexual violence within the conflict (European Parliament, 2014). In addition, 50% of women had experienced sexual violence in a domestic context, evidence of the spread of what some call a "rape epidemic."

The findings of a systematic review in 2014 suggested that approximately one in five refugee or displaced women in complex humanitarian settings had experienced sexual violence (Vu et al., 2014). A study in Uganda found that despite widespread rape, few women spoke of their victimization (Bracken, Giller, & Kabaganda, 1992). Women of all ages had been targeted by sexual violence, but only few of those—and many of them only after several years—sought medical care and psychological help. Interestingly, recent studies have shown that during conflict a surprisingly high proportion of sexual violence is perpetrated on women by domestic intimate partners (Stark et al., 2017).

UN peacekeepers have been hit by a series of accusations of sexual exploitation and abuse of civilians across the globe. At least 55 UN peacekeepers were accused of sexual exploitation and abuse of civilians across UN missions around the globe during the first half of 2017, including a number of allegations against UN peacekeepers in Central African Republic (Essa, 2017).

Women who are on their own during a CE may find it more difficult to assure their safety and that of their children. They may become targets of violence from three sides: from the opposing army, from the armed forces in the country to which they have fled, and sometimes from their own community (Palmer & Zwi, 1998). They may be forced to provide sex in exchange for food, shelter, or other necessities for self and family survival. The experience of Afghan refugees is illustrative: "In the camps in Pakistan, most of which are controlled by one or other of the warring Afghan factions, women have been attacked, particularly those who are unaccompanied by men. If they refuse sexual favors, they are often denied access to vital rations" (Amnesty International, 1995, p. 15).

The issue of safety of women should be carefully considered when planning camp and other facilities: The siting of water, latrines, and cooking fuels should be undertaken in such a way as to reduce risks of abuse and violence, as described in the Sphere Project guidelines.

Human Rights

In times of war, the Universal Declaration of Human Rights and other laws, covenants, declarations, and treaties that constitute the body of human rights law are complemented by international humanitarian law. The latter is "a set of rules aimed at limiting violence and protecting the fundamental rights of the individual in times of armed conflict" (Perrin, 1996, p. 381). These rules are intended to govern the conduct of war by banning the use of certain weapons and by minimizing the effects of armed conflicts, whether international or internal, on noncombatants. The protection of the rights of noncombatants in wartime is based primarily on the Geneva Conventions of 1949 and the two Additional Protocols of 1977. Yet, despite the existence of both of these bodies of international law, CEs are consistently associated with serious infringements of the dignity of individuals and, more specifically, with a major impact on the health status of affected individuals and populations.

General practices that can be considered clear violations of international humanitarian law include the intentional targeting of civilian noncombatants, medical personnel, and civilian health facilities. Protection is also conferred upon prisoners of war, wounded and ill combatants, and military medical installations. Violations by states and individuals occurred with great frequency during the second half of the 20th century. One of the most prominent example was the genocidal activities perpetrated upon the Tutsi population of Rwanda in 1994. In the 1990s, the governments of Serbia and Croatia pursued ethnic cleansing policies against the populations of neighboring republics of the former Yugoslavia and, in the case of Serbia, the province of Kosovo. Arab militia in western Sudan, known as the Janjaweed, terrorized the Zaghawa, Masaalit, and Fur peoples between 2003 and 2005 with active support from the government of Sudan. This overt support led to the issuance in March 2009 of an arrest warrant by the International Criminal Court for the president of Sudan, Omar Hassan al-Bashir. Between 2011 and 2017, numerous human rights abuses were committed by all sides of the Syrian conflict, including torture, summary executions, and sexual violence.

Although violations of human rights law and international humanitarian law are crimes, the legal systems for punishing the perpetrators and compensating the victims are grossly inadequate. To date, four international tribunals have been established to prosecute war criminals from the former Yugoslavia, Rwanda, Cambodia, and Sierra Leone. Although these courts help to move the punishment of war criminals from theory to practice, they have been slow to act and expensive to implement. The establishment of an International Criminal Court—a permanent standing body dedicated to the trial and punishment of individuals accused and convicted of violations of human rights law—is, at least conceptually, another step toward strengthening what has in many respects been a legal system without law enforcement capability.

Reporting and responding to reports of human rights violations pose major problems. Although wars and internal conflicts have been proximate causes of most humanitarian emergencies, few of the individuals and agencies that have been involved in providing relief to the individuals affected are trained in the recognition of human rights violations or know where and how to report them. Until more widespread attention is paid to these crimes against humanity, the victims will continue to suffer from preventable acute and chronic morbidity, and the perpetrators will largely go unpunished. It would be useful to treat human rights violations as a major cause of morbidity and mortality during wars and their aftermath and to establish the epidemiologic characteristics of their distribution (Spirer & Spirer, 1993).

Indirect Public Health Impact of Civil Conflict

This section focuses in detail on the impacts on the health of populations that are not directly consequences of violence. Although the chapter as a whole focuses on the public health consequences of armed conflict, there is a phased evolution of public health effects as a country or region moves from political disturbances, economic deterioration, and civil strife through armed conflict, population migration, food shortages, and the collapse of governance and physical infrastructure. Thus, this section attempts to frame the indirect consequences of civil conflict in the changing context of evolving humanitarian emergencies.

Food Scarcity

As political disturbances evolve in a country, there is generally a significant effect on national and local economies. Under such scenarios, especially in LMICs, one of the first health effects is undernutrition in vulnerable groups, often caused by food scarcity. Local farmers may not plant crops as extensively as usual, or may decrease the diversity of their crops due to the uncertainty created by the economic or political situation. The cost of seeds and fertilizer may increase, and government agricultural extension services may be disrupted, resulting in lower yields. Distribution and marketing systems may be adversely affected. Devaluation of the local currency may drive down the price paid for agricultural produce, and the collapse of the local food processing industry may further diminish demand for agricultural products.

If full-scale armed conflict occurs, the fighting may damage irrigation systems, crops might be intentionally destroyed or looted by armed soldiers, distribution systems may collapse completely, and there may be widespread theft and looting of food stores. In countries that do not normally produce agricultural surpluses or that have large pastoral or nomadic communities, the impact of food deficits on the nutritional status of civilians may be severe. Such has been the case not only in sub-Saharan Africa, but also in Yemen (**EXHIBIT 12-1**). If adverse climatic factors intervene, as often happened in drought-prone countries such as Sudan, Somalia, Mozambique, and Zimbabwe, the outcome may be catastrophic famine.

Famine may be defined as high malnutrition and mortality rates resulting from inadequate availability of food. Lack of food availability may result from either insufficient production or inadequate or inequitable distribution.

EXHIBIT 12-1 Yemen

The civil war in Yemen has exacerbated the country's preexisting challenges, including poverty, poor health, and shortages of basic necessities such as water, fuel, and medications. By late 2016, the price of food had increased by 55% in that country and the gross domestic product (GDP) had contracted by almost 33%. Moving safely and freely was impaired, and access to clean water was a complicating barrier. An outbreak of watery diarrhea that caused much morbidity and mortality in Yemen was attributed to cholera. An estimated 14.4 million people, almost half of whom were severely affected, were unable to meet their food needs (Action Against Hunger, 2017). UNICEF estimated that approximately half of children in Yemen were affected by stunting (chronic malnutrition), and nearly 460,000 of Yemen's children were suffering from severe acute malnutrition (Eshaq et al., 2017).

When food aid programs are established, there may be inequitable distribution due to political factors, food stores may be damaged or destroyed, food may be stolen or diverted to military forces, and the distribution of food aid may be obstructed (Macrae & Zwi, 1994). The resulting food shortages may cause prolonged hunger and eventually drive families from their homes in search of relief. There have been many examples of food aid diversion, including in Sudan, Somalia, and the former Yugoslavia in the 1990s and in South Sudan and Yemen in the 21st century. Indeed, in latter-day CEs, targeting of relief assistance and the use of humanitarian aid as a resource that enables the warring parties to continue their violence is an ongoing challenge for humanitarian agencies.

Population Displacement

A common response by families and communities to civil conflict is to flee the violence. Individuals may flee because they fear persecution due to their particular political beliefs, ethnicity, or religion. In some societies, migration of part of the family to a safer area may be a traditional coping mechanism, with adult males staying behind to care for their land and animals. Some of these men may also be directly involved in the conflict. Mass migration and food shortages have been responsible for most deaths following civil conflicts in Africa and Asia.

Refugees

Refugees are defined under several international conventions as persons who flee their country of origin through a well-founded fear of persecution for reasons of race, religion, social class, or political beliefs (but not for economic reasons).

A number of large mass migrations occurred in the last decade of the 20th century, such as the 1 million Kurdish refugees who fled Iraq for Iran or Turkey in 1991; the 1.5 million refugees or displaced persons within the republics of the former Yugoslavia in 1993; the estimated 2 million Rwandan refugees who fled into Tanzania, Eastern Zaire, and Burundi in 1994; and the approximately 780,000 ethnic Albanians who fled the then Serbian province of Kosovo in 1999. The number of dependent refugees under the protection and care of the UNHCR steadily increased from approximately 6 million in 1980 to more than 17 million in 1992. However, during the first decade of the new millennium, the number of refugees and internally displaced persons remained relatively stable.

This situation changed dramatically following the Arab Spring in 2011, especially in Syria, Iraq, Libya, and Yemen, where fierce armed conflict erupted and continues to this day. In March 2017, the number of registered Syrian refugees exceeded 5 million, with the majority residing in Turkey, Jordan, and Lebanon (UNHCR, 2017b). In addition, an estimated 6.6 million Syrians were internally displaced. Other major sources of refugees and internally displaced persons in 2017 were chronically conflict-affected Afghanistan, Somalia, the DRC, and Colombia, as well as countries recently affected by conflict, such as South Sudan, Central African Republic, and Yemen (UNHCR, 2017a). **TABLE 12-1** notes major refugee populations as of December 2016.

TABLE 12-1 Origins of 10 Major Refugee Populations (Plus Palestinians), December 2016

Refugee Origin	Host Country	Number
Palestinians	West Bank, Gaza, Jordan, Lebanon	5,300,000
Syria	Turkey, Lebanon, Jordan, Egypt	5,500,000
Afghanistan	Iran, Pakistan, India	2,500,000
South Sudan	Ethiopia, Uganda, Kenya, Sudan	1,400,000
Somalia	Kenya, Sudan, Yemen, Djibouti	1,000,000
Sudan	Chad, South Sudan	650,000
Democratic Republic of Congo (DRC)	Uganda, Rwanda, Tanzania	537,500
Central African Republic	Cameroon, Chad, DRC	490,900
Myanmar	Bangladesh, Malaysia	490,300
Eritrea	Ethiopia, Sudan, Saudi Arabia	459,400
Burundi	Tanzania, Rwanda, Uganda	408,100

Data from United Nations High Commissioner for Refugees. (2017). *Global trends: Forced displacement in 2016*. Geneva, Switzerland: Author.

In 2016, the countries hosting the largest numbers of refugees were Turkey (2.9 million), Pakistan (1.4 million), Lebanon (1 million), Iran (980,000), Uganda (940,800), Ethiopia (791,600), Jordan (685,000), and Germany (669,000).

Internally Displaced Persons

In addition to those persons who meet the international definition of refugees, millions of people have fled their homes for the same reasons as refugees but remain *internally displaced* in their countries of origin. Ascertaining the number and location of the world's IDPs presents a stiff challenge due not only to definitional difficulties, but also as the result of institutional, political, and operational obstacles. Despite these difficulties, there is a broad consensus that the global population of IDPs has increased from somewhere in the region of 20 million at the end of 2000 to more than 40 million at the end of 2016.

IDPs lack the protection afforded by the international conventions and protocols on refugees. Nevertheless, the Geneva Conventions and certain articles of the United Nations Charter provide some protection to these individuals. Since 2005, UNHCR has been officially responsible for providing assistance to IDPs.

During 2016, 5.5 million IDPs were newly displaced by conflict and violence in their countries, according to data reported by UNHCR. At the end of 2016, the largest IDP populations were in Colombia (7.4 million), Syria (6.3 million), Iraq (3.6 million), DRC (2.2 million), Sudan (2.2 million), Nigeria (2.2 million), Yemen (2.0 million), South Sudan (1.9 million), Ukraine (1.8 million), Afghanistan (1.8 million), and Somalia (1.6 million). **FIGURE 12-2** shows the trend in IDP and refugee numbers between 1996 and 2015.

The final stages of the Sri Lankan civil war created 250,000 internally displaced persons who were transferred to camps in Vavuyina District and detained against their will for many months (see **EXHIBIT 12-2**).

During the 1990s, the United Nations Security Council sanctioned some extraordinary measures to protect displaced populations in southern Sudan, northern Iraq, the republics of the former Yugoslavia, Somalia, and Timor-Leste. There have been few such United Nations interventions since 2000. One notable exception has been the United Nations–African Union Mission in Darfur (UNAMID), which at the end of 2016 had almost 20,000 uniformed personnel from 116 countries.

Until the past decade, the most common response to mass population movements, either across international boundaries or within countries, had been to establish camps or settlements. In Eastern Europe, many refugees and IDPs have been housed in hotels, resort camps, schools, and hostels, where environmental conditions have been relatively good. In contrast, in LMICs, most refugees and displaced persons have been placed in camps located in inappropriate border areas. Conditions in camps have varied enormously; in general, camps with fewer than 20,000 residents have had more favorable environmental conditions than larger camps. Camps for Rwandan refugees in eastern Zaire in 1994 contained up to 300,000 persons; they were poorly planned and laid out, with inadequate sanitation and poor access to

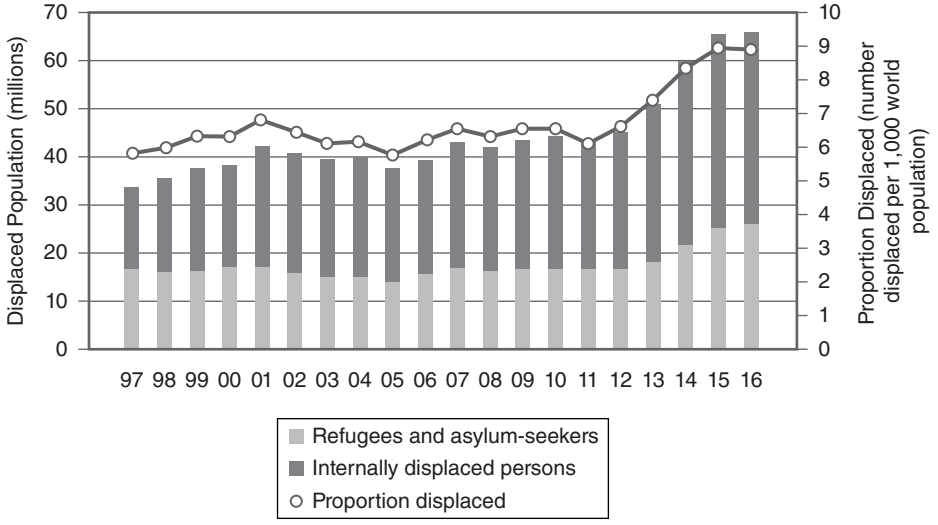


FIGURE 12-2 Estimated global number of refugees and internally displaced persons (millions), 1997–2016.

Reproduced from United Nations High Commissioner for Refugees. (2017). *Global trends: Forced displacement in 2016*. Geneva, Switzerland: Author.

EXHIBIT 12-2 Crisis Traps IDPs in Sri Lanka

During the first half of 2009, Sri Lanka's 26-year conflict in the north of the country intensified. Fighting between the Sri Lankan army and the Liberation Tigers of Eelam (LTTE) trapped approximately 70,000 civilians in a small coastal stretch of land in the north Vanni region. The United Nations estimated that 4,500 civilians were killed and 12,000 injured during the final stages of the conflict. The International Committee of the Red Cross (ICRC) was the only humanitarian aid agency permitted access to the area.

Eventually, LTTE forces were defeated and the fighting ceased. In the immediate aftermath, approximately 250,000 civilians were displaced and held in camps for up to 12 months. In 6 days alone, from May 16 to May 20, 77,000 people emerged from the former conflict zone in northern Sri Lanka and arrived in Vavuniya district, where MSF was one of the aid agencies working. MSF provided surgical and medical support to the 400-bed Vavuniya Hospital, which at one point housed 1,900 inpatients. To ease the congestion, MSF established a 100-bed field hospital. The main cause of hospitalization among the 3,000 patients admitted to the MSF hospital from June to the end of November 2009 was trauma and wounds. In total, more than 1,350 surgical procedures were performed during that time in the MSF hospital.

Data from Medical emergency in Sri Lanka. (2009). *The Lancet*. 373:1399; MSF Australia. (2009, December 14). Field News. Sri Lanka: Situation Retrieved from www.msf.org.au

clean water. Relief program managers found it difficult, if not impossible, to establish equitable systems of distribution of commodities, such as food and shelter materials, and there was a high frequency of violence and other crimes. By contrast, smaller refugee camps in Burundi were more easily managed and the residents there suffered fewer health consequences related to environmental conditions.

In addition to poor environmental conditions, crowded camps promote the spread of many communicable diseases, such as measles, meningitis, and acute respiratory infections. The overwhelming nature of these large camps also tends to create a sense of loss of dignity and independence among mostly rural refugees, and induces mental health disorders such as anxiety and depression.

As noted earlier, in 2016, more than 50% of the world's refugees and IDPs were in urban areas, which has both positive and negative implications for their quality of life and health status. The positives include a diminished risk of communicable diseases; the negatives include difficulties in accessing health services that may rely on user fees. For example, most health services in Lebanon are provided by private practitioners.

Destruction of Public Utilities

Wars often involve the intentional or accidental destruction of public utilities, such as water and sewage systems, electricity sources and distribution grids, and fuel supplies. Although these disruptions mainly affect urban areas, local water supplies have also been destroyed in rural conflicts, such as in Somalia during the early 1990s.

Between 1992 and 1995, in Sarajevo, the capital, and other large cities in Bosnia and Herzegovina, municipal water supplies were destroyed by shelling; similar breakdowns in sewage systems and cross-contamination of piped water supplies led to widespread contamination of drinking water. These problems were compounded by the lack of electricity and diesel fuel needed to run generators.

During a conflict, hospital generators are often able to supply only operating rooms and emergency rooms, further promoting a concentration of services in the area of trauma management. In this kind of environment, routine surgical procedures, inpatient medical care, and pediatric, obstetric and gynecologic, and perinatal care services deteriorate. In addition, the cold chain required to maintain immunization programs is not sustainable.

The Effects of Armed Conflict and Political Violence on Health Services

The model presented in **FIGURE 12-3** offers a framework for describing the health service impact of conflict and CEs. The focus is on the health services within the countries affected by CEs; there are also related pressures and constraints on the health services of host countries to which refugees may flee.

Access to Services

The impact of conflict on health facilities and services depends on their prior availability, distribution, and utilization patterns. Where services

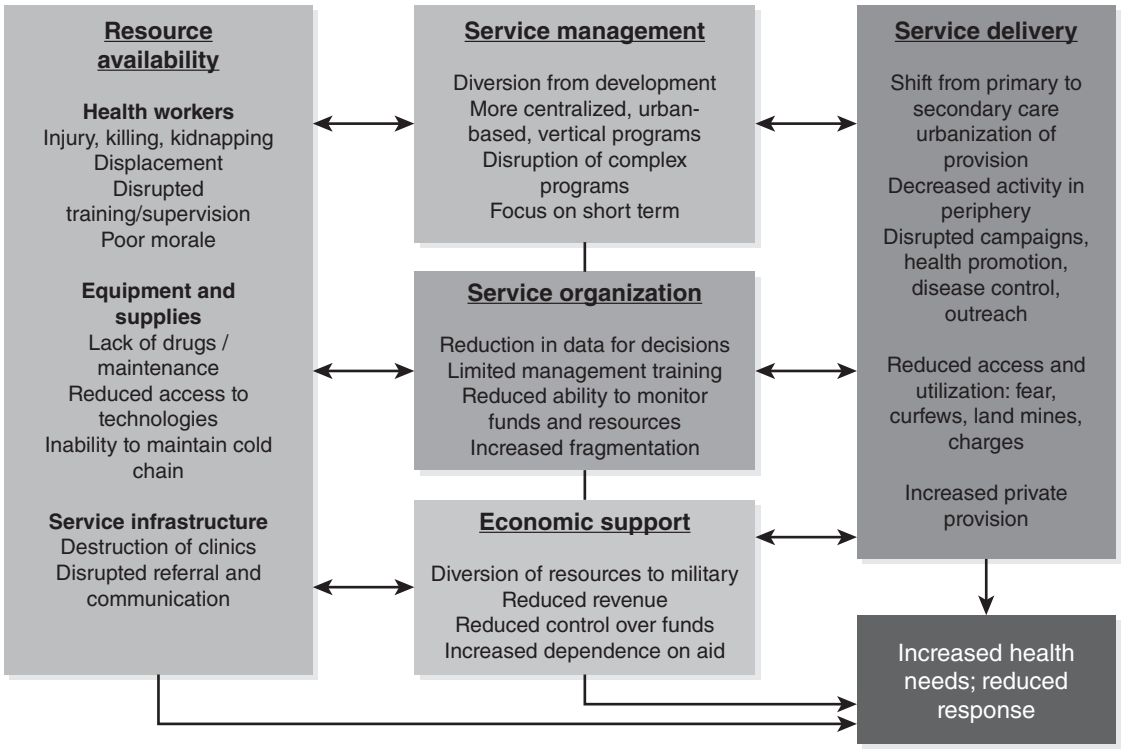


FIGURE 12-3 Impact of conflict and complex emergencies on health.

were originally available, as in Iraq (prior to 1991) and Syria (prior to 2011), the conflict may cause rapid deterioration as a result of infrastructure and distribution systems damage, resource constraints, declining health personnel availability and morale, and reductions in access. The prewar health system in Iraq was extensive, accessible to 90% of the population, and reached 95% of the children requiring immunizations (Lee & Haines, 1991). By the end of the Gulf War, many hospitals and clinics had been severely damaged or closed, those operating were overwhelmed with work, and damage to infrastructure, water supplies, electricity, and sewage disposal exacerbated problems in population health and health services activity.

Six years of civil war in Syria have taken a devastating toll: thousands of lives lost, forced migration at levels unseen since World War II, homes and infrastructure obliterated, health facilities and workers targeted in attacks. Access to basic health care, including surgical care, is severely hampered by restricted ability to move, the destruction and breakdown of health infrastructure, shortage of medical supplies, lack of human resources and threats of kidnappings and killings by different armed groups on medical staff. By the end of March 2015, out of 113 public hospitals in Syria, 44% were reported fully functioning, 36% partially functioning (i.e., shortage of staff, equipment, or medicines, or damage of the buildings in some cases), while 20% were reported to be nonfunctioning (Eastern Mediterranean Regional Office [EMRO], 2015). Reports estimate that more Syrians have died due to health complications resulting from inadequate healthcare services than as a direct consequence of the conflict (European Commission, 2015).

Utilization is determined by geographic access (i.e., the services are not too far away), economic access (i.e., the services are affordable), and social access (i.e., no psychological or other barriers prevent use of services)—all factors that may be disrupted during CEs. Service access may be limited by fear of physical or sexual assault or by physical restrictions on access as a result of antipersonnel land mines, curfews, and, in some cases, the encirclement of areas. In Afghanistan, the Taliban imposed constraints on women accessing services that were previously available to them.

Conflict may seriously disrupt links between services operating at different levels: Referrals will be disrupted by logistical and communication constraints, as well as physical and military barriers to access. Health workers may move to urban areas to seek protection, other opportunities to make a living, or opportunities to provide health services privately with greater financial returns. Towns and cities may be besieged, with entry and exit controlled by militias, as in Aleppo (Syria), Mosul (Iraq), Sanaa (Yemen), and Juba (South Sudan).

Equipment and Supplies

Access to medicines and supplies is typically disrupted during conflicts. Drug shortages, especially where they were previously available, may lead to an increase in medically preventable causes of death, such as asthma, diabetes, and infectious diseases. The Eastern Mediterranean Regional Office of the World Health Organization (WHO) reported from Syria that owing to the reduction of locally produced medicines by 70% of the pre-conflict levels, which led to higher costs for medicines, there were increasingly high numbers of vulnerable people suffering from noncommunicable diseases (NCDs) who could not be adequately treated (EMRO, 2015).

The quality of care available may also suffer greatly during CEs. In Somalia, amputations performed without intravenous antibiotics led to higher rates of infection. In the former Yugoslavia, operations were performed with inadequate anesthesia. Healthcare technologies, including x-rays and laboratories, are undermined through lack of maintenance, spare parts, skilled personnel, chemicals, and other supplies.

Additional problems may emerge as a result of the humanitarian response. Drug donations, if poorly coordinated and standardized, may lead to a large number of expired and inappropriate drugs being offloaded in countries experiencing CEs. These medications may not be usable, but require safe and efficient disposal, placing an additional burden on the recipient country's pharmaceutical services.

Human Resources

Injury, killing, kidnapping, and exodus of health workers are all common events during CEs. Evidence indicates that health workers have been specifically targeted in many recent conflicts; this has been particularly well documented in Syria. Physicians for Human Rights estimates that between 2011 and 2015, at least 610 medical personnel were killed, and there were 233 deliberate or indiscriminate attacks on 183 medical facilities (PHR, 2015). PHR claims that the Syrian government was responsible for 88% of the recorded hospital attacks and 97% of medical personnel killings in that country.

Even if not directly targeted, health workers may flee in search of safety and security. Thousands of physicians once worked in Aleppo, formerly Syria's most populous city, but the assault has resulted in an exodus of 95% of them to neighboring countries and to Europe. Across Syria, millions of civilians have no access to care for chronic illnesses, and the health ministry routinely prevents UN convoys from delivering medicines and surgical supplies to besieged areas.

New Actors

During internal conflicts, due to scarcity of resources and government difficulties in accessing populations under the control of insurgents, NGOs usually fill part of the vacuum left by the public sector. For example, NGOs such as MSF and the Syrian American Medical Society Foundation have been remotely supporting clandestine hospitals and surgical facilities inside Syria. In recent conflict-related emergencies, various military forces have played a direct role in providing relief (e.g., northern Iraq, 1991), as have private companies contracted by government or UN agencies (e.g., Albania and Macedonia, 1999). The entry of these new players has further complicated the response to CEs.

The role of NGOs is extremely important both during and in the aftermath of ongoing conflict. During conflicts, indigenous NGOs and church groups may be among the few service providers that continue to operate during the conflict, especially in rural areas and those more directly affected by violence. A key problem, however, is that these NGOs often provide a patchwork of services that are relatively independent of the state and do not necessarily fit in with other service provision approaches or priorities. They may communicate poorly with one another, adopt different approaches and standards of care and of health worker remuneration, and focus attention mostly at a local level, with some impact on the equity of service availability across large regions.

A key research challenge is to understand how health systems adapt and respond to conflict and to determine whether positive developments can be further reinforced and sustained. Mechanisms to protect and maintain key elements of service provision and functioning, including information systems and supplies, are crucial to assuring ongoing system functioning. How best to promote this state requires further exploration.

Specific Health Outcomes

Mortality

In this section, the impact of civil conflict and humanitarian emergencies on mortality rates is confined to indirect causes, such as food scarcity, population displacement, destruction of health facilities and public utilities, and disruption of routine curative and preventive services. Mortality directly caused by the violence of war was discussed earlier.

The most severe health consequences of conflict—population displacement, food scarcity, and siege situations—have occurred in the acute emergency

phase, during the early stage of relief efforts, and have often been characterized by extremely high mortality rates. Although the quality of the international community's disaster response efforts has steadily improved, death rates associated with forced migration have often remained high, as demonstrated by several emergencies during the 1990s. For example, the exodus of almost 1 million Rwandan refugees into eastern Zaire in 1994 resulted in mortality rates that were more than 30 times the rates experienced prior to the conflict in Rwanda. Most deaths were caused by communicable diseases, such as cholera and dysentery.

In Darfur, Sudan, death rates among civilian populations were as high as 9.5 per 10,000 per day (20 times higher than baseline mortality rates) in 2003, with most deaths being due to violence (Depoortere et al., 2004). However, a study published in 2010 reviewed 63 retrospective mortality surveys conducted after the Depoortere study, spanning from early 2004 to the end of 2008 (Degomme & Guha-Sapir, 2010). The highest CMR was 4 per 10,000 per day in early 2004, when the most common cause of death was still injuries due to violence. The study showed significant reductions in mortality rates from early 2004 to the end of 2008, although rates were higher during deployment of fewer humanitarian aid workers. Also, the main causes of death evolved from violence to communicable diseases. The study estimated that almost 300,000 excess deaths related to conflict and displacement occurred during the 4-year study period.

CMRs have been estimated from burial site surveillance; administrative, hospital, and burial records; community-based reporting systems; and population surveys. The many problems in estimating mortality under emergency conditions have included the following:

- Poorly representative or inaccurate population sample surveys
- Failure of families to report all deaths for fear of losing entitlements
- Inaccurate estimates of affected populations for calculating mortality rates
- Lack of standard reporting procedures

In general, however, mortality rates may be underestimated because deaths are usually underreported or undercounted, and population size is often exaggerated.

Early in an emergency, when mortality rates are elevated, the CMR is usually expressed as deaths per 10,000 population per day (CDC, 1992). The median annual CMR in developing countries is approximately 9 per 1,000, corresponding to a daily rate of approximately 0.25 per 10,000 (Reed & Keeley, 2001). A threshold of 1 per 10,000 per day typically has been used to define an elevated CMR and to characterize a situation as an emergency (CDC, 1992). In one of the most severe refugee emergencies of the 1990s, the CMR among Rwandan refugees during the first month after their arrival in eastern Zaire was between 27 and 50 per 10,000 per day (Goma Epidemiology Group, 1995).

The most reliable estimates of mortality rates have come from well-defined and secure refugee camps where there is a reasonable level of camp organization and a designated agency has had responsibility for the collection of data (**TABLE 12-2**). The most difficult situations have been those where IDPs have been scattered over a wide area and where surveys could take place only in relatively secure zones (**TABLE 12-3**). These safe zones have sometimes acted as magnets for the most severely affected elements of a population. For example, in 1998, a survey in Ajiep, southern Sudan, found that the CMR increased from 17.8 per 10,000 per day during the period June 3 to July 11 to 69.7 per 10,000 per day between July 12 and July 20 (Brown, Moren, & Paquet, 1999). On the one hand, this increase may have been due to an influx of displaced persons reaching Ajiep in a poor condition or to a decrease in the food available within the town. On the other hand, it is possible that the worst affected communities were in areas that were inaccessible to those performing the surveys. In either case, it has proved difficult to extrapolate the findings of surveys on mortality conducted in specific locations to broader populations in conflict-affected countries.

Extensive differences in mortality survey methods have also been identified. For example, an evaluation of 23 field surveys performed in Somalia between 1991 and 1993 found wide variation in the target populations, sampling strategies, units of measurement, methods of rate calculation, and statistical analysis (Boss, Toole, & Yip, 1994).

Since the mid-1990s, CMRs among refugees have tended to be relatively low. In fact, a study found that in conflict-affected emergencies, CMRs are highest among IDPs, followed by nondisplaced residents, followed by refugees who fled to other countries (Heudtlass, Speybroeck, & Guha-Sapir, 2016).

Trends in death rates over time have varied from place to place. In refugee populations where the international response has been prompt and

TABLE 12-2 Estimated Daily Crude Mortality Rates (Deaths per 10,000 per Day) in Selected Refugee Populations, 1991–2013

Period	Country of Asylum	Country of Origin	Mean CMR for Period
June 1991	Ethiopia	Somalia	2.3[a]
March–May 1991	Turkey	Iraq	4.7[b]
March–May 1991	Iran	Iraq	2.0[c]
March 1992	Kenya	Somalia	7.4[a]
March 1992	Nepal	Bhutan	3.0[d]
August 1992	Zimbabwe	Mozambique	3.5[e]
December 1993	Rwanda	Burundi	3.0[f]
August 1994	DRC	Rwanda	19.6–31.3[g]
May 1999	Albania	Yugoslavia (Kosovo)	0.5[h]
November 1999	Indonesia (Tuapukan Camp, West Timor)	East Timor	2.1[i]
September 2004	Chad	Darfur, Sudan	1.3[j]
July–September 2008	Chad (19 camps)	Darfur, Sudan	0.2–0.6[k]
November 2009	Syria	Iraq	0.2[l]
October 2013	Jordan (Za'atri Camp)	Syria	0.1[m]

Abbreviation: CMR = crude mortality rate.
[a] Toole & Waldman (1993).
[b] CDC (1991b).
[c] Babille, de Colombani, Guerra, Zagaria, & Zanetti (1994).
[d] Marfin et al. (1994).
[e] CDC (1993b).
[f] CDC (1994).
[g] Goma Epidemiology Group (1995).
[h] UN ACC/SCN (1999).
[i] WHO Health Information Network for Advanced Planning (2000a).
[j] UN Standing Committee on Nutrition (2004d).
[k] UN Standing Committee on Nutrition (2009).
[l] UN Standing Committee on Nutrition (2010).
[m] Interagency Regional Response for Syrian Refugees (2013).

effective, as was the case for Cambodians in eastern Thailand (1979) and Iraqis on the Turkish border (1991), death rates declined to baseline levels within 1 month. By comparison, among refugees in Somalia (1980) and Sudan (1985), death rates were still well above baseline rates 6 to 9 months after the influx of refugees occurred (Toole & Waldman, 1990). In the case of 170,000 Somali refugees in Ethiopia in 1988–1989, death rates actually increased significantly 6 months after the influx. This increase was associated with elevated malnutrition prevalence rates, inadequate food rations,

TABLE 12-3 Estimated Crude Mortality Rates (Deaths per 10,000 per Day), Internally Displaced Populations, 1991–2017

Period	Country (Region)	CMR for Period
April 1991–March 1992	Somalia (Merca)	4.6[a]
April–November 1992	Somalia (Baidoa)	16.9[b]
April 1992–March 1993	Sudan (Ayod)	7.7[c]
April 1993	Bosnia and Herzegovina (Sarajevo)	1.0[d]
May 1995	Angola (Cafunfo)	8.3[e]
February 1996	Liberia (Bong)	5.5[f]
May 1998	Burundi (Cibitoke)	3.3[g]
June 3–July 20, 1998	Sudan (Ajiep)	26.0[h]
September 2001	Sierra Leone (Kono district)	1.4[i]
June 2002	Angola (Muacanhica, Muahimbo, Luena)	2.9–7.2[j]
February 2002	Sudan (Jonglei)	7.2[j]
April 2003	DRC (Katanga)	1.9[k]
April 2004	Sudan (Wade Saleh, Darfur)	3.6[l]
May 2008	Somalia (Afgoi and Merka)	1.0[m]
June 2008	South Darfur, Sudan (four sites)	0.75[n]
May–July 2009	North Darfur, Sudan (five sites)	0.3–0.7[o]
November 2013–February 2014	Central African Republic	6.0[p]
February–March 2017	South Sudan (all IDP camps)	<1[q]

Abbreviation: CMR = crude mortality rate.
[a] Manoncourt et al. (1992).
[b] Moore et al. (1993).
[c] CDC (1993c).
[d] CDC (1993a).
[e] UN Administrative, Committee on Coordination, Sub-Committee on Nutrition (ACC/SCN) (1995).
[f] UN ACC/SCN (1996).
[g] UN ACC/SCN (1998).
[h] Epicentre and MSF France, in UN ACC/SCN (1999).
[i] UN ACC/SCN (2002a).
[j] UN ACC/SCN (2002b).
[k] UN ACC/SCN (2003).
[l] UN Standing Committee on Nutrition (2004b).
[m] UN Standing Committee on Nutrition (2009a).
[n] UN Standing Committee on Nutrition (2009b).
[o] UN Standing Committee on Nutrition (2010).
[p] Coldiron et al. (2017).
[q] WHO (2017a).

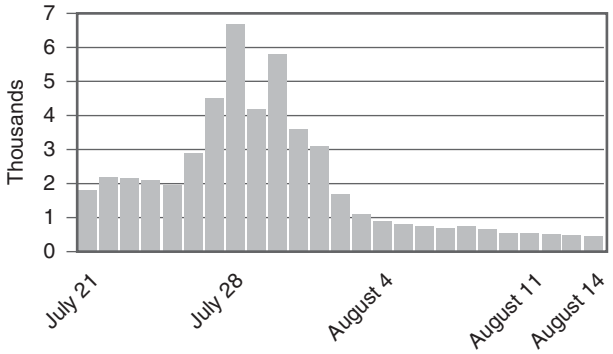


FIGURE 12-4 Number of deaths per day, July 21 to August 14, 1994, Rwandan refugees, North Kivu Camps, Zaire.

Reproduced with permission from Goma Epidemiology Group. (1995). Public health impact of Rwandan refugee crisis: What happened in Goma, Zaire, in July 1994? *The Lancet,* 345, 339–344.

and high incidence rates of certain communicable diseases (Toole & Bhatia, 1992). Although initial death rates among Rwandan refugees in eastern Zaire were extremely high, they declined dramatically within 1 to 2 months (**FIGURE 12-4**). Surveys in Darfur, Sudan, in 2004 showed that mortality was higher in the predisplacement period and generally declined once people reached refugee or IDP camps (Depoortere et al., 2004).

Most deaths have occurred among children younger than 5 years of age; for example, 65% of deaths among Kurdish refugees on the Turkish border occurred in the 17% of the population younger than 5 years (Yip & Sharp, 1993). However, in some refugee situations, such as Goma during the first month after the refugee exodus, mortality rates were comparable in all age groups because the major cause of death was cholera, which is equally lethal at any age. Among IDPs in countries affected by severe famine, high adult mortality has been reported. For example, in the Somali town of Baidoa, 59% of 15,105 deaths reported between August 1992 and February 1993 were among adults (Collins, 1993). In Ajiep, southern Sudan, the CMR among IDPs in August 1998 was equal to the under-5 mortality rate (Salama, Spiegel, & Brennan, 2001).

In most reports from refugee camps, mortality rates have not been stratified by sex. The surveillance system for Burmese refugees in Bangladesh did estimate sex-specific death rates, demonstrating considerably higher death rates in females. Gendered analyses that take into account differences in the sociocultural position of women have been rare in emergency settings.

In the 1980s and 1990s, the major reported causes of death among refugees and displaced populations were diarrheal diseases, measles, acute respiratory infections, and malaria, exacerbated by high rates of malnutrition. These diseases consistently accounted for between 60% and 95% of all reported causes of death in these populations. Measles epidemics caused high death rates among refugees during the 1980s. Epidemics of severe diarrheal disease have been increasingly common and contributed to high mortality. Cholera case-fatality ratios (CFRs) in refugee camps have ranged between 3% and 30%, and dysentery CFRs have been as high as 10% among young children and the elderly.

In eastern European conflicts, a high proportion of mortality among civilians was caused by trauma associated with the violence. Nevertheless, there was also increased mortality in these conflicts due to the collapse of the public health system. Chronic conditions, such as cardiovascular diseases, cancer, and renal conditions, were inadequately treated because the health system was focused on the management of war-related injuries. Preventive health services, including childhood immunization and antenatal care, ceased in many areas.

Over the past two decades, CEs have occurred without excess indirect mortality being reported among displaced populations. For example, mortality rates among Kosovar refugees in Albania and Macedonia remained less than 1 per 10,000 per day. Likewise, mortality rates among Syrian refugees who fled to Turkey, Jordan, and Lebanon have been below the emergency threshold.

Nutrition

Nutritional deficiencies are often the first public health effects of an evolving CE. Among refugees and IDPs, many factors might lead to high rates of nutritional deficiency disorders, including prolonged food scarcity prior to and during displacement, delays in the provision of complete rations, problems with registration and estimation of the size of an affected population, and inequitable distribution systems.

In the emergency phase, acute energy depletion is a life-threatening condition and leads to excess mortality. A critical factor is the synergy between malnutrition and infection; thus, malnutrition prevalence may be increased by high rates of infectious diseases, such as measles, diarrhea, dysentery, acute respiratory infections, malaria, and helminth infestation. Infections, by contrast, lead to decreased appetite and increased metabolic rate, which exacerbate acute malnutrition. These factors may differentially affect certain demographic groups within the population. The most vulnerable groups typically include children younger than 5 years of age, pregnant and lactating women, the elderly, unaccompanied children,

the disabled, the chronically ill (e.g., tuberculosis and human immunodeficiency virus [HIV]–infected patients), households lacking an adult male, and disadvantaged ethnic or religious groups. In industrialized countries, the elderly are often most vulnerable, especially those living alone on fixed incomes.

Malnutrition

In estimating the prevalence of acute malnutrition in a population, the prevalence among children between 6 months and 5 years of age is typically used as a surrogate. This approach is employed because the relationship between weight and height in this age group is generally similar in all ethnic groups, provided those children have access to adequate food. International reference tables developed by WHO (2009) are widely used to define various degrees of acute malnutrition.

The prevalence of moderate to severe acute malnutrition in a random sample of children between 6 months and 5 years of age (or 110 cm in height) is generally a reliable indicator of this condition in a population. Since weight is more sensitive to sudden changes in food availability than height, nutritional assessments during emergencies focus on measuring weight for height. Moderate to severe acute malnutrition is defined as a weight-for-height measurement more than 2 standard deviations below the mean of the WHO reference population (*Z*-score less than –2) (WHO, 2009). Severe acute malnutrition is defined as weight for height more than 3 standard deviations below the reference mean (*Z*-score less than –3). All children with bilateral pitting edema are classified as having severe acute malnutrition.

As a screening measurement, the middle upper arm circumference (MUAC) may also be used to assess acute malnutrition, although there is not complete agreement on which cutoff values should be used as indicators. Field studies indicate that a MUAC between 12.0 cm and 12.5 cm correlates with a weight-for-height *Z*-score of –2; the lower figure (12.0 cm) is more appropriate in children younger than 2 years of age (WHO, 2009).

Acute malnutrition without edema has been termed wasting or marasmus, and acute malnutrition with edema has been termed kwashiorkor; however, a combination of the two may occur in some children. Both are associated with anemia, though this condition is often more severe in children (or adults) with kwashiorkor. The anemia may be exacerbated by local conditions, such as malaria and hookworm infection.

In 2001, the United Nations Sub-Committee on Nutrition convened a meeting in Nairobi to review the assessment of nutritional status in adults and adolescents. The attendees recommended that assessment of adult malnutrition should be considered under the following circumstances:

- If the crude mortality rate increases in relation to the under-5 mortality rate
- If many adults are present at existing supplementary feeding centers
- Where there are high rates of under-5 malnutrition in the absence of an epidemic outbreak
- If there is reasonable doubt that the child malnutrition rates do not reflect the nutritional status of the general population
- If the populations are entirely reliant on food aid and if data are required as an advocacy tool to leverage resources

MUAC is commonly used as a screening tool in adult populations. A MUAC of less than 23 cm in men and less than 22 cm in women has been proposed as indicating moderate malnutrition (corresponding to body mass index [BMI] values of less than 17). MUAC values of less than 20 cm in men and less than 19 cm in women have been proposed as indicators of severe wasting, corresponding to BMIs less than 13 (Ferro-Luzzi & James, 1996). Nevertheless, one reliable clinical sign of severe malnutrition commonly observed in adults is inability to stand.

Prevalence of acute malnutrition among the internally displaced has tended to be extremely high. In southern Somalia during 1992, the prevalence of acute malnutrition among children younger than 5 years in displaced-persons camps in Marka and Qorioley was 75%, compared with 43% among town residents (Manoncourt et al., 1992). Among the estimated 1 million IDPs in Darfur, Sudan, acute malnutrition prevalence ranged between 12.6% and 21.5% in March–April 2004, increasing to 20.6% to 39% in May–June 2004. Following heightened international interest and assistance to this population, rates of malnutrition generally decreased to a range of 10.7% to 23.6% by September 2004. However, 5 years later, after many humanitarian aid agencies had been expelled from Sudan, malnutrition rates increased significantly. In northern Darfur, acute malnutrition prevalence ranged between 16.9% and 34.5% according to surveys conducted between May and July 2009. (**EXHIBIT 12-3** discusses malnutrition among Sudanese refugees.) The prevalence of acute malnutrition among children younger than 5 years of age in various displaced populations is presented in **TABLE 12-4**.

In South Sudan, where 1.9 million people have been internally displaced and 1.6 million have fled to other countries, the malnutrition prevalence has been critically high. An estimated 5.5 million people in

EXHIBIT 12-3 Sudanese Refugees in Chad, 2004–2010

As a result of the armed conflict in Darfur, Sudan, that began in early 2003, approximately 100,000 refugees fled to Chad during 2003. Initially, they lived in makeshift camps and received little international assistance. Early nutrition surveys indicated great variation in nutritional status of refugees between settlements; in November 2003, rates of acute malnutrition among children ranged from 7% to 27% (UN Standing Committee on Nutrition, 2004a). By May 2004, the camps were overcrowded, water and sanitation were inadequate, and an outbreak of the waterborne viral disease hepatitis E had occurred in some camps. In June, a joint agency survey found the overall acute malnutrition prevalence among children to be 35.6%, with 5.5% severely malnourished (UN Standing Committee on Nutrition, 2004c).

By the end of 2004, more than 200,000 Darfur refugees had fled to Chad. International media attention increasingly focused on both the conflict in Darfur and the plight of refugees and IDPs. International assistance increased, and by the end of 2004, most refugees were located in 11 camps run by UNHCR. Blanket supplementary feeding was introduced for all children younger than age 5 and for all pregnant and lactating women. According to food basket monitoring, the average energy content of the food ration reached 1,967 kilocalories per person per day by October 2004 (compared with an official ration of 2,063 kilocalories). A September survey in two camps found that the prevalence of acute malnutrition had dropped to 19.6%; although this represented an improvement, the rate remains high by international standards (UN Standing Committee on Nutrition, 2004d).

Five years later, the CMR in all refugee camps in Chad was less than 1 per 10,000 per day. The prevalence of acute malnutrition ranged between 4% and 10%. However, the prevalence of acute malnutrition in IDP camps across the border in Sudan remained high; between May and July 2009, it was between 17% and 34% in six camps in northern Darfur.

South Sudan (47% of the population) were projected to be severely food insecure at the height of the 2017 lean season between May and July 2017, and more than 1.1 million children were estimated to be acutely malnourished.

Micronutrient Deficiency Diseases

High incidence rates of several micronutrient deficiency diseases have been reported in many refugee camps, especially in Africa. Frequently, famine-affected and displaced populations have already experienced low levels of dietary vitamin A intake and, therefore, may have low vitamin A reserves. Furthermore, the typical rations provided in large-scale relief operations lack vitamin A, putting these populations at high risk of this micronutrient deficiency. In addition, those communicable diseases that are highly incident in refugee camps, such as measles and diarrhea, are known to rapidly deplete vitamin A stores. Consequently, young refugee and displaced children are at high risk of developing vitamin A deficiency.

In 1990, more than 18,000 cases of pellagra, caused by food rations deficient in niacin, were reported among Mozambican refugees in Malawi (CDC, 1991a). Despite the increased awareness of micronutrient deficiencies among relief agencies, niacin was not a component of the general food ration for IDPs in Angola in 2000. A large outbreak of pellagra was documented, with attack rates among IDPs more than twice those observed in the nondisplaced population (Salama et al., 2001).

Numerous outbreaks of scurvy (vitamin C deficiency) were documented in refugee camps in Somalia, Ethiopia, and Sudan between 1982 and 1991. Cross-sectional surveys performed in 1986–1987 reported prevalences as high as 45% among females and 36% among males; prevalence increased with age (Desenclos et al., 1989). The prevalence of scurvy was highly associated with the period of residence in camps, a reflection of the time exposed to rations lacking in vitamin C. Outbreaks of scurvy and beriberi were also reported among Bhutanese refugees in Nepal during 1993 (UN ACC/SCN, 1995).

Iron-deficiency anemia has been reported in many refugee populations, affecting particularly women of childbearing age and young children. For example, a survey among IDPs in Darfur, Sudan, in 2004 found that 55% of children younger than 5 years were anemic, as well as 26.2% of women of reproductive age (UN Standing Committee on Nutrition, 2004d).

Impact of Communicable Diseases

In most CEs in LMICs, the high rates of excess preventable mortality have been attributed primarily to communicable diseases. Acute respiratory infections, diarrhea, measles, and malaria have been most frequently cited as proximate causes. The substandard conditions found in refugee camps do not change the diseases that account for most of the morbidity and mortality in humanitarian emergency settings, but they do alter epidemiologic patterns in two important ways: The incidence, or attack, rates of commonly

TABLE 12-4 Prevalence of Acute Malnutrition Among Children Younger Than 5 Years in Internally Displaced and Conflict-Affected Populations, 1992–2004

Date	Country (Region)	Prevalence of Acute Malnutrition (%)
1992	Southern Somalia	47–75[a]
1994	Sudan (Bahr el Ghazal)	36.1[b]
1994	Afghanistan (Sarashahi)	18.6[c]
1995	Angola (Cafunfo)	29.2[d]
1995	Sierra Leone (Bo)	19.8[d]
1996	Zaire (Masisi)	31.0[e]
1999	West Timor, Indonesia	24[f]
2003	Hiran Region, Somalia	17.1[g]
2008	North Darfur, Sudan	14–17[h]
2009	Bandudu Province, DRC	7–20[i]
2012	Central African Republic	21[j]
2016	Maiduguri, northeast Nigeria	20[k]
2017	Mayendit County, South Sudan	27%[l]

Note: Acute malnutrition is defined as weight for height 2 standard deviations below the reference mean.
Data from
[a] Toole & Waldman (1993).
[b] Médecins sans Frontières (MSF) Belgium, in UN ACC/SCN (1994).
[c] MSF Holland, in UN ACC/SCN (1994).
[d] Action Contre la Faim (ACF), in UN ACC/SCN (1995).
[e] MSF Holland, in UN ACC/SCN (1996).
[f] World Health Organization, HINAP (2000a, 2000b).
[g] UN Administrative Committee on Coordination, Sub-Committee on Nutrition (ACC/SCN) (2003).
[h] UN Standing Committee on Nutrition (SCN) (2009a).
[i] Standing Committee on Nutrition (SCN) (2010).
[j] ReliefWeb (2012).
[k] Loewenberg S. (2017).
[l] USAID (2017).

occurring and potentially fatal diseases are increased, and the case-fatality ratios are higher than usual.

Individual communicable diseases commonly affecting emergency-affected populations are discussed in the section of this chapter dealing with the prevention and control of communicable diseases.

Reproductive Health

Unfortunately, reproductive health services for refugees and displaced persons have often been considered to be secondary priorities, although increased attention has been paid to reproductive health issues in the past 10 years. Although clearly the provision of food, water, sanitation facilities, and shelter are the highest priorities during a complex humanitarian emergency, steps should be taken to ensure that other critical health needs of women, men, and adolescents are met as quickly as possible. Women are a particularly vulnerable subset of the population because the gender-based discrimination that is all too common in stable societies is frequently exacerbated in times

of societal stress and meager resources. Uncontrolled violence and its aftermath are characterized by a number of specific features that negatively affect reproductive health—namely, the breakdown of family networks and the consequent loss of protection and safety, as well as the lack of channels by which to provide information to adolescents and women of reproductive age.

The Guttmacher Institute reviewed maternal deaths that occurred in 2008–2010 in 25 refugee camps in 10 countries (Hynes, Sakani, Spiegel, & Cornier, 2012). Reports were available on 108 deaths, including 68 deaths in Kenya. In every country but Bangladesh, maternal mortality ratios were lower among refugees than among the host population. The proportion of women who had had four or more antenatal care visits was lower among refugee women who had died (33%) than among the general refugee population (79%). Approximately 78% of the maternal deaths followed delivery or abortion, and 56% of those deaths occurred within 24 hours. Delays in seeking and receiving care were more prevalent than delays in reaching care. In Kenya, delays in seeking or accepting care and provider failure to recognize the severity of the woman's condition were the most common avoidable contributing factors.

A minimum initial package of essential reproductive health services (described later in this chapter) has been developed and is recommended by the major relevant international agencies. Interventions beyond this essential package require major investments of time and personnel who should not be diverted from the principal task of reducing excessive preventable mortality as rapidly as possible. In all cases, special care must be taken to ensure that women heads of household are being given equitable quantities of food and nonfood commodities for themselves and their families.

Noncommunicable Diseases

Prior to the last decade, responding to the health needs of refugees and IDPs traditionally emphasized the direct causes of ill health, such as firearms and other weaponry, as well as communicable diseases and nutritional deficiencies, which have been shown to pose a major problem in many CEs. More recently, and in light of the aging of the population generally and of the changing geographic distribution of conflicts to include areas previously well served by health care, such as Syria and Iraq, new problems have emerged.

Noncommunicable diseases are widespread in all populations of older adults worldwide. In those situations where medical care was at some stage available, the withdrawal or destruction of medical facilities and drug distribution mechanisms, and the withdrawal of health workers from areas of active conflict, all have an impact on the treatment and care available for NCDs. Examples of reductions in the quality and availability of care for NCDs come from Syria since the onset of civil war in 2011, from Iraq in the aftermath of the Gulf War and disruptions by ISIL, from Sarajevo during its siege in the Yugoslav civil war, and from the Kosovar Albanians, especially those internally displaced people who lost access to services and care.

Patients with conditions such as cardiovascular disease (including hypertension), diabetes, asthma, and cancer may deteriorate given the lack of access to medical care that typically occurs in conflict settings. A systematic review in 2014 found that, overall, the prevalence of NCDs was high among urban refugees in the Middle East, ranging from 9% to 50%; by comparison, the prevalence among urban refugees in Asia and Africa ranged from 1% to 30% (Amara & Aljunid, 2014). Prior to the conflict, 77% of deaths in Syria were due to NCDs. The prevalence of diabetes in Syria was 8.9% (adults age 20–79) prior to the conflict and the prevalence of hypertension was 28% (Amara & Aljunid, 2014). More than half of public hospitals in Syria are now either partially functional or completely out of service, constraining the treatment of patients with NCDs. Moreover, local production of medicines has declined by 70% associated with the increasing cost of medicines for treatment of NCDs.

Maintaining diagnostic and treatment services, drug supplies, and access to care is extremely difficult given the destruction of infrastructure, targeting of health services, disruption of logistics and supply systems, and absolute resource constraints present in many CEs. In some conflicts, the imposition of sanctions may play some role in reducing access to technologies and drugs necessary for the diagnosis, treatment, and care of NCDs.

Mental Health

War and political violence have direct and indirect mental health consequences for victims, relatives, neighbors, and communities. The severity and type of mental health problems reflect the nature, intensity, and form of the violence; the relationship of the assessed person to others affected—self, family, and community members; the cause of the conflict; and the affected person's relationship to participation, victimization, or causation of the conflict. Anxiety, uncertainty, and fear about the future and about

whether family members and homesteads remain alive and intact are significant sources of distress for affected individuals and communities. Among those who are forced to flee either as refugees or as internally displaced people, the lack of knowledge about relatives and property left behind creates stress and distress. Despite the ongoing challenges of maintaining lives and livelihoods, life as a refugee, especially in a camp situation, may be monotonous and conducive to stress, anxiety, and depression. Finally, it is clear, and little disputed, that some individuals who experience particularly horrific experiences as victims of torture or gross human rights abuses during conflicts may suffer from post-traumatic stress disorder.

Few national surveys of the mental health impact of conflicts are available, and even if such studies could be conducted, prior measurement of the distribution of mental health status within the population would be required to assess the CEs' impact. A number of small studies focused on particular subgroups of conflict-affected populations have been conducted, but their biomedical biases leave them open to challenge. Little is known about the etiology of the symptoms of multiple trauma, the mental health consequences suffered by those who victimize, and the role of coping mechanisms.

In emergencies, not everyone has or develops significant psychological problems. Many people show resilience—that is, the ability to cope relatively well in situations of adversity. Numerous social, psychological, and biological factors interact to influence whether people develop psychological problems or exhibit resilience in the face of adversity.

Depending on the emergency context, particular groups of people may be at increased risk of experiencing social and/or psychological problems. Although many key forms of support should be available to the emergency-affected population in general, good programming specifically includes the provision of relevant supports to the people at greatest risk, who need to be identified for each specific crisis.

All subgroups of a population can potentially be at risk of mental health disorders, depending on the nature of the crisis. The following are groups of people who frequently have been shown to be at increased risk of various problems in diverse emergencies (Inter-Agency Standing Committee, 2007):

- Women (e.g., pregnant women, mothers, single mothers, widows and, in some cultures, unmarried adult women and teenage girls)
- Men (e.g., ex-combatants; idle men who have lost the means to take care of their families; young men at risk of detention, abduction, or being targets of violence)
- Children (from newborn infants to youths 18 years of age), such as separated or unaccompanied children (including orphans), children recruited or used by armed forces or groups, trafficked children, children in conflict with the law, children engaged in dangerous labor, children who live or work on the streets, and undernourished and/or under-stimulated children
- Elderly people (especially when they have lost family members who were caregivers)
- People who have been exposed to extremely stressful events/trauma (e.g., people who have lost close family members or their entire livelihoods, rape and torture survivors, witnesses of atrocities)
- People in the community with preexisting, severe physical, neurologic, or mental disabilities or disorders
- People in institutions (e.g., orphans, elderly people, people with neurologic/mental disabilities or disorders)
- People experiencing severe social stigma (e.g., "untouchables," commercial sex workers, people with severe mental disorders, survivors of sexual violence)

Prevention and Mitigation of Complex Emergencies

The prevention of CEs is primarily the prevention of the conflicts that cause them; thus, the task is largely political. Since 1990, many CEs have had their roots in ethnic and religious conflicts within sovereign states. The United Nations Charter has been ill equipped to intervene in issues deemed to be "internal" by member states. Chapters 6 and 7 of the Charter do allow the Security Council to authorize appropriate action, including the use of force, in situations that threaten international peace and security. During the Cold War, these provisions were rarely used because such action was likely to be vetoed by one of the five permanent members of the Security Council. However, Security Council resolutions supported intervention by the international community to protect civilians in conflicts in Somalia, Bosnia and Herzegovina, Haiti, Iraq, Angola, and Timor-Leste during the 1990s. Since 2001, such international consensus has been rare. Armed interventions by the United States and its allies in Afghanistan and Iraq were not conducted under the auspices of the United Nations but rather through the North Atlantic Treaty Organization (NATO).

In general, the international community has had little success in resolving internal conflicts by using diplomacy. A major victory for these efforts occurred in 2005 when the government of Sudan and the Sudan People's Liberation Front signed a peace agreement in Nairobi. This led eventually to the establishment of the state of South Sudan. Similar peace agreements have been signed since 2001 in Sri Lanka, Angola, Liberia, Sierra Leone, and the DRC (although the last was still fragile in 2017).

The United Nations has a poor record in conflict resolution. The internal conflict in Somalia was allowed to evolve over five years into the total disintegration of the nation state. Only when famine reached appalling levels in 1992 did the Security Council authorize extraordinary action to ensure the protection and care of the civilian population. Within 6 months of their arrival, UN troops became embroiled in the conflict itself, taking sides with or against certain armed factions. This engagement led to heavy loss of life and the eventual withdrawal of the UN forces. Twenty-five years later, Somalia still lacks a functional government and the self-proclaimed Somaliland in the north of the country is essentially independent, though not internationally recognized.

In Bosnia and Herzegovina, between 1992 and 1995, the United Nations mobilized peacekeeping troops to safeguard the delivery of humanitarian supplies. However, these forces were not authorized to intervene to protect civilians from the violence intentionally directed at various ethnic groups; as a result, the international community's armed representatives were forced to silently witness gross abuses of human rights. This dilemma has been termed the "humanitarian trap."

When the genocide began in Rwanda in 1994, several Belgian peacekeepers already in the country with a UN contingent were killed by extreme Hutu nationalists. Instead of increasing the level of UN presence, the entire peacekeeping force was withdrawn, leaving civilians defenseless against these extremists. Eventually, more than half a million Rwandans were killed. It is likely that this string of UN failures led to NATO taking unilateral action in the form of a massive bombing campaign against Serbia to protect ethnic Albanians in 1999. Although well intentioned, this action seemed to accelerate the pace of atrocities and ethnic cleansing instigated by the Serbs against the Kosovar Albanians.

The basis of protection of civilians in time of conflict is the Geneva Conventions of 1949 and the Additional Protocols of 1977. In addition, the 1951 Convention on the Prevention and Punishment of the Crime of Genocide was intended to protect civilians from the type of slaughters that occurred in Cambodia in the 1970s and in Rwanda in 1994. General Assembly resolutions in 1971, 1985, and 1986 also elaborated on the protection of civilian populations. However, international human rights and humanitarian (armed conflict) laws are only as good as their enforcement. The Office of the UN High Commissioner for Human Rights in Geneva, Switzerland, is one of the primary official bodies that oversees this sometimes overwhelming task, relying on the documentation and testimony provided by accredited NGOs. Today, UNCHR passes influential resolutions on human rights abuses in member states and is the subject of intense lobbying by governments and NGOs. It is thanks to the efforts of a number of human rights advocacy groups that resolutions continue to be adopted by the commission condemning human rights violators. Such resolutions occasionally lead to action by the General Assembly and Security Council, which may dispatch peacekeeping forces to a troubled region.

Between 2004 and 2008, a number of UN Security Council resolutions demanded that the government of Sudan remove its support for the genocidal militia in Darfur. Finally, in 2008 the Security Council authorized a United Nations–African Union hybrid operation in Darfur (UNAMID). In December 2016, there were 19,248 uniformed personnel from 116 countries deployed in the region, though peace and well-being for the displaced population remain elusive.

Since 2008, the United Nations has been unable to resolve any armed conflicts within states. The international response to these conflicts has mainly been military, with various coalitions bombing different armed factions in Syria, Iraq, Yemen, and Somalia. The UN Security Council has been unable to reach a consensus on resolving the vicious civil war in Syria. While France intervened to bring about an uneasy peace in Mali in 2013, a vicious armed conflict continues to rage in nearby Central African Republic.

The UN Office for the Coordination of Humanitarian Affairs (OCHA), based in New York and Geneva, is responsible for coordinating efforts in early warning, prevention, mitigation, and response to disasters, including CEs. ReliefWeb's purpose is to strengthen the capacity of the humanitarian relief community through the timely dissemination of reliable information on prevention, preparedness, and disaster response. As with many other similar projects, this information is made available via the Internet (www.reliefweb.int).

Early Warning and Detection

Efforts to prevent and mitigate the impact of CEs on populations must rely on accurate and timely information to be effective. Given the enormous cost of military intervention and major relief and rehabilitation programs, it is surprising that so little has been invested in early warning, emergency detection, preparedness, and mitigation projects. In the late 1990s, a number of systems collected, aggregated, and disseminated information on a number of indicators relevant to CEs. Most of these systems relied on information collected by other agencies, such as governments, UN agencies, and NGOs. ReliefWeb, a project of OCHA, was mentioned earlier as an exemplar.

Unlike natural disaster early warning systems and preparedness programs, monitoring and detecting CEs is fraught with political obstacles. Indeed, the adage that "the first casualty in war is truth" applies to attempts to collect accurate data on the health outcomes of war. The existence of armed conflicts is no secret; the political response is still inadequate. What would be valuable as CEs evolve would be a more accurate picture of which health interventions will be the highest priorities and the most effective in preventing excess mortality and morbidity. Stopping the violence is a public health issue that can be addressed only by the world's leaders. In LMICs, the priorities are most likely to be nutrition and communicable disease control. Thus, key indicators to monitor in early warning systems include food availability, nutritional status, immunization coverage, incidence of vaccine-preventable diseases, and antenatal program coverage.

Responses to Complex Emergencies

Primary Prevention

Primary prevention is the basic strategy of public health, and epidemiology is one of its essential tools. In situations of armed conflict, however, epidemiology can be practiced safely and reliably in very few areas. Hence, the traditional documentation, monitoring, and evaluation elements of disease prevention may be ineffective in these settings. Because war and public health are essentially incompatible pursuits, the provision of adequate food, shelter, potable water, and sanitation, as well as vaccination and other primary healthcare services has proved problematic in countries disrupted by war. Primary prevention in such circumstances, therefore, means stopping the violence. More effective diplomatic and political mechanisms need to be developed that might resolve conflicts early in their evolution prior to the stage when food shortages occur, health services collapse, populations migrate, and significant adverse public health outcomes emerge.

Secondary Prevention

Secondary prevention involves the early detection of evolving conflict-related food scarcity and population movements, preparedness for interventions that mitigate their public health impact, and the development of appropriate public health skills to enable relief workers to work effectively in emergency settings. Preparedness planning needs to take place both at a coordinated international level and at the level of countries where CEs might occur. It might be argued that preparedness measures may be particularly important at the most peripheral levels—in communities and households. Relief agencies need resources not only to respond to emergencies when they occur, but also to implement early warning systems, maintain technical expertise, train personnel, build reserves of relief supplies, and develop their logistic capacity. At the country level, all health development programs should have an emergency preparedness component that includes the establishment of standard public health policies (e.g., rapid detection and management of epidemics, vaccination), treatment protocols, staff training, and the maintenance of reserves of essential drugs and vaccines for use in disasters.

Tertiary Prevention

Tertiary prevention involves prevention of excess mortality and morbidity once an emergency has occurred. The health problems that consistently cause the most deaths and severe morbidity as well as those demographic groups most at risk have been identified. Most deaths in refugee and displaced populations are preventable using currently available and affordable technology. Relief programs, therefore, must channel all available resources toward addressing measles and other vaccine-preventable diseases, diarrheal diseases (including cholera), malnutrition, acute respiratory infections, and, in some cases, malaria, especially among women and young children. Violence, especially sexual and gender-based violence, mental health problems and, as socioeconomic progress leads to longer life expectancy in many parts of the world, the ongoing control of noncommunicable diseases such as diabetes, cardiovascular conditions, and cancer, also need to be addressed. The challenge is to institutionalize this knowledge within the major relief organizations and to ensure that relief management and

logistical systems provide the necessary resources to implement key interventions in a timely manner.

Initially, many refugees and displaced persons may find themselves in crowded, unsanitary camps in remote regions where meeting their basic needs is extremely difficult. Recently, there has been a more noticeable settlement of displaced individuals and families in urban areas, where they are integrated with the local population. In either case, prolonged exposure to the violence of war and the deprivations of long journeys by refugees cause severe physical and psychological stress. Upon arrival at their destination, refugees—most of whom tend to be women and children—may suffer severe anxiety or depression, compounded by the loss of dignity associated with complete dependence on the generosity of others for their survival. If refugee camps are located near borders or close to areas of continuing armed conflict, the desire for security is an overriding concern. Therefore, the first priority of any relief operation is to ensure adequate protection; camps should be placed sufficiently distant from borders to reassure refugees that they are safe (Sphere Project, 2011).

To diminish their sense of helplessness and dependency, refugees should be given an active role in the planning and implementation of relief programs. Nevertheless, giving total control of the distribution of relief items to so-called refugee leaders may be dangerous. For example, leaders of the former Hutu-controlled Rwandan government took control of the distribution system in Zairian refugee camps in July 1994, resulting in relief supplies being diverted to young male members of the former Rwandan Army. Even when their intentions are good, the priorities of communities may differ from those of international relief agencies. Whereas the latter may offer preferential treatment of children because mortality rates have been shown to be highest in younger age groups, a community may feel that the elderly, for example, deserve special care because they carry with them the traditions and customs of the culture. In general, the targeting of food and other supplies to communities and to vulnerable groups within those communities should be the subject of discussion between relief authorities and the communities (Jaspars & Shoham, 1999).

In the absence of conflict resolution, those communities that are totally dependent on external aid for their survival either because they have been displaced from their homes or because they are living under a state of siege must be provided with the basic minimum resources necessary to maintain health and well-being. The provision of adequate food, clean water, shelter, sanitation, and warmth will prevent the most severe public health consequences of CEs. It would seem that the temporary location of refugees in small settlements or villages in the host country would have fewer adverse public health consequences than their placement in crowded, often unsanitary camps. Public health priorities include a rapid needs assessment, the establishment of a health information system, measles vaccination or other mass vaccination programs developed in accordance with WHO guidance (WHO, 2017e), the control of diarrheal and other communicable diseases, maternal and child health services, and nutritional rehabilitation. Critical to the success of the response is coordination of the many agencies involved in the relief effort.

Following the Goma crisis in 1994, the aid community recognized that aid agencies needed to demonstrate the effectiveness of their actions. This consensus led to the Sphere Project, which through a comprehensive, inclusive process developed the *Sphere Project Handbook: Humanitarian Charter and Minimum Standards in Disaster Response* in 2000. The *Handbook* includes minimum standards for shelter, water and sanitation, and a range of health and nutrition interventions and outcomes. A fourth edition of the *Handbook* was scheduled to appear in 2018.

Rapid Assessment

Displacement is the final, desperate act of a threatened population. Whenever possible, assessments of the public health needs of the population should be conducted prior to the act of migration or resettlement, whether it is within the country of origin or beyond its borders. Impending emergencies—even the development of CEs—can frequently be predicted. Knowing the size of the population and its age and gender distribution, having baseline data concerning its health status and the level of health services available to it, and being aware of the characteristics of the place or places to which refugees are most likely to move can be of immense help in knowing which relief supplies will be needed and which kinds of health programs should be implemented. Needless to say, such predisplacement assessments have been rare.

Early assessments can be made by a variety of means. Technology-dependent methods such as satellite surveillance can provide information regarding crop growth, population densities, and even troop movements, although such technology is often unavailable to those agencies that need it the most. Reviews of existing documents and other information provided by a variety of UN agencies, bilateral governments, NGOs, and national authorities familiar with the situation can be helpful. On-the-ground

economic evaluations, including a description of trends in market prices for food and essential commodities in food-basket analyses, can be helpful. More detailed information can frequently be obtained from visual inspection of the affected area, including mapping, key informant interviews, and observation of the affected population. Information-gathering projects and services, such as the Famine Early Warning System Network, ACAPS, and the UN-based ReliefWeb, have proved invaluable for early dissemination of information. For CEs, however, where political instability and increased violence are almost always compounding factors, more direct means of assessment prior to displacement are often impossible and the earliest assessment can be conducted only after the displaced persons have reached a relatively safe area of resettlement.

Early, rapid assessments have multiple purposes. They can provide important information about the evolution of the emergency, identify groups and areas at greatest risk, evaluate the existing local response capacity, determine the magnitude of external resources required, and indicate which health programs will be required in the short and medium terms. Every CE is characterized by a different set of causes and consequences, and each should be assessed for its impact on the health of the affected population.

For CEs, early assessment should include both a description of the conflict and its sequelae, in terms of the affected areas and populations, and a characterization of the health consequences of displacement. In some cases, affected populations may not have migrated but rather may be trapped in a siege-like setting, such as in Sarajevo (Bosnia), Aleppo (Syria), Mosul (Iraq), Juba (South Sudan), and Sanaa (Yemen). In regard to the conflict itself, variables of particular significance include the duration of the conflict, the progress of negotiations (and the likelihood of an early return for the displaced populations), the patterns of violence, the size and location of inaccessible areas and populations, and the state of remaining available health services.

The highest priorities for early assessment are the availability and adequacy of drinking water, food, and shelter. Minimum standards described later in this chapter must be met. Regarding the health status of the population, perhaps the most important and most sensitive indicator is the mortality rate. Early documentation of mortality will establish an indispensable baseline and allow for monitoring of trends that will attest to the overall effectiveness of the relief program. In an emergency, crude mortality rates are expressed as deaths per 10,000 per day. A CMR of greater than 1 has been used to define the existence of a public health emergency, and a CMR of greater than 2 indicates a critical situation. Of course, these thresholds are gross estimates. Whenever pre-crisis mortality rates are available, they should be taken into account (Guha-Sapir & van Panhuis, 2004). Age- and gender-specific mortality should be assessed to identify those population groups at the highest risk.

Rates of diseases commonly associated with high rates of preventable mortality should be assessed as early as possible. These include diarrheal illnesses, acute respiratory infections, and diseases with high epidemic potential, such as cholera, dysentery, measles, and meningitis. Where appropriate, the occurrence and risk of locally endemic diseases such as malaria and dengue should be analyzed as well. In middle-income countries, the prevalence of noncommunicable diseases should also be assessed, and attempts made to ensure ongoing care of those persons receiving long-term treatment for NCDs.

Complex emergencies are often accompanied by food shortages that can lead to malnutrition. Assessment of protein-energy malnutrition among children should be undertaken as soon as possible, using one of the variety of methods available for that purpose. Mass screening of all children is optimal, but an initial random sample of the population can establish the prevalence of malnutrition and indicate the need for targeted screening and feeding interventions. Vaccination coverage of children should also be assessed to determine the urgency of mounting vaccination campaigns. Notably, in the absence of any coverage data, measles vaccination should be considered a priority.

Rapid assessments require detailed planning and may fail because of the inadequacy of transport, maps, communications equipment, and fuel. In addition, attention needs to be given to the security situation in the affected area. An assessment is of limited value unless its results are communicated in a timely and effective manner to those who can act upon them. Presentation of the findings should be organized and clear, and the recommendations of the assessment team should indicate which actions are of highest priority, what a reasonable time frame for action would be, and which resources will be required. Without these data, essential information required for the survival of large populations of displaced individuals may not be acted upon in time to prevent high levels of excess preventable mortality.

The potential usefulness of rapid assessment should not be underestimated. Unfortunately, in the past there have often been too many assessments done by different agencies in an uncoordinated manner. It is essential that a designated lead agency coordinate rapid assessments, ensuring that sectoral assessments

(e.g., water and sanitation, medical services, and food) are integrated and that the findings are used to inform program policies and planning. For example, in the post-tsunami relief effort of January 2005, WHO organized teams from a number of UN and voluntary agencies, the Indonesian Ministry of Health, and the Indonesian military forces, and, using U.S. military assets for support, led a series of rapid assessments of stranded populations along the west coast of Aceh province on the island of Sumatra. These efforts proved quite helpful in identifying remote populations and their needs and in targeting appropriate relief. More recently, ACAPS (www.acaps.org), an NGO founded in 2009, has been providing assessment data to the emergency relief community. Of course, the collection and dissemination of useful information during the early days of an emergency, especially when security is a major concern, can be quite challenging.

In contrast to the post-tsunami assessments, the rapid assessments attempted following the Haiti earthquake of 2010 did not yield results that guided the relief effort in a timely manner. The extent of the destruction; the need to reserve means of transport, both ground and air, for the delivery of humanitarian goods; and the lack of a clear and creditable plan for exactly what to assess and where all resulted in the early days of the relief effort being a data-free period. Most important in that setting, perhaps, would have been to determine where hospital beds were available in the mostly destroyed capital city and throughout the country. Unfortunately, identification of these resources, and especially appropriate and effective sharing of the information with those who had the greatest need for it, was not rapidly forthcoming.

Health Information Systems

Epidemiologic surveillance is the ongoing and systematic collection, analysis, and interpretation of health data. This information is then used for planning, implementing, and evaluating public health interventions and programs. Surveillance data are used both to determine the need for public health action and to assess the effectiveness of programs. In CEs, after the response to an initial rapid assessment has been instituted, the development and implementation of ongoing health information systems immediately becomes a high-priority activity. Although data on many of the subjects included in the rapid assessment will continue to be collected on a regular basis, routine health information systems will allow for the monitoring of a significantly larger number of other potentially important health conditions and health programs.

Characteristics of Effective Health Information Systems

To be useful, surveillance systems must be relevant, especially in CEs, where time and resources are frequently in short supply. Data collection should be restricted to the most important of actual and potential public health problems. Equally important, data should be collected only if it will be useful in stimulating and guiding a response—if no intervention is feasible, there is little need to encumber the system with information on the problem.

The best health information systems are the simplest. In a number of CEs, difficulties have arisen in explaining the importance of the data to the local staff responsible for its collection, as well as to decision makers, who frequently do not appreciate the limitations of data of variable accuracy collected under the most difficult of circumstances. Case definitions must be clear, consistent, and suited to the local capacity to make accurate diagnoses. Where no microscopes or rapid diagnostic tests are readily available, for example, malaria may have to be represented by "fever and chills." The data generated from simple systems must be recognized for what it is worth and should not be over-interpreted. Nevertheless, good surveillance systems rely on laboratory confirmation of suspect or probable cases of diseases of public health importance. Efforts should be made to have available a basic public health laboratory, with appropriate supplies and technical expertise.

Representativeness is another essential element of health information systems that is related to the quality of the data. Careful interpretation of data collected from a surveillance system is required before extrapolations can be made to the general population. For example, in South Sudan during the 2000s, nutrition and mortality assessments were difficult to interpret because health information has been collected and reported from food distribution centers to which the displaced and most severely malnourished members of the population were drawn. Thus, the mortality and malnutrition rates derived from these sites were not necessarily representative of the entire population of the region.

The organization and implementation of health information systems should be made the responsibility of one individual or agency, which should also be responsible for ensuring widespread cooperation and coordination of data collection and use. If a host government is in place and guiding the response, it should establish appropriate policies and guidelines that can be implemented with help from appropriate technical partners. It is important for the interpretation

and response to the data that the health information system establishes standardized case definitions, data collection methods, and conditions of reporting. This sort of across-the-board consistency can help avoid the problems described earlier in which different agencies used different data collection methods.

In emergencies, both reporting and the response to reports must be timely. When the goal is to prevent excess mortality, undue delays between any two links of the surveillance chain, from the peripheral data-collection level to the more central, policy-making level and back to the periphery where action needs to be taken, can result in an unnecessary loss of life. Depending on the nature of the data, especially when an epidemic illness is deemed likely or is occurring, daily reporting of selected information by telephone, text messaging, or messenger is not necessarily excessive. Today, widely available technologies such as cell phones and digital data transfer have become the methods of choice for data transmission from more peripheral to central levels. For other conditions, data are generally reported and analyzed on a weekly basis during the emergency period and monthly during the less acute phases of the crisis. Of course, data needs may change rapidly as an emergency evolves to a more steady state. For this reason, information systems must have a high degree of flexibility, and their response to new demands should be achieved with minimal disruption.

Methods of Data Collection

As long as they possess the qualities mentioned previously, surveillance systems may combine active and passive reporting mechanisms. Active reporting can include randomized population-based surveys aimed at gathering data on one or a selected few parameters, such as vaccination coverage or nutritional status. Alternatively, it can involve the hiring of personnel for the specific purpose of monitoring important health events that might occur outside the bounds of the healthcare system itself, such as hiring gravediggers to report on burials so as to determine mortality.

Passive reporting generally refers to the routine collection and relaying of health statistics within the system itself, whether it is from community-based health posts, from primary care clinics, or from hospitals. In some cases, access to the healthcare system may be limited and utilization may be low owing to a variety of factors, such as fear or mistrust of the healthcare providers, unfamiliarity with the system, physical destruction of facilities, and other reasons that are frequently unrecognized by relief agencies. Thus, it is especially important that emergency health information systems be regularly evaluated for the characteristics described earlier.

Sometimes a combination of active and passive reporting can be used. In Haiti, for example, the system that had been established in 25 Pepfar-funded hospitals that had been treating patients for acquired immunodeficiency syndrome (AIDS) was adapted for use during the earthquake emergency, and those reporting sites began to report on a number of conditions for which information was required on a daily basis.

Content of Health Information Systems in Complex Emergencies

Trends in crude mortality remain an important feature of surveillance throughout the emergency phase and beyond. In many cultures, death is a family and religious matter, such that deaths are not normally brought to the attention of the healthcare system. In fact, severely ill patients in hospitals are frequently taken home to die. For this reason, active surveillance is best for estimating mortality. Grave watchers—often those who dig the graves—can be hired on a 24-hour basis to report new burials and, if possible, to ascertain the age and gender of the deceased. In some cultures, the free distribution of burial shrouds or other materials used for burial or funerals can provide a useful incentive for reporting. At times, mortality can be determined by means of a population-based survey, but the data derived from these surveys are subject to different sorts of bias and are frequently dated.

Sometimes large numbers of corpses are left unburied by families, as happened in Goma in 1994 because the ground was composed of volcanic rock that made it impossible to dig graves. Likewise, following extensive natural disasters such as the Indian Ocean tsunami of December 2004, bodies may be buried under rubble or swept away. In contrast, in epidemics such as the Ebola outbreak in West Africa that make safe burial procedures imperative, it may be possible to ask those responsible for body collection and mass burial to keep count of the number of dead they see. In other cases, as in Haiti in 2010, dead bodies cannot be found or are so quickly disposed of by the authorities that only gross estimates of the number of deaths can be given.

Health information systems should collect morbidity data on commonly occurring diseases and on diseases of epidemic potential. Diseases that have been prominent in all CEs include watery diarrhea, acute respiratory infections, malaria and other important endemic conditions, and malnutrition. Measles, meningitis, cholera, and shigellosis (dysentery or bloody diarrhea) have all been responsible for major

epidemics in emergency settings, and guidelines for the establishment of sensitive thresholds for the detection of each need to be followed where they already exist, or need to be developed. The detection of an epidemic should trigger an immediate and aggressive response.

In addition, at least two health programs—treatment of malnutrition and vaccination—need to be regularly monitored. Indicators of the numbers of patients in intensive or supplementary feeding programs need to be tracked. Vaccination coverage rates also need to be estimated, and, when deemed appropriate, measles vaccination should be offered to all children age 6 months to 15 years regardless of prior vaccination status as soon as resources permit. Vaccination with the antigens of the national routine immunization program should be established when feasible. Other vaccines may be offered after careful consideration by public health authorities of the epidemiology of the diseases, the characteristics of the available vaccine, and the context in which the CE is occurring.

Two areas that have been relatively neglected in CEs are reproductive health and psychosocial health (Spiegel et al., 2010). Ample evidence shows that in emergencies, women of reproductive age, and especially pregnant women, need special attention, and their health conditions, including pregnancy, should be carefully monitored by surveillance mechanisms. Gender-based violence is also a major concern, and recent evidence suggests that domestic or intimate-partner violence is an important factor in many populations that have been subjected to undue stress (Stark & Ager, 2011). Forced migration is itself a traumatic event. When its stresses are compounded by the ethnic strife and violence that frequently accompany CEs, close attention should be paid not only to individuals who might be seriously affected by post-traumatic stress disorders but also to the reestablishment of community structures. Finally, when the emergency has subsided (e.g., when crude mortality rates drop to less than 1 per 10,000 per day), increased attention can be paid to dealing with more chronic or less fatal diseases, such as tuberculosis, sexually transmitted infections (STIs, including HIV), diabetes, hypertension, and elective surgical conditions.

The establishment of a useful health information system, with all of the characteristics described here, is an essential function of the health service. Without one, programs will be developed by guesswork and the effectiveness of program implementation will remain a matter of conjecture. For a recent comprehensive review of health information needs and their use in complex emergencies, see Checchi et al. (2017). Information and surveillance systems are also covered in the *Public Health Infrastructure* chapter of this text.

Shelter and Environment

As mentioned earlier, the placement of refugees and IDPs in small settlements or integration into local villages is preferable to the establishment of large camps. Health outcomes are probably better in these small settlements because environmental conditions are more favorable and there is less crowding. However, the provision of relief assistance to a large number of scattered settlements may pose difficult management and logistical challenges and may provoke resentment in the surrounding communities. In Guinea in the early 1990s, food aid and other relief items were provided to communities for distribution to both Liberian refugees and local inhabitants. This system may have worked well because the refugees and the local population were of the same ethnic origin and many were related. In contrast, measles vaccination coverage of Mozambican refugees in Malawi who were absorbed into communities and were dependent on the national immunization program was considerably lower than that in refugee camps, where services were provided by the United Nations and international voluntary agencies.

An interesting case in point involves the crisis in Syria, which began in 2011. In addition to a large number of internally displaced persons, an estimated 5 million refugees have fled, mostly to neighboring countries. In Jordan, many of the refugees have been settled in camps. In contrast, in Lebanon, no camps have been built and refugees must settle in small clusters or try to integrate with the indigenous population. This leads to competition for jobs, housing, and other necessities and can result in an increase in xenophobic tensions. In Turkey, the situation is mixed: Approximately 3 million Syrian refugees have moved to that country, but only 260,000 are housed in camps, most camps ranging in size between 5,000 and 10,000 people.

When camps are unavoidable, three key determinants of location should be safety and protection of residents, access to essential supports such as adequate quantities of clean water for drinking and fuel for cooking, and all-weather access. In many instances, local politics determines the site, and the location is often less than desirable. For example, in 1988, the Ethiopian government placed Somali refugees in a large camp of 180,000 persons on a site that had no local supply of drinking water. The nearest source with an adequate quantity of water was situated in a

town 100 km from the camp. For many years, water was delivered daily in convoys of trucks at enormous cost to the relief agencies. The Ethiopian government refused to move the refugees closer to town for fear of exacerbating political problems that it was experiencing with local ethnic Somalis.

Sites should be chosen with ease of water drainage in mind, though sometimes drainage systems have to be created at the time of camp construction. This is critical to ensure access by vehicles, limit disease vector breeding, and facilitate refugees' access to services, such as health clinics and food distribution. When hundreds of thousands of refugees fled ethnic cleansing in Kosovo in 1999, spontaneous camps were established near the border of Macedonia, where as many as 45,000 refugees camped on a muddy, snowy field sheltered only by their vehicles and makeshift tents. An added hazard was a large quantity of land mines laid by Serb forces along the border. Other inappropriate locations have included the swampy, malarious areas on the Thai border where Cambodian refugees were housed, the inhospitable mountainous area on the border between Iraq and Turkey to which the Kurdish population of Iraq fled in the wake of the 1991 Gulf War, and the volcanic, rocky ground in eastern Zaire where Rwandan refugee camps were settled, precluding both latrine construction and burial of the dead.

The issue of shelter for refugees is a thorny one for Western European authorities as well. A case in point is the "Calais jungle" that was established in 2015–2016 during the ongoing European refugee crisis. Approximately 6,000 refugees settled around Calais, France, hoping to reach the United Kingdom. French authorities were reluctant to provide decent living conditions for fear of attracting additional refugees. Healthcare and other essential services were provided by NGOs that are usually involved in refugee crises in LMICs.

Ideally, the size of camps should be limited to 20,000 residents for reasons of security and ease of administration. Such camps should be further divided into sections of 5,000 persons for the purpose of services delivery. Shelter is an urgent priority. On average, the covered area provided per person should be 3.5 to 4.5 m^2 (Sphere Project, 2011). In warm, humid climates, shelters should have optimal ventilation and protection from direct sunlight. In cold climates, shelter material should provide adequate insulation combined with sufficient clothing, blankets, bedding, space heating, and caloric intake. Ideally, houses should be built using a traditional design and local materials. This may pose local environmental problems such as the destruction of trees, so building materials should be trucked in from areas remote from the camp. Waterproofing is essential and may be achieved with plastic sheeting or tarpaulins. Tents may provide temporary shelter; however, they deteriorate in rain and wind and should be replaced with local materials as soon as possible.

To limit further environmental damage through deforestation, cooking fuel, such as charcoal, wood, oil, or kerosene, should be brought to the site from remote areas and fuel-efficient stoves provided or constructed. Camps may easily become fire hazards, and fire prevention should be an objective of proper camp design.

While many proposals for shelter continue to be made, many focus solely on the design and materials used for housing structures themselves. Such a narrow focus does not take into account that the issue of providing adequate shelter to refugees and internally displaced persons is a process, much more than it is a product.

Water and Sanitation

When refugee camps are unavoidable, proximity to safe water sources needs to be recognized as the most important criterion for site selection. The Sphere Project's key indicator for minimum water supply is 15 liters of clean water per person per day for all domestic needs—cooking, drinking, and bathing. Other indicators include at least one water collection point per 250 people and a maximum distance from shelter to the nearest water point of 500 meters. Ideally, both the quantity and the quality of water provided to refugees and displaced persons should meet international standards; however, in many cases this is not possible. In general, ensuring access to adequate quantities of relatively clean water is probably more effective in preventing diarrheal disease, especially bacterial dysentery, than providing smaller quantities of microbe-free water.

The usual options for supplying water to refugees and IDPs include surface water (e.g., lakes, rivers, streams), shallow wells, springs, bore wells, and trucked water from remote sources. Although surface water is often abundant, it needs to be treated, usually through a system of sedimentation and chlorination, and sometimes with filtration. A system of piped distribution to bladders or other means of storage, as well as outlet taps needs to be developed to avoid crowding and drainage problems. Shallow wells and springs need to be protected and, unless gravity-fed, provided with a mechanism for drawing water, such as pumps. Deep-bore wells provide clean water at the source, but it may take some time to bring in the necessary equipment to establish them.

In addition, measures to prevent post-source contamination need to be implemented, including treatment at the source (e.g., "bucket chlorination"). Sufficient collection and storage containers (at least three 20-liter containers per family), especially containers with narrow openings that prevent post-collection contamination, should be made available. A study in a Malawian refugee camp in 1993 demonstrated a significant reduction in fecal contamination of water stored in such buckets compared with standard buckets. In addition, the incidence of diarrhea among children younger than 5 years of age in the households with the improved buckets was considerably lower than in control households (Roberts et al., 2001). Recently, newer methods of disinfecting water in the home, including simple chlorination and/or flocculation systems, such as the Safe Water System (CDC, 2008), have been moderately successful in CEs.

Adequate sanitation is an essential element of diarrheal disease prevention efforts and a critical component of any relief program. While the eventual goal of sanitation programs should be the construction of one latrine per family, interim measures may include the designation of separate defecation areas and the temporary provision of neighborhood latrines. The Sphere Project minimum indicator of acceptable household-level sanitation is a maximum of 20 persons per latrine. Toilets should be segregated by sex and be no more than 50 meters from dwellings for security reasons. To achieve maximal impact, these measures should be complemented by community hygiene education and regular distribution of at least 250 g of soap per person per month. Hygiene education has been shown to significantly increase the impact of sanitation programs. An analysis of data gathered in the Malawi study cited earlier demonstrated that the presence of soap in households significantly reduced the risk of diarrhea (Peterson, Roberts, Toole, & Peterson, 1998).

Achieving the goals associated with these indicators can be difficult. In Goma, where the ground was volcanic rock and digging was difficult, sanitation remained inadequate and undoubtedly contributed to the virulence of the cholera outbreak. In Haiti after the 2010 earthquake, in an almost totally destroyed urban area, reconstituting even rudimentary plumbing was out of the question during the emergency period. Moreover, the construction of latrines was painfully slow and greatly hampered by the prevailing conditions, even when the importance of improved sanitation was clearly recognized. The objective of post-emergency sanitation measures should be to restore the pre-disaster levels of environmental services rather than attempting to improve on the original levels, although longer-term reconstruction efforts can certainly aim to "build it back better"—that is, to improve upon the status quo ante.

The provision of water and sanitation in emergency-affected populations in urban and rural areas of Eastern Europe has posed different challenges. In general, the goal is to repair existing systems. Even so, interim measures may be required, such as rehabilitating old wells, providing generators to pump water from distant sources such as rivers, and providing containers and security for residents to collect water at available sources. In the "Calais jungle" mentioned earlier, sanitation was notoriously poor and the difference between the living conditions of the refugees and the surrounding population was stark.

Food Rations and Distribution

The quantity and quality of food rations is one of the most critical determinants of health outcomes in emergency-affected populations. During the early evolution of CEs, measures should be taken to increase access to food without forcing people to leave their homes. The establishment of "feeding camps" may act as a magnet for hungry families and lead to the spontaneous establishment of settlements and camps, with the subsequent health problems described earlier in this chapter. In addition, in conflict zones, warring parties may target areas where populations regularly congregate. Subsidized food shops; food-for-work or cash-for-work programs; emergency support for home food production, such as the distribution of fast-growing seed varieties and agricultural tools; and other measures may be effective in ensuring food provision prior to the onset of armed conflict.

Once armed conflict has commenced, either forcing people to flee their homes or be trapped in siege-like conditions, it is usually necessary to distribute food. General food rations should contain at least 2,100 kilocalories (kcal) of energy per person per day as well as the other nutrients listed in **TABLE 12-5**. Rations should take into consideration the demographic composition of the population, the specific needs of vulnerable groups, cultural food preferences, and access by the population to alternative sources of food and income. In cold climates, the minimum energy value of the ration should be adjusted upward. Pregnant women require, on average, an extra 285 kcal per day, and lactating women an extra 500 kcal. These extra requirements should be provided through distribution of rations within the household; however, this may need to be monitored, perhaps indirectly via the prevalence of low-birth-weight babies.

TABLE 12-5 Minimum Nutritional Requirements for Emergency-Affected Populations, for Planning Purposes During the Initial Stages of an Emergency

Nutrient	Minimum population requirements
Energy	2,100 kcals
Protein	53 g (10% of total energy)
Fat	40 g (17% of total energy)
Vitamin A	550 µg RAE
Vitamin D	6.1 µg
Vitamin E	8.0 mg alpha-TE
Vitamin K	48.2 µg
Vitamin B1 (Thiamin)	1.1 mg
Vitamin B2 (Riboflavin)	1.1 mg
Vitamin B3 (Niacin)	13.8 mg
Vitamin B6 (Pyridoxine)	1.2 mg
Vitamin B12 (Cobalamin)	2.2 µg
Folate	363 µg DFE
Pantothenate	4.6 mg
Vitamin C	41.6 mg
Iron	32 mg
Iodine	138 µg
Zinc	12.4 mg
Copper	1.1 mg
Selenium	27.6 µg
Calcium	989 mg
Magnesium	201 mg

Reference: RNI from FAO/WHO (2004). Vitamin and Mineral Requirements in Human Nutrition. Second edition, were used for all vitamin and mineral requirement calculations except copper, as requirements for this mineral were not included in FAOA/VHO (2004). Requirements for copper are taken from WHO (1996), Trace Elements in Human Nutntion and Health.
Abbreviations: Alpha-TE = alpha-tocopherol equivalents; DFE = dietary folate equivalents; NE = niacin equivalents; RAE = retinol activity equivalents.

Food should be distributed regularly as dry items to family units, taking care that socially vulnerable groups, such as female-headed households, unaccompanied minors, and the elderly, receive their fair share. Achieving this aim requires an accurate registration system listing all residents in family groups. Because of the indignity of being counted and the fact that many refugees are afraid to have their identities known and precise locations reported, attempts at registration for the distribution of humanitarian aid should be carried out with the security and cultural concerns of the population taken into careful consideration (Harrell-Bond, Voutira, & Leopold, 1992). If a refugee or displaced population is organized in well-defined communities, food may be distributed to community leaders, who then divide it further to the heads of households in quantities based on the number of family members. Sometimes, as was the case during the Kurdish exodus to camps on the Iraq–Turkey border in 1991, this arrangement has been established in voluntary exchange for demographic information that allows health authorities to enumerate the births and deaths occurring. In other situations, food is distributed directly to heads of households, based on a ration card system.

In LMICs, food rations usually comprise a staple cereal, such as rice, wheat, or maize, or a mix such as corn–soy blend; a source of dense fat, such as vegetable oil; a source of protein, such as beans, lentils, groundnuts, or dried fish; and extra items such as salt, tea, and spices. Experience has shown that women are fairer than men in distributing each food item in the correct quantity. Standard serving containers based on a known weight or volume of food are essential for distribution centers.

There is a widespread belief that currently available refugee rations, especially those distributed in non-European populations, are inadequate to meet nutrient requirements, which may be higher than the traditional recommended daily allowances. The best way to combat depletion of essential nutrients is to take active steps to increase the variety of the diet. This is not possible with relief foods that have to be capable of bulk storage and shipment—cereal, pulses, sugar, oil, and salt. For this reason, recipients of international aid should be encouraged to barter food items for local produce. Market facilities should be established, supported, and controlled by camp leaders. Seeds for leafy vegetables should be distributed as part of all relief activities, even in the acute phase of the CE. Every effort should be made to provide spices and herbs, used traditionally by the population, with all food baskets. There is no need for a special food basket for children or pregnant and lactating women—only a sufficiently varied diet for the whole family (Golden, 1999).

Increasingly, especially in urban areas, cash vouchers or ATM cards are provided to refugees to purchase food. At the first World Humanitarian Summit, held in Istanbul in 2015, one of the major elements of the "Grand Bargain" reached by attendees was to vastly increase the use of cash programming. That is, it was agreed that populations in need would be provided cash and allowed to budget for themselves and their households in accordance with their own priorities. Buying food in local markets based on individual preferences represents an important use of this new way of providing humanitarian assistance, albeit one that has not yet been adequately implemented or evaluated.

In addition to food, adequate cooking fuel, utensils, and facilities to grind whole-grain cereals need to be made available to all families. Fuel-efficient stoves, often made of mud, may lead to more efficient use of scarce fuel. In children younger than 2 years, breastfeeding will provide considerable protection against communicable diseases, including diarrhea. Thus, attempts to introduce or distribute breast milk substitutes and infant feeding bottles should be strongly opposed in an emergency situation. The evidence that vitamin A deficiency is associated with increased childhood mortality and disabling blindness is now so convincing that supplements of vitamin A should be given routinely to all refugee children younger than 5 years of age at first contact and every 3 to 6 months thereafter (Nieburg et al., 1988).

In Eastern Europe, the same principles have been followed in providing food rations, though the types of food have varied and have included cheese, meat, powdered orange juice, and fruit. In some industrialized countries experiencing economic and political instability, ration vouchers have been distributed to vulnerable persons, such as elderly pensioners, which can be redeemed for food at designated stores. More recently, cash transfers to cell phones or to debit cards have been made.

One of the main problems in Africa, and in some parts of Asia, has been providing refugees and IDPs with foods containing adequate quantities of micronutrients, especially niacin, riboflavin, thiamin, iron, and vitamin C. Epidemics of pellagra, scurvy, and beriberi have been common in African refugee camps. For many years, this was a blind spot in emergency food and nutrition planning; more recently, the problem has been acknowledged, although solutions are still inadequate.

In southern Africa, niacin has been added to maize flour during the milling process. However, vitamin C is water-soluble and sensitive to heat, light, and bruising; thus, the transport, storage, and distribution of large quantities of foods such as citrus fruit have been problematic. One solution to this problem is to allow and encourage people to swap some of their ration items in local markets for vitamin C–containing foods, such as tomatoes, onions, potatoes, green chili peppers, and other fruits and vegetables. In addition, the provision of seeds to enable refugees to grow small amounts of vegetables in kitchen gardens is an effective measure. The provision of fortified blended cereals has also been proposed as a vehicle for ensuring adequate micronutrient intake. Studies in Ethiopia, Nepal, and Tanzania have shown that these cereals are generally acceptable, although overcooking and consequent depletion of the vitamin C content may be a problem in some communities (Mears & Young, 1998). WHO has published a series of guidelines on the prevention of scurvy (WHO, 1999a), thiamine deficiency (WHO, 1999b), and pellagra (WHO, 2000a).

A food ration monitoring system is important to ensure that families are receiving fair and adequate quantities of food. On a food ration distribution day, monitoring teams can take up positions at several points not far from the distribution center. They should be equipped with weighing scales and food composition tables. Families should be stopped randomly as they return to their homes and asked to participate. Each of the items in their ration should be weighed and converted to calories and other nutrients using the tables. The total weight of each food item (in grams) and the total nutrients provided should be divided by the number of family members and the number of days until the next distribution day. This calculation will provide the average quantity per person per day and may be compared with the official ration and with standard tables of recommended daily allowances of nutrients.

Nutritional Rehabilitation

In general, the goal of an emergency feeding program is to provide adequate quantities of nutrients through the general household distribution of food rations. However, due to the many factors described earlier, some population subgroups may be either already acutely malnourished or at high risk of becoming malnourished. In general, children are defined as acutely malnourished according to their weight-for-height index, although mid-upper arm circumference is also frequently used as a screening tool. Adolescents and adults may be defined as acutely malnourished according to their MUAC, body mass index, or clinical signs.

Currently malnourished and at-risk groups may require targeted feeding, or what is termed *selective feeding*. **FIGURE 12-5** demonstrates the various kinds of selective feeding. Notably, these programs should be seen as additions to, rather than substitutes for, the general feeding program. They need clear objectives and criteria for opening, admission, discharge, and closure that should be based on population-based anthropometry surveys and agreed-upon nutritional indices. Selective feeding programs should be complemented by measures to improve the food ration distribution system, provide adequate clean water and sanitation, and control measles and other communicable diseases. They should be integrated into community health programs that offer other prevention and care services.

During the first decade of the 21st century, a major change evolved in the way in which selective feeding programs are delivered. There was a trend away from reliance on fixed sites to prepare and administer supplementary and therapeutic food. Community management of acute malnutrition (CMAM) programs are now the norm *except* for the management of severely malnourished children with complications such as dehydration, acidosis, sepsis, and hypothermia. Instead of labor-intensive on-site preparation of food, ready-to-use therapeutic foods (RUTF) manufactured by commercial companies have become the most preferred product. The most common is Plumpy'nut, which is produced by the French company Nutriset. Plumpy'nut requires no water preparation or

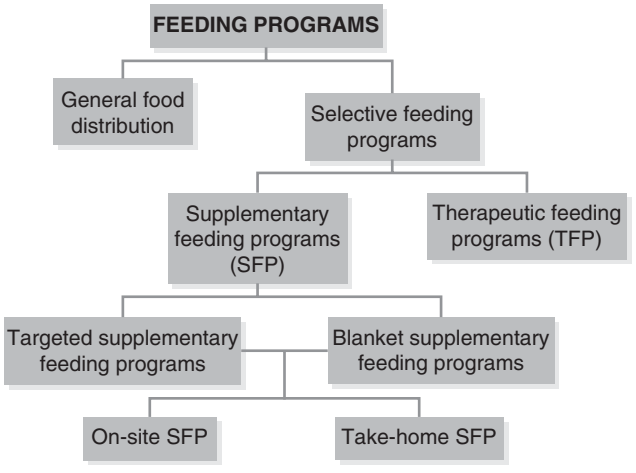


FIGURE 12-5 Overall feeding program.

Reprinted with permission from United Nations High Commissioner for Refugees (UNHRC), & World Food Programme. (1999). *Guidelines for selective feeding programmes in emergency situations.* Geneva, Switzerland: UNHCR.

refrigeration and has a 2-year shelf life, making it easy to deploy in difficult conditions to treat severe acute malnutrition. The ingredients of this product are peanut paste, vegetable oil, powdered milk, sugar, vitamins, and minerals, combined in a foil pouch. Each 92-g pack provides 500 kcal or 2.1 MJ.

RUTF has proved effective in the rehabilitation of malnourished children. In a study in Malawi, 219 children age 1 to 5 years with severe malnutrition and edema (but without other complications) were provided with RUTF for feeding at home. More than 83% of the children recovered, while 5% died, meaning that this intervention exceeded international minimum standards (Ciliberto, Manary, Ndekha, Briend, & Ashorn, 2006). Average weight gain was roughly 2.8 g/kg/day, which is less than international standards, but nonetheless resulted in good overall recovery. Research has further shown that effective RUTF can be manufactured in most settings worldwide at low cost using local foods and fairly common equipment.

Supplementary feeding programs (SFPs) provide nutritious foods in addition to the general ration, with the aim of rehabilitating moderately malnourished individuals and decreasing the population prevalence of acute malnutrition. There are two kinds of SFPs: targeted and blanket. Each type may provide food supplements as on-site, precooked ("wet"), or take-home ("dry") rations. Take-home rations are usually preferred because they require fewer resources, carry lower risks of cross-infection among recipients, and retain the family's primary responsibility for feeding. On-site feeding may be justified when the security situation is poor, when the general ration is inadequate such that food is likely to be shared within the household, or when cooking fuel is scarce.

A targeted SFP aims to prevent moderately malnourished individuals from becoming severely malnourished and to rehabilitate them. Such programs generally focus on children younger than 5 years and on pregnant and lactating women. Care should be taken not to apply strict guidelines for the opening of selective feeding programs. Like all decisions taken in an emergency, it should be based on a review of both current public health priorities and available resources.

Current guidelines suggest that a targeted SFP should be implemented when one or more of the following situations occur:

- The prevalence of acute malnutrition among children younger than 5 years is 10% to 14%.
- The prevalence of acute malnutrition among children is 5% to 9% in the presence of aggravating factors (e.g., inadequate general food ration, CMR greater than 1 per 10,000 per day, epidemics of measles or pertussis, and high incidence of acute respiratory infections and diarrheal diseases).

Criteria for admission should include the following:

- Children younger than 5 years with a weight-for-height *Z*-score between –3 and –2
- Older children, adolescents, and adults who are moderately malnourished according to their MUAC, BMI, or clinical condition
- Referrals of children graduating from therapeutic feeding programs
- Selected pregnant women and nursing mothers

Blanket SFPs are meant to provide extra nutritious food to all members of groups at risk of malnutrition. They are started under the following conditions:

- General food distribution systems not yet in place
- Problems in delivering and distributing the general ration
- Prevalence of acute malnutrition equal to or greater than 15%
- Prevalence of acute malnutrition of 10% to 14% in the presence of aggravating factors
- Anticipated increase in acute malnutrition rates due to seasonally induced epidemics
- In the case of micronutrient-deficiency disease outbreaks, to provide micronutrient-rich supplements to a target population

The primary recipients of blanket SFPs are all children younger than 5 years, pregnant women and nursing mothers (up to 6 months after delivery), and other groups such as the elderly or medically ill. As with RUTF, ready-to-use supplementary foods (RUSF) have been developed and are gaining popularity. Their judicious use has the potential to replace on-site supplementary feeding programs to a large extent.

Therapeutic feeding programs (TFPs) aim to rehabilitate severely malnourished persons, the majority of whom are children with severe wasting or nutritional edema—conditions commonly known as marasmus and kwashiorkor, respectively. A TFP should be initiated when the number of severely malnourished individuals cannot be treated adequately in other facilities. Although current guidelines for opening a TFP are based on population malnutrition prevalence, the decision should be based on an assessment of the public health needs and the resources available. In a large camp, the percentage of malnourished children may not be an appropriate basis for this decision. Perhaps a more practical guide would be the presence of, say, 50 potential beneficiaries in a settlement.

The following criteria are used for admission (UNHCR & World Food Programme, 2011):

- Children younger than 5 years with weight-for-height *Z*-scores less than –3 or with edema, with complications such as dehydration, sepsis, hypothermia, inability to take foods orally, or poor conscious state
- Severely malnourished children older than 5 years, adolescents, or adults assessed by other means and with complications
- Low-birth-weight babies (less than 2,500 g)
- Orphans younger than 1 year of age when traditional caring practices are inadequate
- Mothers of children younger than 1 year with breastfeeding failure where relactation counseling has failed

Rehabilitation of severely malnourished individuals is a medical procedure demanding careful attention to detail; therefore, well-trained staff are essential. TFPs usually have two phases:

1. The intensive care phase includes 24-hour care with medical treatment to control infection and dehydration and correct electrolyte imbalance, and nutritional treatment. Frequent feeds with therapeutic milk (F-75) are essential to prevent death from hypoglycemia and hypothermia. This phase usually lasts 1 week.
2. The rehabilitation phase provides at least 6 meals per day along with medical and psychological support, including for mothers of malnourished children. This phase usually lasts 5 weeks. This phase can take place at home using RUTF.

Details of the quantity and types of food recommended, discharge criteria, and reasons for stopping SFPs and TFPs are provided by UNHCR and the World Food Programme (2011). In addition, detailed information on the management of severe malnutrition is provided in WHO manuals (WHO, 1999c, 2000b). It should be noted that the increasing use of RUTF like Plumpy'nut has revolutionized the care of severe undernutrition; hospitalization with 24/7 supervision of the administration of detailed parenteral nutrition protocols is no longer regularly required and has, to a significant degree, been replaced by community therapeutic care.

Community Management of Acute Malnutrition

Collins et al. (2006) outline three levels of CMAM: (1) inpatient stabilization care for people with acute malnutrition and serious medical complications; (2) outpatient therapeutic programs for people with severe acute malnutrition but no complications; and (3) supplementary feeding for moderate acute malnutrition and no complications. Under this definition, inpatient care is provided according to WHO guidelines. It is worth noting that other studies have categorized CMAM programs differently. For example, a recent review of the effectiveness of CMAM separated programs into four categories: (1) residential nutrition centers; (2) day-care nutrition centers; (3) clinic-based care; and (4) domiciliary (home) care (Ashworth, 2006). Although the term CMAM is frequently used in reference to slightly different models, the three-level program structure outlined previously—inpatient care, outpatient care, and supplementary feeding program—is a useful one.

CMAM approaches are also becoming more effective as techniques and protocols develop. In an extensive literature review, comprising 33 community-based nutrition rehabilitation programs implemented between 1980 and 2005, only one-third (33%) were considered effective, using low mortality and acceptable weight gain as indicators of success (Ashworth, 2006). In contrast, among those programs implemented between 1995 and 2005, a much greater proportion (62%) was considered effective.

Health Services: Response to Complex Emergencies

Complex emergencies severely disrupt normal healthcare service activity (see Figure 12-3). In developing appropriate responses to this disruption, it is worth considering the impact on and response by at least three different sets of services:

- Service provision in the country affected by the CE
- Service provision in countries to which refugees have fled
- Service provision by multilateral agencies and NGOs

Service Provision in Countries Directly Affected by Complex Emergencies

The main challenge in these settings is to seek to minimize the direct and indirect adverse impacts of the conflict on the health services personnel and resources available. In many conflicts, as highlighted earlier, health services may be specifically targeted, as may health services providers. In such situations it may be extremely difficult to maintain services due to absent

or fleeing health personnel, lack of drugs and equipment, disruption of referral systems, and destruction of physical infrastructure, including hospitals and clinics. In recent years, the purposeful destruction of health facilities in Syria, Yemen, Afghanistan, and Iraq has represented clear violations of international law. In 2016, the United Nations Security Council adopted Resolution 2286, deploring the short- and long-term consequences of such attacks for the civilian populations and healthcare systems of the countries concerned. To date, no action has been taken on the basis of this resolution.

Even if the official health system is destroyed, however, the health workers who previously worked within it may still be present within the community (unless they have themselves fled to safer areas) and may still be able to offer services and advice. The extent to which services can be maintained is in part dependent on earlier disaster preparedness activities, in terms of both training and of prepositioning of drugs and other supplies in areas where logistic support was predicted to be vulnerable to disruption.

In some situations, it has been possible to ensure that areas where services are likely to be disrupted have received prior stocks of drugs and equipment so that they can maintain services despite disruption of the normal supply chains. Although this might be possible for certain forms of equipment and drugs, others, such as vaccination facilities (which are dependent upon maintaining an intact cold chain), and support activities, such as training and supervision, are invariably disrupted during CEs. Nevertheless, local-level responses are possible. In Afghanistan, despite the conflict in Kabul, the medical school was able to relocate to another city to continue its training activities. In relation to the availability of drugs for treatment, prior distribution of necessary drugs to responsible community members—such as community and other health workers, teachers, and in some cases patients with chronic conditions—may ensure their availability despite the population having to move suddenly. Patients with chronic conditions such as leprosy or tuberculosis could be issued with drug supplies for extended periods, assuming that the drugs can be procured and distributed and that patients with these conditions will be able to adhere to the treatment regimens prescribed for them. The potential to distribute insecticide-treated mosquito nets to populations on the move might be an option in areas of high malaria risk; some experience in this approach has been gained among the highly mobile Burmese refugees on the Thai border.

Service Provision in Countries to Which Refugees Have Fled

When refugees flee from their country and cross borders, their normal sources of health care are no longer available to them. Therefore, they may be dependent on what they can provide for themselves, which is often desperately little if they have been forced to flee suddenly, as was the case with the Kosovar Albanians in 1999 and with Syrians in 2011 and the ensuing years. Alternatively, they may depend on what can be provided in the host country to which they have fled, either through existing host country services or through additional services mobilized by NGOs and other organizations. Host government services may rapidly become overwhelmed if large numbers of refugees suddenly move into an area and seek to use the local health system. In addition to their absolute numbers, the health condition of refugees may be poor, especially if their journey has been traumatic and unplanned or if their prior state of health was poor.

In most cases, at least in the short term, these local host-country services will not receive additional resources and, therefore, will have to cope as best they can with the additional service load. In some instances, such as with significant flight from middle-income countries like Syria, services that require a moderate level of technical expertise and that may be expensive, such as renal dialysis or oncologic services, may no longer be able to cope with demand and need. This trend may disadvantage local community members who ordinarily utilize such services, because the supply of drugs and usual access to health workers may be compromised. Communities will typically be willing to accept these hardships for a time, but if the situation becomes protracted and leads to a lasting decline in services, tension may develop between the host and refugee communities. Another point of friction occurs when refugees are offered access to host services at no cost, while local community members may be required to pay user charges, both informal and formal, to obtain care.

Finally, different policies may be applied to refugee and local populations. For malaria treatment, for example, international organizations providing care to the emergency-affected population might opt for more effective drugs than those designated by existing national policies. Such was the case in Burundi in 2000–2001, where disagreements between MSF and the government of Burundi over malaria treatment policies took several months to resolve (MSF, 2001). Devoting attention to issues of equity and to ensuring that host and refugee communities gain similar access to services will be in the longer-term best interests of both groups.

Service Provision by Multilateral Organizations and Nongovernmental Organizations

The key challenges for these organizations are, in the short term, to reduce excess preventable loss of life and to reestablish an environment in which maintaining and promoting health is possible. Many of the subsequent sections deal with the nature of the health-related interventions that are necessary in these settings.

In camp settings, health services should be organized to ensure that the major causes of morbidity and mortality are addressed through fixed facilities and outreach programs. An essential drug list and standardized treatment protocols are necessary elements of a curative program. It is not necessary to develop totally new guidelines in each refugee situation; several excellent manuals already exist, from which guidelines can be adapted to suit the local conditions (MSF, 2017; WHO, 2003). Curative services should be decentralized by establishing a camp system consisting of community health workers, health posts, first-line outpatient clinics, and a small inpatient facility to treat severe emergency cases. Patients requiring surgery or prolonged hospitalization should be referred to a local district or provincial hospital; this facility will require provision of extra drugs and other medical supplies to cope with the greater patient load.

Organizations that typically espouse a development approach rather than relief-oriented approach have sought to place the issue of developmental relief onto the CE management agenda, arguing that early attention to the difficult issues of efficiency, effectiveness, sustainability, equity, and local ownership will be beneficial in the longer term. If the latter approach is adopted, greater effort will be devoted to activities such as training, building local capacity, and keeping costs down, rather than seeing these as desirable, but not practical, goals given the acute needs faced in relation to saving lives.

An additional concern facing the range of organizations offering services in response to humanitarian crises is the importance of coordination. Organizations need to work together if they are to reinforce each other's actions, maximize the use of available resources, minimize duplication and overlap, and enhance effectiveness, equity, and efficiency. The Code of Conduct for Humanitarian Organizations highlights the need for effective coordination, usually under the aegis of a lead organization from the United Nations system, such as OCHA, UNHCR, or, in some cases (especially for the health sector), WHO or UNICEF. Optimal levels of coordination have, unfortunately, been an elusive goal in many emergency settings.

Lastly, considerable concern has arisen in recent years regarding the quality of much of the service provision by NGOs operating in CEs. A key problem is the lack of transparency and accountability of such services, and the fact that despite meaning well, some organizations may do more harm than good. Recent initiatives, such as the Sphere Project (see Exhibit 12-4 later in this chapter), seek to establish standards and minimum indicators of acceptable performance for organizations operating in these environments. Although no enforcement mechanisms, or even effective incentives, have been developed to ensure compliance with these suggested standards, many voluntary organizations and even institutional and governmental donors have agreed to do what they can to comply. The increasing professionalization of the field of humanitarian assistance, briefly discussed toward the end of this chapter, similarly reflects these trends toward improved practice and accountability.

Refugee Health Workers

Community health workers (CHWs) are seen in many countries as the mainstay of primary healthcare promotion, although some critics have questioned their value to national primary healthcare programs (Chou et al., 2017; Walt, 1988). In refugee and displaced person settings, the selection and training of refugee health workers has been considered as one key mechanism by which health programs can work more closely with affected communities.

The principal advantages of refugee health workers may be related to their role as intermediaries between the affected community and the services provided by the humanitarian agencies, which are often led by expatriate staff. Refugee workers are more likely to understand the cultural, behavioral, and environmental influences on health status, to contribute to realizing the potential for self-care and refugee-provided services within the community, to share the health service provision workload, to build capacity and skills that will potentially be available after repatriation, and to enhance the dignity of both the community and the healthcare providers themselves. CHWs who are relatively unskilled and trained within the community may be the mainstay of service provision. However, it is important to recognize that the presence of trained health workers within the affected community—whether these be midwives, nurses, doctors, or others—represents an extremely valuable resource whose role should be promoted in whatever services are developed with expatriate agency support.

The role of CHWs in refugee settings will depend on the public health needs, the availability of host country– and NGO-provided services, the prior level of skill of the workers, and the extent and quality of training received. In many cases, refugee health workers have worked outside of health center settings and have performed a range of tasks, including the following:

- Identifying sick and malnourished community members and assisting them in obtaining assistance
- Collecting and reporting demographic data such as births and deaths
- Providing first aid and basic primary care such as oral rehydration for children with diarrhea
- Assisting in mass vaccination campaigns
- Encouraging participation in health campaigns and disease control programs
- Ensuring that the needs and perspectives of refugees are taken into account in the development of health programs

On a different level, in many conflict-affected countries, professional health workers, including, among others, radiologists, anesthesiologists, and surgeons, remain in their communities and, under extremely trying circumstances, attempt to provide at least essential services. The conflicts in Syria, Afghanistan, Yemen, and Iraq have all demonstrated the remarkable dedication and professionalism of health workers. As discussed earlier, these professionals have often been targeted by armed forces in an attempt to destabilize their communities.

One of the first refugee programs that systematically trained large numbers of CHWs as the basis of health services delivery was established in Somalia in the early 1980s. At that time, almost 1 million refugees from Ethiopia were housed in 35 camps scattered throughout three regions of the country. The Ministry of Health's Refugee Health Unit coordinated the training of approximately 2,000 CHWs and 1,000 traditional birth attendants with the help of NGOs working in the country. The training curriculum was standardized, as were treatment protocols, essential drugs, and salaries and conditions. As NGOs gradually withdrew from the country, health services were largely provided by CHWs and traditional birth attendants, supervised by Somali Ministry of Health trainers, nurses, and doctors.

Despite the many potential advantages of developing and working with a cadre of refugee health workers, this approach also faces numerous problems. It may be difficult to recruit health workers, especially in unstable settings or if potential workers have other pressing priorities, and they will require training and ongoing supervision. Their selection may be highly political, and identifying the affiliations of potential health workers may be difficult for agencies with limited knowledge of the area in which they are working. There is potential to stimulate rivalry and exacerbate perceptions of inequity if an inappropriate mix of workers from different areas and different backgrounds is selected. In the presence of adequately resourced host-country services, especially those that have sufficient numbers of trained staff, establishing a cadre of CHWs with more limited skills may not be seen as a priority.

Among the prerequisites for effective CHW training programs are a clear description of the tasks to be performed, an adequate level of logistic and supervisory support, a transparent system of selection and remuneration, and the consent and involvement of the affected and, where relevant, host communities.

Women's and Children's Healthcare Services

Children and women, especially pregnant women, have been repeatedly shown to be the most vulnerable members of refugee and displaced populations, especially during the emergency period. Among Rohingya refugees in Bangladesh (1992–1993), women and girls were seen less frequently at health facilities than men and boys, but had much higher mortality (**FIGURE 12-6**).

In Goma in 1994, households headed by women were found to have substantially less food and nonfood commodities that had been issued at general distribution points by international relief officials than those in which an adult male was present. For these reasons, health services oriented to the specific needs of children and women are essential in reducing morbidity and mortality within a population to a minimum level. Women's and children's health (WCH) care should begin within the community, at the household level, and not depend entirely on established health facilities. Often, as discussed earlier, community members, such as traditional birth attendants or others previously trained as CHWs, can be recruited from within the affected population itself to provide basic services.

For children, routine screening and preventive services are important. Growth monitoring, and referral of children whose growth is faltering to

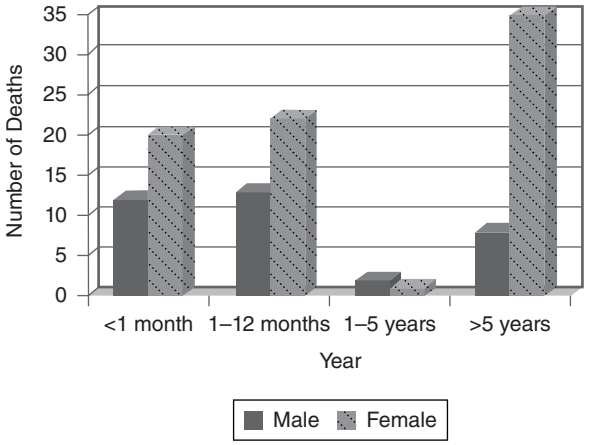


FIGURE 12-6 Number of deaths among Burmese refugees in the Gundhum II Camp, Bangladesh, May 6 to June 26, 1992 by age and sex.

Data from Centers for Disese Control and Prevention (CDC). (1992). Famine affected, refugee, and displaced populations: Recommendations for public health issues. *Morbidity and Mortality Weekly Report, 41*(RR-13).

supplementary feeding programs, is an essential function of WCH services. A WCH program will also ensure that all children are vaccinated on schedule and are receiving regular supplements of vitamin A. When curative care is required, as for diarrhea and acute respiratory infections, it can be offered at the household level by trained CHWs, or the child can be referred to peripheral, first-line health facilities.

Pregnant women (who typically constitute approximately 3% of the refugee population) should be regularly monitored. At least four prenatal examinations should be conducted to identify high-risk pregnancies. All women should be vaccinated with tetanus toxoid to prevent neonatal tetanus in their newborns. Iron and folic acid should be distributed (and their ingestion monitored, if possible). Insecticide-treated bed nets for the prevention of malaria should be distributed, and the presumptive treatment of malaria, if appropriate, should be undertaken in the second and third trimesters. In the postnatal period, counseling services addressing a variety of issues, from family planning to child care, and especially breastfeeding, should be offered.

Although many elements of women's and child health can be instituted in the post-emergency phase, a critical service that must be provided during the earliest stages of a relief effort is the establishment of emergency obstetric care. Cesarean section and transfusion facilities and the ability to give parenteral antibiotics are essential if maternal mortality is to be kept low. Provisions for emergency delivery are part of an overall minimum initial service package, discussed in the next section.

Reproductive Health

UNHCR, WHO, and the UN Population Fund (UNFPA) state that although food, water, and shelter remain priorities in an emergency assistance program, reproductive health care is among the crucial elements that give refugees the basic human welfare and dignity that is their right (UNHCR, 1995a). The response to reproductive health problems during emergencies consists of a constellation of assessment, services, and regular monitoring that addresses the implementation of the following programs:

- A minimum initial service package (MISP)
- Safe motherhood
- Prevention and treatment of sexual and gender-based violence
- Prevention and care for sexually transmitted diseases
- Family planning
- Reproductive health needs of adolescents

Minimum Initial Service Package

Although resources should not be diverted from attempts to control the diseases that have traditionally been the leading causes of death in CEs, five interventions related to reproductive health should be implemented even in the acute phase. In addition, a reproductive health coordinator should be designated to ensure that these measures are adequately addressed (http://www.who.int/reproductive-health/). Note that specific guidelines have been developed for the prevention and response to gender-based violence (UNHCR, 1995b).

Emergency Contraception

Forced migration is frequently accompanied by sexual violence. To prevent unwanted pregnancies resulting from rape, emergency postcoital contraceptive supplies should be available to women who request them. Although the extent of this problem has never been adequately documented, and although it is likely to vary from one situation to the next, the current recommendation is for sufficient supplies for 1% of women of reproductive age to be immediately available. Two methods of emergency contraception are currently available: (1) the combined oral contraceptive (two formulations are used), which must be taken within 3 days of unprotected sexual intercourse, and (2) the copper intrauterine device, which must be implanted within 5 days of unprotected intercourse.

Universal Precautions

To prevent the transmission of HIV, universal precautions must be respected from the outset of an emergency. Although chaotic conditions are frequently prevalent and although health services are implemented under stressful conditions during CEs, the threat of HIV infection can and must be minimized. Universal precautions include the following:

- Washing hands with soap and water after contact with body fluids or wounds
- Wearing gloves for procedures involving contact with blood or other body fluids
- Using protective clothing when exposure to large amounts of blood is likely
- Handling sharps safely
- Disposing of waste materials safely (burning and/or burial)
- Disinfecting or sterilizing medical equipment
- Wearing gloves when handling corpses, and washing with soap and water afterward

Delivery Services

All populations affected by CEs will include women in the later stages of pregnancy or who are at high risk for complicated deliveries. These women need services even during the acute phase of the emergency. In a population of 2,500 with a crude birth rate of approximately 3%, there will be 5 to 8 births per month. To deal with these deliveries, simple supplies must be made available. Both simple delivery kits and midwife kits are readily available from UNICEF and other suppliers of health supplies. Skilled attendants should be present at every delivery.

In addition, although it is not a formal part of the MISP, sites where complicated deliveries can be performed should be established as early as possible during an emergency. Cesarean section for obstructed delivery, transfusion for excessive hemorrhage, and parenteral antibiotics for the treatment of sepsis are the only ways to reduce maternal mortality, and these measures should be available and easily accessible from the onset of a relief operation.

Sexually Transmitted Infections

HIV and AIDS and other sexually transmitted infections (STIs) are major problems among emergency-affected populations from areas where there is a high prevalence of these conditions. Until recently, there was a paucity of data on the prevalence of HIV in refugee populations. It had been postulated that refugees and other conflict-affected populations might be at greater risk of acquiring HIV because of sexual exploitation, the breakdown of traditional societies and values, and the disruption of STI treatment and condom promotion programs. Nevertheless, some evidence indicates that conflict might actually inhibit the spread of HIV (Spiegel et al., 2007). For example, when the 20-year conflict in Angola ended in 2002, the country had a significantly lower HIV prevalence (5% to 10% in Luanda and 1% to 3% in rural areas) than that in all other southern African countries (15% to 40%).

Between 2002 and 2004, UNHCR surveys of pregnant women in 20 refugee camps in Kenya, Rwanda, Sudan, and Tanzania found lower HIV prevalence among refugees than in the surrounding population in three of the four countries (e.g., 5% versus 18% in Kenya), and in Sudan there was no significant difference (Spiegel, 2004). The UNHCR report notes that most refugees moved from low- to high-prevalence countries, that most refugees lived in remote rural areas with restricted freedom of movement, and that NGOs had mounted HIV prevention programs targeted at easily accessible, "captive" populations.

To minimize the transmission of STIs, including HIV, an adequate supply of condoms should be available on request to all members of the target population. Several guidelines have been developed for the prevention of HIV in refugee and post-emergency settings (Holmes, 2003; Inter-Agency Standing Committee Task Force on HIV/AIDS in Emergency Settings, 2003; Sphere Project, 2011).

Comprehensive Reproductive Health Services

The last element of the MISP is planning for the provision of comprehensive reproductive health services as rapidly as is feasible. To do this, reproductive health indicators should be included in health information systems to allow for the collection of baseline data on maternal, infant, and child mortality; prevalence of STIs; and population contraceptive prevalence rates. Suitable sites for the delivery of reproductive health services should be identified, though they need not be separate from other health facilities. These sites should be secure and allow for safe passage of women between their homes and the clinic. They should be easily accessible to all who wish to use them, guarantee privacy and confidentiality, and have access to clean water and latrines. In addition, an adequate referral system should be identified or established to provide care to women with obstetric complications or other health emergencies.

Post-Emergency Reproductive Health Programs

When adequate food, water, sanitation, and shelter have been provided, programs to address the initial emergency health priorities have been implemented (e.g., measles vaccination), and mortality rates have declined, the range of reproductive health services should gradually be expanded. In the post-emergency phase, health care would be quite incomplete unless the entire range of programs aimed at maintaining or improving reproductive health was included. The key to successful reproductive health programs is soliciting active participation of as many adolescents and women of reproductive age as possible. Although training female health workers (including traditional birth attendants) is usually the focus of such programs, male health workers should also be trained in the basic principles of reproductive health care. In addition, because adequate reproductive health services are frequently unavailable to the local, non-refugee population, it is a good idea to try to extend whatever programs are made available to refugees and displaced persons by expatriate relief workers to women and adolescents in the host population.

Communicable Disease Control

Concern for the potential impact of communicable diseases has dominated the public health response in many emergency settings. As discussed previously, this attention has frequently been warranted. Although many of the technical interventions and public health programs used in emergencies draw heavily from their counterparts in stable settings, a few important differences should be considered. Most important among them include addressing the needs of the local, nondisplaced, population; maintaining respect for national health policies when dealing with refugees (but not necessarily adhering to them); and promoting substantial community involvement as early as is feasible.

Measles

Measles was traditionally among the most feared of communicable diseases in emergency settings. During the 1970s and 1980s, measles epidemics were common, and it was not unusual for measles to be the major cause of mortality in large, displaced populations. Major outbreaks of measles have been uncommon in refugee camps since the late 1980s because of the high priority afforded to mass measles vaccination campaigns in the acute phase of humanitarian emergencies. However, major measles outbreaks, resulting in high mortality rates, were reported among IDPs in Ethiopia and the Democratic Republic of Congo during 2000 and 2001 because of delays in initiating vaccination campaigns (Salama et al., 2001).

Measles remains a significant problem in populations that cannot be reached for immunization because of armed conflict. In the case of Somalia, child immunization has been banned by the radical jihadist group Al Shabab. There have been major outbreaks of measles in Somalia with thousands of cases in 2014 and 2016, continuing into 2017 (WHO, 2017b).

Of the 1.8 million Syrian children born since the conflict in that country began, more than 50% are unvaccinated against measles. Measles has swept through Syria, including Aleppo and the northern regions, with more than 7,000 confirmed cases occurring between 2015 and 2016, and cases also reported in refugee camps in Jordan and Lebanon. Large outbreaks of measles have also occurred in war-torn South Sudan and Yemen (UNICEF, 2017; ReliefWeb, 2017b).

Because of the devastating impact that measles has had in many emergencies, it has become almost universally accepted that mass measles vaccination, regardless of vaccination history or place of provenance of the displaced, should be instituted as early during an emergency as possible. Leading reference publications (CDC, 1992; MSF, 1997; Sphere Project, 2011) accord measles vaccination the highest priority of all health-specific interventions and recommend that it be undertaken immediately after an initial rapid assessment regardless of the circumstances. If children cannot be vaccinated upon arrival or registration, a mass vaccination campaign should be undertaken. In general, this rule should be followed unless extenuating circumstances exist, such as the cholera epidemic in Goma in 1994 that required the full attention of health services workers.

Any mass vaccination program must be planned carefully if it is to be implemented successfully. Sufficient vaccine should be on hand (with a reserve in case of excess wastage) and stored in a functioning cold chain of adequate capacity. Only autodestruct syringes should be used, and safety boxes must be available for their storage and disposal.

The target population for measles vaccination in emergencies is usually children age 6 months to 15 years. WHO has recommended reducing the usual minimum age for measles vaccination from 9 months

to 6 months during CEs, because high attack rates and high CFRs have occurred in younger children, especially in large displaced populations living in relatively crowded conditions. Still, because vaccine efficacy in children age 6 to 9 months may be lower than optimal owing to the persistence of passively transferred maternal antibodies to measles virus, children in this age group should be vaccinated again at the age of 9 months. The upper age limit for mass vaccination is more flexible and depends, to a large extent, on the amount of measles vaccine, injection equipment, health personnel available, and the pressure of competing healthcare priorities. Because age and undernutrition are such important risk factors for complicated measles and for high CFRs, all children up to 15 years old who are eligible for selective feeding programs or who are hospitalized with other illnesses should be vaccinated against measles on a priority basis. Then, depending on the factors mentioned earlier, all children younger than 2 years should be considered for vaccination, followed by all children younger than 5 years. Finally, if the circumstances allow, the target population can be expanded. In any case, a mass vaccination campaign should seek to achieve at least 95% coverage of the target population.

Because a mass vaccination campaign can reach such a high proportion of the most vulnerable population, there are frequently demands to attach other services to it. Vitamin A, for example, can be offered to the same target group during the course of the campaign. Insecticide-treated bed nets and hygiene kits can be distributed to each household in which a child resides. There have been suggestions to provide polio vaccination along with measles vaccine, although the logistical burden of doing so must be carefully considered. In any case, a routine vaccination program for all children using the standard antigens recommended by WHO should be established during the post-emergency period. Other vaccines, such as yellow fever and meningitis, are effective in interrupting transmission of these infections during an epidemic, but should not be offered routinely at the time of measles vaccine. Cholera vaccination has also been recommended at the time of measles vaccination, as discussed in the following section.

The early detection of measles cases when they occur is an important feature of an effective community-based surveillance system. Measles treatment includes the administration of two doses of vitamin A and the appropriate treatment of common complications such as pneumonia, diarrhea, malnutrition, and meningoencephalitis. Children with measles should be closely monitored in regard to their nutritional status and, if indicated, should be enrolled in supplementary feeding programs during their convalescence.

Measles remains an important threat to the health of children in many emergency settings. However, as vaccination programs in many parts of the world have progressed and as vaccination coverage levels increase, measles vaccination should be considered alongside other priority interventions. In northern Iraq and in the Balkans, as in Goma, measles vaccination was delayed to address other, more urgent problems. In contrast, in Afghanistan (2002), in Darfur (2004), and in Aceh province, Indonesia, in the wake of the tsunami (2005), measles vaccination was among the first interventions organized and implemented. Nevertheless, despite the clear threat that measles poses to the health of populations in emergency settings, it is always appropriate to weigh the public health needs in light of the available resources and to order priorities accordingly. The guidelines titled *Vaccination in Emergencies: A Framework for Decision-Making*, which have been developed and disseminated by WHO (2017e), should be consulted at the start of any emergency health sector intervention.

Diarrheal Diseases

Unlike measles, which can be easily prevented with a vaccine, diarrheal diseases are often among the top three causes of mortality in humanitarian emergencies. Although diarrhea is mainly a condition of young children, cholera and dysentery—the major epidemic forms of diarrhea—affect people of all ages. Of all disease conditions, diarrhea is the most closely linked to poor sanitation, inadequate water quantity, and contaminated water.

The detection and reporting of diarrhea should be part of the routine surveillance system in all emergencies. Both acute watery diarrhea and bloody diarrhea should be reported separately by age (younger than 5 years and older than 5 years are minimum age groups—the elderly should be disaggregated if possible) and by gender.

All health personnel should be sensitized to the potential impact of diarrhea and should be skilled in most aspects of prevention and in treatment. The key to prevention lies in providing adequate sanitation facilities and at least the minimum recommended quantity of water of acceptable quality (see the "Water and Sanitation" section earlier in this chapter).

The mainstay of diarrhea case management is oral rehydration therapy (ORT). Although any fluids can be used to prevent the development of dehydration, low-osmolar oral rehydration salts (ORS) can be used

in all cases and are the treatment of choice for all levels of dehydration (see the *Infectious Diseases* chapter). In fact, the first large field trial of ORS took place in a refugee camp in West Bengal, India, where it was shown that cholera patients treated with the standard treatment at the time were 3.8 times as likely to die from dehydration as those treated with ORS (Mahalanabis et al., 1973). Zinc should be administered to all children with diarrhea for 10 days, if possible, to reduce the severity and duration of each episode, and to lengthen the time between episodes (Baqui et al., 2002).

Rehydration facilities should be available in all health facilities, including health posts and outreach sites within the community. Keys to the success of ORT in emergencies, where the caseload can be substantial, include careful organization of ORT centers and the presence of concerned and skilled staff. Mothers or other caretakers are important contributors to ORT and must be instructed as to the quantity of fluid that their children require. Assertive administration of ORT, especially for children who are tired and reluctant to drink, is essential. Breastfeeding should be continued, and the nutritional status of children recovering from diarrhea must be carefully monitored. Rehydration of unaccompanied children should be carefully overseen and appropriate follow-up ensured (see the *Understanding and Acting on Social Determinants of Health and Health Equity* chapter).

Cholera

Cholera epidemics have occurred frequently in emergency settings. Although deaths due to non-cholera watery diarrhea have been far more numerous, cholera, in addition to being able to cause death rapidly from dehydration, incites fear and even panic in many populations. Its ability to affect other relief activities and to divert health personnel and supplies from other activities may even contribute to higher death tolls due to other diseases. Outbreaks of cholera have occurred in all parts of the world; large outbreaks were recorded in India (1971), Thailand (1979), Sudan (1985), Somalia (1985), Ethiopia (1984), Malawi (1988–1991), northern Iraq (1991), Goma (1994), and Rwanda (1996).

In October 2010, nine months after the January 2010 earthquake, an outbreak of cholera started in Haiti, which quickly spread across the country. As of January 2014, 697,256 cholera cases and 8,534 deaths had been reported by the Haitian Ministry of Health (2014). One of the largest cholera epidemics in history, it occurred in a country where there had not been a cholera case for more than a century. Six years later, the United Nations acknowledged that cholera had been introduced to Haiti by UN peacekeepers from Nepal.

Most recently, outbreaks have occurred in Tanzania, Ethiopia, and Iraq in 2015, and in South Sudan, Somalia, and Yemen in 2016 and 2017. WHO has described the cholera outbreak in war-torn Yemen as the worst ever in the world (WHO, 2017c). By mid-September 2017, WHO was reporting 750,000 cumulative cases. In the first three months of 2017, WHO reported a cumulative number of 17,211 cases and 388 deaths from cholera in Somalia, with a case-fatality rate of 2.25%. Cholera is also widespread in South Sudan: As of May 2017, a total of 7,735 cholera cases including 246 deaths (CFR 3.23%) had been reported in South Sudan, involving 19 counties in 11 states, since the initial case was reported on June 18, 2016 (WHO, 2017d).

Investigations of cholera outbreaks have documented numerous modes of transmission and risk factors, including contaminated water, shared water containers, inadequately heated leftover food, insufficient soap, and funeral gatherings for cholera victims. One of the most lethal cholera epidemics ever recorded occurred among refugees in Goma, Zaire, in 1994, when it was estimated that 45,000 people (approximately 9% of the total population) died in a 3-week period. The source of contamination is believed to have been Lake Kivu, the principal source of water on which the population depended.

Early detection of possible cases of cholera is key to the effective management of an epidemic. Although non-cholera diarrhea is a far more common cause of morbidity and mortality in children, the death of an adult from dehydration should raise suspicions of cholera. Attack rates can be higher in refugee camps than in non-camp situations, but in all situations case-fatality rates can be kept as low as 1%. Laboratory confirmation should be obtained as quickly as possible at the start of a suspected epidemic, but need not be continued—even though most cases in a cholera epidemic, especially in areas with a high background rate of watery diarrhea, may not actually be due to cholera. Whenever cholera is suspected, aggressive attempts to educate the community should be made so as to limit the panic that frequently accompanies this disease. During the course of an epidemic, cases and deaths should be reported on a daily basis through the institution of an active surveillance mechanism.

In populations located in areas where cholera is a risk (e.g., if fecal culture-confirmed cholera diarrhea has occurred in three of the previous five years), WHO now recommends mass immunization with killed oral

cholera vaccines.[1] Although all age groups are vulnerable to cholera, immunization in resource-limited areas should be targeted to high-risk children age 1 year or older (Shanchol or Euvichol) or age 2 years or older (Dukoral) (WHO, 2017f). Although the availability of oral cholera vaccines has increased rapidly in recent years, mass immunization campaigns should not disrupt or delay other critical prevention programs, such as clean water and sanitation.

Cholera vaccine is now used both prophylactically and reactively in many emergency settings. Large shipments of vaccine from the well-established global vaccine stockpile have been sent to CE situations as diverse as those in Sierra Leone (in anticipation of cholera after a series of floods and mudslides devastated parts of Freetown in 2017), northern Nigeria (to prevent outbreaks in the Boko Haram–menaced areas of Borno State), and war-torn Yemen, where initial shipments of vaccine were refused by the Yemeni government.

The need to establish rehydration facilities at multiple sites within the community has been dramatically highlighted by the occurrence of epidemics of cholera. In an outbreak in Somalia in 1985, a new camp with only a centralized treatment facility and no trained CHWs reported a case fatality rate of 23.3%. In contrast, in seven camps in which peripheral ORT corners with trained CHWs had been established in the framework of a primary healthcare system, the CFR was limited to 2.4%. Even more dramatically, during the devastating outbreak in Goma in 1994, more than 90% of the approximately 45,000 deaths that occurred during a 3-week period occurred beyond the reach of the health system. Active case finding and rehydration therapy within the community, rather than reliance on overwhelmed and understaffed health facilities, may have averted a significant fraction of these deaths.

Although as many as 90% of patients during a cholera epidemic can be treated orally, intravenous rehydration will be required for the most severe cases. A referral system for such care must be in place, and cholera treatment sites should be identified and prepared with adequate bed capacity, human resources, water, drugs and other supplies, and disposal facilities. The treatment of cholera is the same as is described in the *Infectious Diseases* chapter. In emergency settings, however, selective chemoprophylaxis is usually not indicated. Resources can be used more efficiently and effectively in other ways, such as by establishing adequate water and sanitation and ensuring that all patients are identified and treated quickly and appropriately.

Dysentery Due to *Shigella dysenteriae* Type 1

The management of epidemics of dysentery in emergency settings is difficult. As is true of other diarrheal diseases, ensuring adequate water and sanitation facilities is essential, but because of the highly communicable nature of *Shigella dysenteriae* type 1, this measure's role in reducing transmission may be limited, especially in the crowded conditions of refugee camps. Nevertheless, the use of narrow-neck containers for water storage to reduce contamination and the distribution and use of soap for hand washing have been shown to be useful. Early case detection and prompt treatment are the keys to limiting spread. An epidemic of *S. dysenteriae* type 1 should be suspected whenever a case of diarrhea with blood in the stool is reported. Laboratory confirmation and sensitivity to antibiotics should be obtained immediately, with careful attention paid to the transport of stool specimens.

The key to dysentery case management is antibiotic therapy, but severe limitations may prevent effective case management on a large scale. These include the large caseloads that may require treatment, the resistance of organisms to first-line and even second-line antibiotics, and the difficulty of ensuring patient compliance with 3- to 5-day courses of treatment.

In general, during epidemics of *S. dysenteriae* type 1, an effective antibiotic such as ciprofloxacin should be given to all patients, under close supervision of health staff. If supplies of an effective antibiotic are limited, patients who are severely ill or most vulnerable (e.g., children, pregnant women, the elderly) should be given antibiotics, and others given only supportive treatment—that is, nutritional support, rehydration when necessary, and other specific measures. Oddly, no outbreaks of *S. dysenteriae* type 1 have been reported from emergency settings in recent years.

Acute Respiratory Infection

Acute lower respiratory infection (ALRI), or pneumonia, has been an important cause of morbidity and mortality in emergency settings. Risk factors have included crowded conditions, inadequate shelter, vitamin A deficiency, and indoor air pollution, especially in societies that cook indoors (such as Nepal). ALRI is undoubtedly a major cause of morbidity and mortality in cold climates, such as northern Iraq, the Balkans, and the war-torn former Soviet republics. Such infection is the leading cause of death among

[1] Dukoral, Shanchol, or Euvichol.

children in LMICs, but has been less consistently reported and investigated than many other communicable diseases in emergency settings.

The management of ALRI is the same as in non-emergency settings, with prompt antibiotic treatment and supportive care. Community health workers need to be trained to diagnose ALRI in children, based on signs such as chest indrawing and a rapid respiratory rate.

Malaria

In endemic areas, including Southeast Asia, the Indian subcontinent, and most of Africa, malaria is consistently among the leading causes of morbidity and mortality. It has been responsible for incidence rates as high as 1,037 per 1,000 per year (Burmese refugees in Thailand, 1984) and for as many as 30% of all deaths in displaced populations (Rwanda, 1994). It has been well established that populations that are displaced to areas of higher malaria endemicity than their place of origin have higher incidence rates and higher mortality from malaria.

A study of refugee sites in Burundi, Chad, Cameroon, Ethiopia, Kenya, Sudan, Tanzania, Thailand, and Uganda to describe trends in malaria incidence and mortality was published in 2011 (Anderson, Doocy, Haskew, Spiegel, & Moss, 2011). An average of 1.18 million refugees resided in 60 refugee sites, with at least 50 cases of malaria occurring per 1,000 refugees during the study period 2008–2009. The highest incidence of malaria was in refugee sites in Tanzania, where the annual incidence of malaria was 399 confirmed cases per 1,000 refugees and 728 confirmed cases per 1,000 refugee children younger than 5 years. Annual malaria mortality rates were highest in sites in Sudan (0.9 death per 1,000 refugees), Uganda and Tanzania (0.7 death per 1,000 refugees each). Malaria was the cause of 16% of deaths in refugee children younger than 5 years of age across all study sites.

Major risk factors for malaria in refugee situations include the lack of adequate housing, poor siting of camps (especially when they are placed in marshy areas), overcrowding, proximity to livestock (which may be the primary targets of mosquito vectors), and a general lack of competently trained health personnel. Malaria control in emergencies depends to some extent on knowledge of the local vectors. In any case, site planning and selection should be done with the possibility of malaria in mind, and areas with swamps, marshes, and other characteristics that favor vector breeding should be avoided. Where mosquito density is high and immunity of the population is low, periodic residual spraying of interior walls can be undertaken, although it is less effective where temporary and shoddy shelters are in use and where mosquitoes bite and rest outdoors. Aerial spraying should usually be avoided except in special circumstances, to avoid further traumatizing the population. In all circumstances, long-lasting insecticide-treated bed nets should be distributed (Bloland & Williams, 2003). In emergencies where people are displaced with their livestock, periodic permethrin sponging of the animals has been shown to reduce vector density and malaria transmission.

The presumptive treatment of pregnant women with pyrimethamine/sulfadoxine in a single dose during the second and third trimesters (especially during women's first and second pregnancies) is also a commonly recommended strategy. The presumptive treatment of children attending feeding centers can also be considered, and presumptive treatment of children for malaria at the time of mass measles vaccination or vaccination during routine programs may have a role in malaria control after further investigation. Whenever feasible, it is highly recommended that the diagnosis of malaria should be confirmed, either microscopically or, even better, with one of several available rapid diagnostic tests. If facilities are not available to do so, malaria should be treated on the basis of a presumptive diagnosis, although other causes of fever should be suspected as well.

In determining who should be treated, with which drugs, and according to what dosage schedule, it is important to consider the national guidelines of the host country, although malaria control policies may also be updated in accordance with the latest information regarding antimalarial drug sensitivities. In most parts of the world, artemisinin combination therapy is currently the most effective, and most recommended, treatment. Note, however, that artemisinin-resistant malaria has recently been reported from Cambodia, Laos, and Myanmar (WHO, 2017g). Strategies for uncomplicated malaria, for severe malaria, and for treatment failures should be developed and explained to all health service personnel.

Meningitis

Although not a consistent problem in emergencies, the threat of meningococcal (group A) meningitis is a formidable one. Overcrowding, especially during the drier seasons of the year, can be an important risk factor for this disease, which is transmitted via the respiratory route. Epidemics have occurred in Thailand (1980), Sudan (1989), Ethiopia (1993), Guinea (1993), and Goma; in the last case, attack rates ranged from 94 to 137 per 100,000 population over a period of

2 months. Outbreaks of meningitis tend to be protracted, lasting from 1 to 2 months. Unless they are detected and controlled at an early stage, they can be directly responsible for high mortality; in addition, patients can require resource-intensive care, which diverts attention from other high-priority health programs. In early 2005, an outbreak of the emerging W135 strain of meningococcus led to the vaccination of more than 150,000 Sudanese refugees in Chad.

The detection of outbreaks of meningococcal meningitis at an early stage is essential. During emergencies, a high level of suspicion should be maintained. All clinically suspicious cases should be diagnosed by either visual inspection of cerebrospinal fluid or, where available, by the appropriate microscopic, serologic, and bacteriologic analyses. Background rates of meningococcal disease vary considerably from one area to another, and the occurrence of disease is highly seasonal. For these reasons, the detection of an epidemic requires a sensitive surveillance system. It has become customary to institute epidemic control measures when a threshold incidence rate of 15 cases per 100,000 population per week has been exceeded for 2 consecutive weeks. In small populations, or where the population has not been accurately determined, a weekly doubling of the number of cases over a 3-week period can signal the early stages of an epidemic. WHO has established both alert and response thresholds that should be reviewed and adhered to when emergencies occur in meningitis-prone areas.

Vulnerability to meningococcal meningitis has decreased markedly since the development and widespread distribution of an effective, affordable vaccine. In December 2010, a new inexpensive and effective meningococcal A conjugate vaccine (MenAfriVac) was introduced, first in Burkina Faso and selected regions of Mali and Niger. By 2016, all 26 countries in the African meningitis belt had introduced this vaccine. High coverage of the target age group of 1–29 years is expected to eliminate meningococcal A epidemics from this region of Africa.

In non-immune populations, mass vaccination campaigns are the intervention of choice in areas in which an epidemic is occurring. Vaccination campaigns usually target the entire population age 1 year and older, although resource constraints may require limiting the age group to be vaccinated. Neither mass chemoprophylaxis nor prophylaxis of household contacts has proved to be an effective intervention during outbreaks, and neither should be instituted.

When epidemics occur in Africa, ceftriaxone is the drug of choice for its treatment. Penicillin, ampicillin, and chloramphenicol are also effective.

Hepatitis E

Like meningitis, outbreaks of hepatitis E have not been frequent occurrences in emergencies but have had major consequences when they arose. An enterically transmitted disease, usually linked to contaminated drinking water, especially when water quantity is compromised, hepatitis E is associated with a particularly high CFR in pregnant women. Clinical attack rates appear to be higher in adults, with children being relatively spared.

A severe epidemic occurred in 2004 among both IDPs in Darfur, Sudan, and refugees from Darfur in Chad. Starting in West Darfur in May of that year, a total of 2,431 cases and 41 deaths (CFR 1.7%) of suspected hepatitis E were reported from health clinics in the Greater Darfur region. A survey conducted in November 2004 found a CFR of 8.2% among pregnant women with hepatitis (Boccia, Klovstad, & Guthmann, 2004). Between June and September 2004, a total of 1,442 cases of suspected hepatitis E and 46 deaths (CFR 3.2%) were reported from two refugee camps and neighboring villages in Chad. Investigation of the epidemic was aided by the use of a newly licensed rapid diagnostic kit.

The most recent outbreak of hepatitis E in Africa occurred in the Lake Chad Basin region, including hundreds of cases in Chad and Niger. The security of this area has been considered unstable for many years due to the presence of Boko Haram forces (jihadists based in northeast Nigeria).

There is no treatment for hepatitis E. Thus, the response to an outbreak is prevention through ensuring access to adequate quantities of clean drinking water and hygiene promotion in communities.

Tuberculosis

Tuberculosis (TB) is one of the most important communicable diseases to control in the post-emergency phase. Its re-emergence as a public health problem in many parts of the world has been characterized by its close association with immune deficiency disorders, especially HIV infection, and with the identification of multiple drug-resistant strains.

Prior to the large movements of refugees in the Middle East, more than 85% of refugees originated from, or remained within, countries with high burdens of TB, such as Pakistan, DRC, Ethiopia, Myanmar, Kenya, Afghanistan, and South Sudan (WHO & UNHCR, 2007). In populations displaced by conflict in the Middle East, TB has been a lesser problem. For example, TB screening of 69,000 Syrian refugees in Jordan from January to June 2014 found only three smear-positive cases and a total of 33 culture-confirmed cases (Cookson et al., 2015).

Tuberculosis control should be instituted only after mortality rates have fallen below 1 per 10,000 per day or when an emergency situation has stabilized and it is apparent that the displaced population will remain in its current location for at least 6 months. However, for individuals being treated for tuberculosis at the time of onset of the emergency, arrangements should be made to allow them to continue their treatment in an uninterrupted manner.

From a public health standpoint, the objectives of TB control are to treat patients so that they cannot infect others, and to restore health to infected individuals. For this reason, only sputum smear–positive individuals are usually included in TB control programs initially, although individuals who are severely ill with noninfectious, extrapulmonary forms of TB should also be treated if they are identified. WHO guidelines for treatment stress directly observed therapy with a short course (6 to 8 months) of a combination of antituberculosis drugs (WHO & UNHCR, 2007). Because both infection with HIV and malnutrition are associated with TB, the presence of these conditions should be determined and dealt with appropriately.

Programs geared toward treating and preventing TB are complicated, and the decision to implement one should not be made unless there are clear written guidelines that will be followed. Laboratory facilities must be available and the regular provision of supplies ensured. Drugs must be stocked and their resupply guaranteed. Finally, a system for tracing those persons who are unable to adhere to treatment regimens must be in place so that they can be identified and assistance provided to ensure treatment completion. Ultimately, successful implementation of a TB control program requires a high level of community awareness, education, and involvement. Every element of such a control program needs to be carefully and meticulously developed and nurtured over time. Agencies that intend to implement TB control programs in post-emergency settings should have a clear commitment to continue for at least 12 to 15 months, have an adequate budget, and have the personnel and material resources necessary to run a successful program.

Other Communicable Diseases

Other communicable diseases that have occurred in emergency or post-emergency settings have had a relatively minor impact on refugees. In the individual setting in which they occur, however, they command an important allocation of resources and may be important contributors to morbidity and mortality. Yellow fever, typhoid fever, relapsing fever, Japanese B encephalitis, dengue hemorrhagic fever, typhus, and leptospirosis are all real threats. Outbreaks of cutaneous leishmaniasis have occurred in war-torn Syria. The last reservoirs of wild poliovirus are found in conflict-affected populations in Afghanistan, Pakistan, and Nigeria. Nevertheless, morbidity and mortality in CEs has been shown time and again to be due to the same conditions that are responsible for the bulk of the disease burden in LMICs in nonemergency settings.

Noncommunicable Diseases

As noted earlier in this chapter, NCDs such as diabetes and hypertension have been the main cause of morbidity and mortality among Syrian, and many Iraqi, refugees and IDPs. Research and evidence are lacking on how to best address care for NCDs in emergencies. To make interventions more effective and sustainable, further NCD-oriented research is needed (Perone et al., 2017). Given the lack of sophisticated medical resources in most emergency settings, it is suggested that NCDs be addressed through a primary healthcare setting combining health education messages with clinical monitoring of patients.

Another important element of a humanitarian response is to address mental health issues. A growing body of evidence has been used to inform effective mental health programs in emergencies, and interagency guidelines have been developed (Inter-Agency Standing Committee, 2007).

Role of International, National, and Nongovernmental Organizations

The vast and complex array of organizations involved in the various stages of humanitarian emergency preparedness and response reflect the complexity of the international community itself. It is difficult to imagine any other situation that attracts such a range of players: heads of state; diplomats; celebrities; bilateral foreign assistance agencies; UN political, social, economic, and technical organizations; military forces; and a broad variety of nongovernmental organizations, including an increasing number of commercial interests.

The UN Security Council plays a critical role in determining how and when the world will respond to the conflicts that lead to CEs and how emergency humanitarian assistance programs will be protected from the forces that fuel the conflicts themselves.

Security Council decisions are, in turn, determined by the leaders of its member states—in particular, the permanent members (the United States, Russia, France, the United Kingdom, and China). The UN Office for the Coordination of Humanitarian Affairs (OCHA) coordinates emergency preparedness and response within the UN system and is steered by an Inter-Agency Standing Committee of relevant UN agencies, including UNHCR, WHO, UNICEF, the World Food Programme (WFP), and UNFPA, as well as NGOs. OCHA launches joint appeals for funds to support coordinated UN agency response programs; however, these may sometimes compete with appeals launched by individual agencies.

The UN Refugee Agency, UNHCR, is responsible for the protection and care of all refugees who cross international borders. In some emergency situations, such as the crises in the former Yugoslavia, the UN secretary-general has designated UNHCR as the lead relief agency; UNHCR has referred to those who do not qualify for assistance as refugees as "persons of interest." Nowadays, UNHCR is the primary agency responsible for the care of most IDP populations.

In general, UNHCR and other lead relief agencies need to be invited to provide assistance by the government of the affected country. In situations of internal conflict, the Security Council may authorize involvement by a UN agency without the approval of the host government, as was the case concerning the displaced Kurdish population in northern Iraq in 1991. The lead UN agency is mandated to coordinate the activities of other relief agencies, including NGOs, in cooperation with the host government where that type of collaboration is appropriate. However, in certain chaotic emergency settings, individual NGOs have sometimes negotiated involvement directly with government authorities and ignored efforts by the lead agency to coordinate activities.

United Nations Reforms

Between 2005 and 2010, the United Nations instituted a set of humanitarian reforms. The three major points of the reforms, which are intended to improve the coordination of humanitarian relief efforts and to streamline the chaotic funding streams, are discussed here.

First, the reforms call for the appointment of a humanitarian coordinator in every situation from which an appeal for emergency funds will emanate. Every UN agency, and NGOs that wish to participate, can submit an estimate of its financial needs to the humanitarian coordinator, who aggregates these requests into a single consolidated appeal. Ideally, this should reduce the level of competition between the individual UN agencies, and between the UN agencies and the NGO community, and provide donors with a clear and agreed upon approximation of financial needs for the relief effort.

Second, the United Nations established the United Nations Central Emergency Response Fund (CERF). CERF was created by the United Nations General Assembly in 2005. CERF is, essentially, a bank account into which donors deposit funds that the Emergency Relief Coordinator (ERC) of the UN, who is also an Under-Secretary-General of the United Nations and the director of OCHA, can draw upon to provide funds to UN agencies in a timely and efficient manner when humanitarian relief needs arise. Since its inception, CERF has provided, on average, more than $400 million per year for humanitarian relief in all kinds of emergencies—natural disasters and conflict, acute and protracted. In addition, CERF funds can be used, at the discretion of the ERC, to address "chronically underfunded" or "silent" emergencies, such as those that have plagued central Africa for most of the past decade.

Third, a system of technical "clusters" has been developed to bring a systematic approach to a previously cluttered humanitarian scene. For a range of technical areas, a UN agency has been assigned to function as the "lead agency" to bring together donors, NGOs, other UN agencies, and other interveners to develop and review policy, to formulate plans for coordinated responses, and to review and evaluate past performance. The current global clusters include, but are not restricted to, health (WHO lead), logistics (WFP), shelter (UNHCR), water and sanitation (UNICEF), nutrition (UNICEF), and food security (WFP). In the field, each cluster is co-chaired by an NGO that, together with the appropriate UN agency, works in cooperation with and under the guidance of the host country's government. The performance of the clusters has been variable to date, and external evaluation will, it is hoped, bring about improvements to what is a promising system, but one that is difficult to implement.

Host Country Role

In many emergencies, especially those involving large refugee populations, the host government has granted temporary asylum to the refugees and has become actively engaged in the relief effort. Many countries have coordinating bodies, such as relief commissions, that take a lead role in mobilizing, organizing,

and delivering relief services. For example, in Somalia in the early 1980s, the Ministry of Health formed the Refugee Health Unit (RHU), which coordinated public health and nutrition assistance to the 800,000 Ethiopian refugees scattered in 35 camps throughout the country. NGOs wishing to provide assistance to refugees had to agree to follow the technical guidelines developed by the RHU and sign a tripartite agreement with the RHU and UNHCR. During the tsunami relief effort of 2004–2005, the picture was more mixed. In India and Thailand, most of the relief effort was run entirely by national authorities. In contrast, in Indonesia and Sri Lanka, a combination of national and international organizations collaborated on policy development and on the day-to-day implementation of relief services.

The International Committee of the Red Cross (ICRC), based in Geneva, Switzerland, is mandated to carry out the protection and care of civilian populations during armed conflict, as outlined in the Geneva Conventions. The ICRC relies on low-key and confidential negotiations with all parties to a conflict to allow it to provide humanitarian assistance. The ICRC is committed to carrying out its mission with independence, impartiality, and neutrality. Although it was once the only organization to operate within areas affected by conflict, in recent years many other NGOs have joined the ICRC in taking on this challenge. Some of these NGOs, while providing relief impartially to all those in need, believe that they should also speak out in the face of gross human rights abuses and have become advocates for more effective international responses. This action may sometimes jeopardize their ability to remain in the affected area, so it is not taken lightly. One of the most hotly debated issues within NGOs is whether to provide humanitarian relief and remain silent about human rights abuses and the diversion of relief resources or to speak out and risk having to leave the area.

Many NGOs are engaged in providing humanitarian assistance in emergencies; they include national Red Cross and Red Crescent societies, international secular and religious agencies, and local churches and community-based organizations in the affected country. Specialized public health agencies such as the U.S. Centers for Disease Control and Prevention, a government agency, and the Paris-based Epicentre, the epidemiology arm of Médecins sans Frontières, provide technical advice to a range of bilateral, UN, and nongovernmental operational agencies. The level of technical skills, experience, management, and logistics capacity within NGOs varies enormously. In an effort to promote coordination and the implementation of best practices among NGOs, a number of initiatives have been introduced, including the Code of Conduct for the International Red Cross and Red Crescent Movement and NGOs in Disaster Response and the Humanitarian Charter and Minimum Standards in Disaster Response (Sphere Project, 2011). In addition to relief NGOs, an increasing number of human rights advocacy NGOs have been active in recent years, including Amnesty International, Human Rights Watch, and Physicians for Human Rights.

Funding for international humanitarian assistance programs generally comes from the governments of high-income countries, such as the United States, Japan, members of the European Union, and other OECD states. The generosity of such governments varies enormously. Although the principle of "humanity"—the provision of assistance to those in need—should govern their commitments, in reality the size and sustainability of foreign aid often depends more on the perceived geopolitical importance of the conflict than on the actual needs of the affected populations. High-profile media, such as CNN, *The New York Times*, Al Jazeera, and the BBC, often play an influential role in determining the size of the response to an emergency. For example, the blanket media coverage of emergencies in the Middle East ensures that relief programs targeting this region are relatively well funded. By comparison, the conflict in Central African Republic has largely been ignored. What is certain is that until there is a consistent response to conflict-related emergencies around the world, the quality and timeliness of humanitarian responses will be unpredictable.

Rehabilitation, Repatriation, and Recovery

In countries emerging from conflict, the costs of reconstruction may be staggering. In the immediate aftermath of the first Gulf War, it was estimated that Iraq would require $110 billion to $200 billion and Kuwait $60 billion to $95 billion to repair the war damage from the actions of the United States–led coalition (Lee & Haines, 1991). Initial estimates for the reconstruction of Afghanistan were on the order of $1 billion to 2 billion per year for at least a decade. As these conflicts have dragged on toward the end of the second decade of this century, those projections have fallen by the wayside, and there is no end in sight.

A particularly important issue for those engaged in dealing with the aftermath of conflicts is to determine the extent to which the prior health system will be simply reestablished along the lines of its previous existence, or whether it will be reformed in an effort to improve its efficiency, effectiveness, and equity. The usual response is to seek to reconstruct what has been destroyed in the conflict, although the mantra for the reconstruction

of Haiti and subsequent natural disasters has been to "build it back better." This apparently logical response may be deeply flawed, however. One key impediment is that the resource base available for reconstructing and operating the health system may be vastly inferior to what it was previously. In Uganda, the post-conflict resource base was less than 10% of that available prior to the civil war, making simple reestablishment of prior services totally unfeasible. Furthermore, much changes during the period of ongoing conflict: the range of providers operating, the role of the state, the attitudes and demands of the community that uses services, and the approach of the international community and key donor organizations. Members of the last group, for example, may seek to promote a radically different state—one that functions as a steward of the health system, running the civil service, setting policies, establishing evaluation frameworks, and so on, but not providing services, and that adopts cost-effectiveness and value for money as its core concerns. This has been the case in Afghanistan since 2003: All health services, except those in Kabul, have been contracted out to NGOs that promise to provide a standard basic package of health services. During the post-conflict phase, international financial institutions and donor governments may greatly influence policy direction, and frequently do so in favor of reducing state expenditures and enhancing the role of the private sector.

In the period between the onset of conflict and its resolution, which may last decades, approaches to the nature of health services and who purchases and provides them may change dramatically. In many conflicts, other providers enter the scene to fill the gap left by retreating and undermined state-provided health services—these include for-profit and not-for-profit providers, as well as the indigenous and traditional sector. There has been little documentation of how and why the private sector emerges to play an important role in these settings. Nor is there any clarity about how best to control and regulate such activities in the interests of ensuring that minimum standards are adhered to and that the medical treatment offered by different providers does not compromise public health goals and objectives. The emergence and changing role of the nonprofit, nongovernmental sector is easier to appreciate: It fills gaps resulting from withering state services and is often supported by donor country funds that ensure that humanitarian relief services are provided in acute emergencies and that development-oriented services are offered where suitable funding and partner organizations, including the government, can be identified.

A major weakness of NGO-provided services is that they are often poorly coordinated, act in parallel with the state systems, have a different vision of the system they are seeking to bolster or reestablish, and compete for partners, resources, and publicity. During CEs, failure to support the indigenous capacity may increase the risk of little being left behind when humanitarian agencies withdraw from the area; increasingly, there is debate regarding how best such services could interface with host government services and policy and could reinforce the limited capacity often present. A particularly important and difficult challenge is to establish the policy framework within which health services and the health system will operate.

Key issues to be debated in the aftermath of periods of conflict include the financing of health services, the extent to which those services can and should be decentralized, the role of the private sector, and the priority to be accorded to issues of equity. These concerns need to be addressed within the broader context of promoting and consolidating the peace, reestablishing the economy, facilitating the demobilization of troops and their absorption into the economy, and facilitating the return of refugees and internally displaced people. Ensuring that, to as great an extent as possible, existing inequities in distribution and access to health services are resolved in the aftermath of conflict may assist in lowering tensions between groups. Promoting the development of a more equitable health and social system may provide an important opportunity for bringing together different groups within affected populations, and may provide early opportunities to stimulate debate, exchange of ideas, and the rekindling of trust.

Gender inequalities permeate many societies; in the post-conflict environment, it may be possible to address these inequities given that the conflict itself may have changed gender relations and modified the traditional roles of men and women. Conflict often leads to women taking on a more important role in relation to making household decisions and controlling household resources, given that even in patriarchal societies, men will often be absent during periods of conflict and women will absorb a multitude of usually male-dominated roles.

An initiative to focus on the health system challenges facing countries emerging from conflict has identified a number of key priorities for intervention in this setting (Zwi, Ugalde, & Richards, 1999):

- Maximizing the contribution of both government and donors to the formulation and development of health policy
- Developing a clear conceptual framework, informed by multidisciplinary approaches, to guide health system development

- Establishing inclusive processes that involve a range of stakeholders in a participatory and transparent process of identifying needs and priorities and agreeing on models and approaches to health system development
- Appreciating the limitations and constraints operating upon the range of stakeholders (government at central and local levels, UN agencies, NGOs, traditional public- and private-sector providers) involved in financing, providing, and overseeing health services
- Promoting evidence-based policy and planning to ensure that more good than harm results from interventions and that resources are used as equitably and efficiently as possible

The Role of the Military and Other Humanitarian Actors

Traditionally the domain of international agencies and not-for-profit nongovernmental humanitarian organizations, CEs have evolved into major geopolitical theaters in which many diverse and disparate actors have sought to carve out new roles for themselves. Because an increasing number of CEs have been precipitated by armed conflict within and between nations, third-party military forces, especially those of Western nations, have been prominently involved in recent relief operations. In addition, during the 21st century, interstate warfare has become less common and the rise of nonstate militias bent on using nonconventional warfare methods has introduced new complications to the provision of humanitarian services to beleaguered populations.

Following the Gulf War of 1991, an extensive international humanitarian effort for the Kurdish population of northern Iraq was coordinated by the U.S. military, which operated under the auspices, but not the command, of the United Nations. For almost the first time, NGOs were to a large degree dependent on Western military forces (including those from Germany, the United Kingdom, France, and the Netherlands, in addition to the United States) for security, transportation, and logistics. The establishment of a secure operational area and the delegation of the delivery of relief services to the humanitarian community were important elements in bringing about a rapid response to the plight of the internally displaced Kurds. However, many NGOs, including the ICRC, were uncomfortable working so closely with the military and were forced to confront and reassess their notions of political neutrality. In addition, although their presence was positive in a number of ways, the military authorities proved to be novices when it came to humanitarian relief. They were ignorant of its basic principles, unfamiliar with appropriate relief services, and unable to promptly deliver essential supplies, such as measles vaccine, to meet the public health needs of a civilian population where maternal and child health problems were the main priority.

The military intervention in Somalia by the allied armed forces, including the U.S. military, in 1992 was launched for humanitarian reasons, with the assent of the UN Security Council. In a chaotic situation characterized by general lawlessness, severe factional combat, and the total collapse of governance, compounded by crop failure and ensuing famine, the only way to secure the delivery of essential relief was with the protection of armed forces. From a military standpoint, the intervention was perceived as a fiasco. However, many humanitarian organizations believed that the military operation contributed to decreasing the high mortality rate, at least initially. Prior to this episode, military forces had steadfastly maintained that their role should be limited to providing security for humanitarian supplies; following the Somali experience, they began to review in earnest the broader role of the military in humanitarian relief.

Following the invasion of Iraq by the United States in 2003 and the overthrow of the Saddam Hussein regime, private contractors hired by the U.S. Department of Defense assumed much of the relief and reconstruction burden. For the first time, NGOs were for the most part sidelined. This event has been the subject of intense debate in the humanitarian community, and it is not clear whether this episode is an exceptional one or whether the entire nature of humanitarian relief is in the process of tumultuous change.

Important lessons for current and future peace and stability operations can be found in the experiences of Provincial Reconstruction Teams (PRTs) in Afghanistan. PRTs are small, joint civilian–military organizations whose mission is to promote governance, security, and reconstruction throughout the country. They are managed by the United States and other members of the International Security Assistance Force (e.g., the United Kingdom, Italy, Germany, France, Australia, the Netherlands). The U.S. PRTs have stressed governance, force protection, and quick impact development projects to "win hearts and minds." Their contribution to long-term health development is dubious.

In 2017, the World Health Organization, in an attempt to provide better access for health services providers, especially emergency trauma teams, to civilian populations trapped in fighting that was taking place in Mosul, Iraq, embedded health professionals from both for-profit and nonprofit organizations

with Iraqi security forces. This allowed for the early evacuation of wounded civilians to trauma services, but also rendered relief personnel indistinguishable from one of the fighting parties. Again, this represents a clear lack of respect for the existing humanitarian principles of neutrality, impartiality, and independence, even while it may have resulted in reducing the loss of life. The ICRC and Médecins sans Frontières have argued against this clear collaboration between a UN agency and bilateral armed forces, while others have suggested that the nature of war has changed to the point that the classical humanitarian principles, as enunciated in the Geneva Conventions and other documents from which international humanitarian law is derived, may be outdated and in need of revision.

In addition to the military and other participants listed earlier in this chapter, other important agencies in addressing humanitarian issues have emerged in recent years. Organizations that specialize in monitoring, detecting, and publicizing human rights abuses and prosecuting their perpetrators have become increasingly active during CEs. These include the Office of the UN High Commissioner for Human Rights; regional organizations such as the Office for Security and Cooperation in Europe; private organizations such as Amnesty International, Human Rights Watch, and Physicians for Human Rights; and national committees supporting the international war crimes tribunals. In some situations, spanning from the Rwanda genocide in 1994 through the ongoing fighting in Syria, where human rights abuses were extremely common, these organizations have tried to ensure that public health assistance programs addressed the sequelae of these abuses.

Finally, there has been increasing criticism of existing relief organizations because of their perceived inability to implement relief programs on the scale that is frequently necessary. Some have suggested that the rapid construction, maintenance, and management of large refugee camps, global logistical support, organization of healthcare services for delivery to large populations, and even the provision of security services might be done more effectively, rapidly, and efficiently by commercial companies contracted by governments or the United Nations. This challenge to the existing relief mechanisms, based as they are on the humanitarian motive, has yet to be resolved. In many ways it could lead to the transformation of humanitarian relief into a business enterprise—one that might inevitably become more closely linked to the donor agencies and used by them as agents of foreign policy. This has been the experience of bilateral development programs.

Professionalization

Partly to stave off this challenge, and partly to correct perceptions of incompetence and amateurism, efforts have been made to establish certain minimum standards of performance for relief workers (**EXHIBIT 12-4**). Due to the transient nature of NGO relief programs and the high personnel turnover both in the field and at headquarters, experiences are not easily institutionalized and lessons need to be learned repeatedly. Limited field experience, a poor understanding of the public health priorities of emergencies, and inadequate skills to carry out the most essential tasks, such as organizing large-scale vaccination and ORT programs, have been frequently observed problems. After what is widely regarded as an initially ineffective relief effort in Goma in 1994, major efforts were undertaken to improve the technical abilities of relief workers in the public health sector.

A number of short-term training courses have been developed and implemented by schools of public health, government disaster relief agencies, and the

EXHIBIT 12-4 The Sphere Project

Perhaps the largest single effort to establish minimum standards of care in emergency settings has been the Sphere Project (www.sphereproject.org). Launched in 1997 by a group of private humanitarian agencies, Sphere recognized that humanitarian relief would be increasingly required for many years and that the existing capacity to respond with high-quality interventions was, for the most part, lacking. To address this situation, a large consortium, including more than 228 private humanitarian organizations from around the world, participated in the development of the Sphere Humanitarian Charter and Minimum Standards in Disaster Response. First published in 1999, and undergoing a third revision in 2018, these standards are intended to govern the overall conduct of relief NGOs and to provide benchmark levels of performance in the areas of water supply and sanitation, food security, nutrition, food aid, shelter and site management, and health services, including mental health and sexual and reproductive health. The Sphere Project does not intend to establish new standards, but rather seeks to consolidate and reach agreement based on existing information. Standards will continue to be developed and existing standards will be modified in accordance with new findings, both from research and from experience gained in the field.

NGOs themselves. Master of public health programs in humanitarian assistance and public health in CEs have been established in schools of public health in the United States and Europe. Although emergency public health workers are not yet required to have accredited qualifications, the quality of health care may improve as more of these training courses become available and as more people complete them.

Research

The acquisition of new knowledge relevant to public health practice in displaced populations has been scant. Although most emergency public health programs rely on the safe and effective interventions that already exist (e.g., vaccines, ORS, water purification, essential drugs, and the like), the implementation of these interventions in emergency settings may be affected by the size of the populations and the urgency of the circumstances. Little is known about the impact of rapid, forced migration on human behavior, disease transmission, and the delivery of effective services in emergency settings. For many years, it had been considered unethical to conduct research of any kind among emergency-affected populations, who could be characterized as the most vulnerable members of the world's population. More recently, however, it has been acknowledged that without applied research studies designed specifically to address operational issues in the context of emergencies, it will be difficult to reduce morbidity and mortality levels from their current, excessively high levels.

Existing standards are largely based on field experience; few are based on rigorously designed and evaluated observational field trials. Although policies in some areas, such as measles vaccination and aspects of food and nutrition, are based on field research, this is not the case in other important public health areas, such as reproductive health and the control of sexually transmitted infections. Similarly, little reliable information is available on which to base policies and programs to promote psychosocial health, despite its rapid emergence as a consistent major public health problem. Of course, research is truly useful only where there is genuine concern for improved performance.

Unfortunately, much of what is learned in humanitarian response is rapidly lost from the short-lived knowledge base. Of the many people who have worked in the field, few forge careers in humanitarian assistance. Data that are collected and reported by field workers are often either discarded or filed in internal agency reports and never seen again. There is no professional society for humanitarian public health workers, and few peer-reviewed journals in which the results of high-quality studies can be published. Although the number of people affected by CEs continues to grow, a solid body of research on which to base policy and practice is still sadly lacking. Without such a database, relief policies will remain relatively uninformed, and mistakes will continue to be made. The relatively new Research for Health in Humanitarian Crises (R2HC) initiative, funded by the U.K. Department for International Development, the U.K. Department of Health and Social Care, and the Wellcome Trust, is a potentially important contributor to the generation of new knowledge regarding what works in humanitarian relief (http://www.elrha.org/r2hc/home/).

Conclusion

Significant progress has been made during the past two decades toward the provision of effective, focused, needs-based humanitarian assistance to conflict-affected populations. Greater emphasis is now placed on the impact, including health outcomes, of international aid. The quantity of aid delivered is no longer considered a valid indicator of effectiveness; instead, its relevance, quality, coverage, and equitable distribution are accepted as more pertinent factors. As public health in emergency settings has developed as a specialized technical field, a number of relief agencies, especially NGOs, have developed technical manuals, field guidelines, and targeted training courses geared toward improving performance in this area. Ability to meet the standard performance indicators developed recently by the Sphere Project and adherence to the international NGO code of ethics are arguably valid criteria by which to assess the quality of specific agencies.

Although a considerable body of knowledge has accumulated specifically relating to the health needs of emergency-affected populations, many areas require further development. Donor agencies should acknowledge the need to support applied health research in emergency settings if more effective interventions are to be developed against both old problems, such as cholera, and emerging issues, such as noncommunicable diseases, TB, mental health, and reproductive health.

In planning for responses to future humanitarian emergencies, there is a need to recognize that improving the technical and management capacity of operational agencies will not be sufficient. Recent experience has dramatically demonstrated that populations in

need will not benefit unless the international community ensures that there are mechanisms to permit secure access by those agencies. The means by which this access is provided is critical and most likely to be the central focus of international policy dialogue. The varied nature of the responses to emergencies in Iraq, Somalia, Bosnia, Rwanda, Kosovo, Sierra Leone, East Timor, Darfur, Syria, Central African Republic, Yemen, and many others demonstrates the lack of consistency and predictability of CE-related humanitarian relief.

Finally, there remains the issue of primary prevention. The perceived differences between communities are generally tolerated in prosperous societies; conflict and all its consequences tend to arise in times of economic distress and political instability. Although programs in good governance proliferate, the reality is that governments everywhere today are perceived to have failed to provide for the basic needs of their peoples. Unless these root causes of conflict are seriously addressed, all that will be accomplished is the perpetuation of a perennial relief industry that inevitably will experience only patchy success.

Discussion Questions

1. What are the major objectives in the initial management of a refugee emergency?
2. What does the word *complex* as used in the term *complex emergency* imply?
3. What is the best indicator of the general health of a refugee population during an emergency?
4. Why are female-headed households in refugee camps at special risk of food scarcity?
5. What are the minimum standards in emergency relief operations for the provision of water and latrines?
6. What roles do general rations, supplementary feeding programs, and therapeutic feeding programs play in maintaining population nutrition?
7. At what age should children in emergency-affected populations be vaccinated against measles?
8. What are the immediate measures that can be taken in an emergency-affected population to prevent HIV/AIDS?
9. How adequate are the existing international legal statutes in protecting internally displaced persons?
10. What roles may community health workers play in an emergency public health program?
11. What are the immediate interventions that should be in place to address the reproductive health needs of both women and men?
12. Which population-based interventions have been developed to address the mental health needs of emergency-affected populations?
13. If there is conflict between adhering to humanitarian principles and saving lives, what are possible options to resolve the dilemma?

References

Action Against Hunger. (2017). Yemen crisis. Retrieved from https://www.actionagainsthunger.org.uk/europe-middle-east/yemen-crisis

Amara, A. H., & Aljunid, S. M. (2014). Non-communicable diseases among urban refugees and asylum-seekers in developing countries: A neglected health care need. *Global Health, 10*, 24.

Amnesty International. (1995). *Women in Afghanistan: A human rights catastrophe.* London, UK: Author.

Anderson, J., Doocy, S., Haskew, C., Spiegel, P., & Moss, W. J. (2011). The burden of malaria in post-emergency refugee sites: A retrospective study. *Conflict and Health, 5*, 17.

Ashworth, A. (2006). Efficacy and effectiveness of community-based treatment of severe malnutrition. *Food and Nutrition Bulletin, 27*(3), S24–S48.

Babille, M., de Colombani, P., Guerra R., Zagaria, N., & Zanetti, C. (1994). Post-emergency epidemiological surveillance in Iraqi-Kurdish refugee camps in Iran. *Disasters, 18,* 58–75.

Baqui, A. H., Black, R. E., El Arifeen, S., Yunus, M., Chakraborty, J., Ahmed, S., & Vaughan, J. P. (2002). Effect of zinc supplementation started during diarrhoea on morbidity and mortality in Bangladeshi children: Community randomised trial. *British Medical Journal, 325*(7372), 1059.

Bloland, P. B., & Williams, H. A. (Eds.). (2003). *Malaria control during mass population movements and natural disasters.* Washington, DC: National Academies Press.

Boccia, D., Klovstad, H., & Guthmann, J. P. (2004, December). *Outbreak of hepatitis E in Mornay IDP Camp, Western Darfur, Sudan.* Paris, France: Epicentre.

Boss, L. P., Toole, M. J., & Yip, R. (1994). Assessments of mortality, morbidity, and nutritional status in Somalia during the 1991–1992 famine. *Journal of the American Medical Association*, *272,* 371–376.

Bracken, P. J., Giller, J. E., & Kabaganda S. (1992). Helping victims of violence in Uganda. *Medicine and War*, *8*(3), 155–163.

Bradol, J.-H., & Le Pape, M. (2017). Humanitarian aid, genocide and mass killings: Médecins sans Frontières, *The Rwandan experience, 1982–1997*. Manchester, UK: Manchester University Press.

Brown, V., Moren, A., & Paquet, C. (1999). *Rapid health assessment of refugee or displaced populations. Annex 3.* Paris, France: Epicentre & Médecins sans Frontières.

Burkholder, B. T., & Toole, M. J. (1995). Evolution of complex disasters. *The Lancet, 346,* 1012–1015.

Centers for Disease Control and Prevention (CDC). (1991a). Outbreak of pellagra among Mozambican refugees—Malawi, 1990. *Morbidity and Mortality Weekly Report*, *40*, 209–213.

Centers for Disease Control and Prevention (CDC). (1991b). Public health consequences of acute displacement of Iraqi citizens: March–May. *Morbidity and Mortality Weekly Report, 40*, 443–446.

Centers for Disease Control and Prevention (CDC). (1992). Famine affected, refugee, and displaced populations: Recommendations for public health issues. *Morbidity and Mortality Weekly Report, 41*(RR-13).

Centers for Disease Control and Prevention (CDC). (1993a). Status of public health—Bosnia and Herzegovina, August–September 1993. *Morbidity and Mortality Weekly Report. 973,* 979–982.

Centers for Disease Control and Prevention (CDC). (1993b). Mortality among newly arrived Mozambican refugees, Zimbabwe and Malawi, 1992. *Morbidity and Mortality Weekly Report, 42,* 468–469, 475–477.

Centers for Disease Control and Prevention (CDC). (1993c). Nutrition and mortality assessment—Southern Sudan, March 1993. *Morbidity and Mortality Weekly Report, 42,* 304–308.

Centers for Disease Control and Prevention (CDC). (1994). Health status of displaced persons following civil war—Burundi, December 1993–January 1994. *Morbidity and Mortality Weekly Report*, *43*(38), 701–703.

Centers for Disease Control and Prevention (CDC). (2008). Safe water systems. Retrieved from https://www.cdc.gov/safewater

Checchi, F., & Roberts, L. (2005, September). Interpreting and using mortality data in humanitarian emergencies: A primer for non-epidemiologists. Humanitarian Practice Network. Retrieved from http://odihpn.org/resources/interpreting-and-using-mortality-data-in-humanitarian-emergencies/

Checchi, F., Warsame, A., Treacy-Wong, V., Polonski, J., van Ommeren, J., & Prudhon, C. (2017, June 7). Public health information in crisis-affected populations: A review of methods and their use for advocacy and action. *The Lancet.* doi:10.1016/S0140-6736(17)30702-X

Chou, V. B., Friberg, I. K., Christian, M., Walker, N., & Perry, H. B. (2017). Expanding the population coverage of evidence-based interventions with community health workers to save the lives of mothers and children: An analysis of potential global impact using the Lives Saved Tool (LIST). *Journal of Global Health, 7*(2), 020401.

Ciliberto, M. A., Manary, M. J., Ndekha, M. J., Briend, A., & Ashorn, P. (2006). Home-based therapy for edematous malnutrition with ready-to-use therapeutic food. *Acta Paediatrica, 95*, 1012–1015.

Coghlan, B., Ngoy, P., Mulumba, F., Hardy, C., Nkamgang Bemo, V., Stewart, T., . . . Brennan, R. J. (2009). Update on mortality in the Democratic Republic of Congo: Results from a third nationwide survey. *Disaster Medicine and Public Health Preparedness, 3*, 88–96.

Coldiron, M. E., Roederer, T., Llosa, A. E., Bouhenia, M., Madi, S., Sury, L., . . . Porten, K. (2017). Retrospective mortality among refugees from the Central African Republic arriving in Chad, 2014. *Conflict and Health, 11*, 7. doi:10.1186/s13031-017-0110-4

Collins, S. (1993, August 4). The need for adult therapeutic care in emergency feeding programs. *Journal of the American Medical Association,* 637–638.

Collins, S., Sadler, K., Dent, N., Khara, T., Guerrero, S., Myatt, M., . . . Walsh, A. (2006). Key issues in the success of community-based management of severe malnutrition. *Food and Nutrition Bulletin, 27*(3), S49–S82.

Cookson, S. T., Abaza, H., Clarke, K. R., Burton, A., Sabrah, N. A., Rumman, K. A., . . . Naoum, M. (2015). Impact of and response to increased tuberculosis prevalence among Syrian refugees compared with Jordanian tuberculosis prevalence: Case study of a tuberculosis public health strategy. *Conflict and Health, 9*, 18.

Degomme, O., & Guha-Sapir, D. (2010). Patterns of mortality rates in Darfur conflict. *The Lancet, 375*, 294–300.

Depoortere, E., Checchi, F., Broillet, F., Gerstl, S., Minetti, A., Gayraud, O., . . . Brown, V. (2004). Violence and mortality in West Darfur, Sudan (2003–04): Epidemiological evidence from four surveys. *The Lancet, 364*(9442), 1315–1320.

Desenclos, J. C., Berry, A. M., Padt, R., Farah, B., Segala, C., & Nabil, A. M. (1989). Epidemiologic patterns of scurvy among Ethiopian refugees. *Bulletin of the World Health Organization, 67*, 309–316.

Eastern Mediterranean Regional Office (EMRO). (2015). Regional situation report, March 2015: WHO response to the Syrian crisis. Retrieved from http://www.emro.who.int/images/stories/syria/WHO_SitRep_March2015.pdf?ua=1)

Eshaq, A. M., Fotham, A. M., Jensen, E. C., Khan, T. A., & AlAmodi, A. A. (2017). Malnutrition in Yemen: An invisible crisis. *The Lancet, 389*(10064), 31–32.

Essa, A. (2017). UN peacekeepers hit by new allegations of sexual abuse. Retrieved from http://www.aljazeera.com/news/2017/07/peacekeepers-hit-allegations-sex-abuse-170701133655238.html

European Commission, Humanitarian Aid and Civil Protection Department (ECHO). (2015). ECHO factsheet: Syria crisis: July 2015. Retrieved from http://ec.europa.eu/echo/files/aid/countries/factsheets/syria_en.pdf

European Parliament. (2014, November). Sexual violence in the Democratic Republic of Congo. Retrieved from http://www.europarl.europa.eu/EPRS/EPRS-AaG-542155-Sexual-violence-in-DRC-FINAL.pdf

Ferro-Luzzi, A., & James, W.P. (1996). Adult malnutrition: Simple assessment techniques for use in emergencies. *British Journal of Nutrition*, *75*(1), 3–10.

Golden, M. H. (1999). *Preventing nutritional deficiency in emergencies.* Presentation at the Enhancing the Nutritional Status of Relief Diets meeting, Washington, DC, April 28–30, 1999.

Goma Epidemiology Group. (1995). Public health impact of Rwandan refugee crisis: What happened in Goma, Zaire, in July 1994? *The Lancet, 345,* 339–344.

Goodhand, J., & Hulme, D. (1999). From wars to complex political emergencies: Understanding conflict and peace building in the new world disorder. *Third World Quarterly, 20*(1), 13–26.

Guha-Sapir, D., & van Panhuis, W. (2004). Conflict-related mortality: An analysis of 37 datasets. *Disasters*, *28*(4), 418–428.

Haiti Ministry of Health. (2014). Cholera update. Retrieved from http://mspp.gouv.ht/site/downloads/Rapport%20Web_07.01_Avec_Courbes_Departementales.pdf

Harrell-Bond, B. E., Voutira, E., & Leopold, M. (1992). Counting the refugees: Gifts, givers, patrons and clients. *Journal of Refugee Studies*, *5*(3/4), 205–225.

Heudtlass, P., Speybroeck, N., & Guha-Sapir, D. (2016). Excess mortality in refugees, internally displaced persons and resident populations in complex humanitarian emergencies (1998–2012): Insights from operational data. *Conflict and Health, 10,* 15.

Holmes, W. (2003). *Protecting the future: HIV prevention, care, and support among displaced and war-affected populations.* New York, NY: International Rescue Committee & Kumarian Press.

Hynes, M., Sakani, O., Spiegel, P., & Cornier, N. (2012). A study of refugee maternal mortality in 10 countries, 2008–2010. *International Perspectives on Sexual and Reproductive Health, 38*(4), 205–213.

Inter-Agency Standing Committee (IASC). (2003). *IASC Guidelines for HIV/AIDS Interventions in Emergency Settings*. Geneva, Switzerland: Author.

Inter-Agency Standing Committee (IASC). (2007). *IASC guidelines on mental health and psychosocial support in emergency settings.* Geneva, Switzerland: Author.

Inter-Agency Regional Response for Syrian Refugees (2013). Health and Nutrition Biweekly Iraq, Jordan, and Lebanon July 7–20, 2013 Issue no. 7. Retrieved from https://reliefweb.int/report/jordan/inter-agency-regional-response-syrian-refugees-health-and-nutrition-biweekly-iraq-0

Jaspars, S., & Shoham, J. (1999). Targeting the vulnerable: A review of the necessity and feasibility of targeting vulnerable households. *Disasters, 23*(4), 359–372.

Lee, I., & Haines, A. (1991). Health costs of the Gulf War. *British Medical Journal,* 303–306.

Loewenberg, S. (2017, January 28). World report: Famine fears in northeast Nigeria as Boko Haram fight rages. *The Lancet, 389*. Retrieved from http://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736(17)30198-8.pdf

Macrae, J., & Zwi, A. (Eds.). (1994). *War and hunger: Rethinking international responses to complex emergencies.* London, UK: Zed Books.

Mahalanabis, D., Choudhuri, A. B., Bagchi, N. G., Bhattacharya, A. K., & Simpson, T. W. (1973). Oral fluid therapy of cholera among Bangladesh refugees. *Johns Hopkins Medical Journal, 132*(4), 197–205.

Manoncourt, S., Doppler, B., Enten, F., Nur, A. E., Mohamed, A. O., Vial, P., & Moren, A. (1992). Public health consequences of civil war in Somalia, April 1992. *The Lancet, 340,* 176–177.

Marfin, A. A., Moore, J., Collins, C., Biellik, R., Kattel, U., Toole, M. J., & Moore, P. S. (1994). Infectious disease surveillance during emergency relief to Bhutanese refugees in Nepal. *Journal of the American Medical Association, 272,* 377–381.

Mears, C., & Young, H. (1998). *Acceptability and use of cereal-based foods in refugee camps: case studies from Nepal, Ethiopia, and Tanzania*. Oxford, UK: Oxfam.

Médecins sans Frontières (MSF). (1997). *Refugee Health: An approach to emergency situations.* London, UK. Macmillan.

Médecins Sans Frontières (MSF) (2017). *Clinical guidelines - Diagnosis and treatment manual.*: Author. Retrieved from https://medicalguidelines.msf.org/viewport/CG/english/clinical-guidelines-16686604.html

Médecins sans Frontières (MSF). (2001). Activity report 2000–2001. Burundi: Fever, hunger, and war. Retrieved from http://www.msf.org/en/article/burundi-fever-hunger-and-war

Moore, P. S., Marfin, A. A., Quenemoen, L. E., Gessner, B. D., Ayub, X. X., Miller, D. S., . . . Toole, M. J. (1993). Mortality rates in displaced and resident populations of Central Somalia during the famine of 1992. *The Lancet, 341,* 935–938.

Nieburg, P., Waldman, R. J., Leavell, R., Sommer, A., & DeMaeyer, E. M. (1988). Vitamin A supplementation for refugees and famine victims. *Bulletin of the World Health Organization, 66,* 689–697.

Palmer, C. A., & Zwi, A. B. (1998). Women, health and humanitarian aid in conflict. *Disasters, 22*(3), 236–249.

Perone, S. A., Martinez, E., du Mortier, S., Rossi, R., Pahud, M., Urbaniak, V., & Beran, D. (2017). Non-communicable diseases in humanitarian settings: Ten essential questions. *Conflict and Health, 11,* 17. doi:10.1186/s13031-017-0119-8

Perrin, P. (1996). *War and public health.* Geneva, Switzerland: International Committee of the Red Cross.

Peterson, E. A., Roberts, L., Toole, M. J., & Peterson, D. E. (1998). Soap use effect on diarrhea: Nyamithutu refugee camp. *International Journal of Epidemiology, 27,* 520–524.

Physicians for Human Rights (PHR). (2015). Doctors in the crosshairs: Four years of attacks on health care in Syria. Retrieved from https://s3.amazonaws.com/PHR_Reports/doctors-in-the-crosshairs.pdf

Physicians for Human Rights (PHR), & Syrian American Medical Society. (2016). Madaya: Portrait of a Syrian town under siege. Retrieved from https://s3.amazonaws.com/PHR_Reports/madaya-portrait-of-a-syrian-town-under-siege.pdf

Project Ploughshares. (2002). Home page. Retrieved from www.ploughshares.ca/

Reed, H. E., & Keeley, C. B. (Eds.). (2001). *Forced migration and mortality.* Washington, DC: National Academies Press.

ReliefWeb (2009). The tsunami legacy—Innovation, breakthroughs and change. Retrieved from https://reliefweb.int/report/indonesia/tsunami-legacy-innovation-breakthroughs-and-change

ReliefWeb. (2012). Major food crisis in Central African Republic with malnutrition rates far above emergency levels. Retrieved from https://reliefweb.int/report/central-african-republic/major-food-crisis-central-african-republic-car-malnutrition-rates

ReliefWeb. (2017a). Hospital bombings: Building evidence with images. Retrieved from https://reliefweb.int/report/syrian-arab-republic/hospital-bombings-building-evidence-images

ReliefWeb. (2017b). Sudan/South Sudan: Measles outbreak—Mar 2015. Retrieved from https://reliefweb.int/disaster/ep-2015-000038-sdn

Roberts, L., Chartier, Y., Chartier, O., Malenga, G., Toole, M. J., & Rodka, H. (2001). Keeping Clean Water Clean in a Malawi Refugee Camp: a Randomized Intervention Trial. *Bulletin of the World Health Organisation,79,* 280–287.

Salama, P., Spiegel, P., & Brennan, R. (2001). No less vulnerable: The internally displaced in humanitarian emergencies. *The Lancet, 357*, 1430–1432.

Sphere Project. (2011). *Humanitarian charter and minimum standards in disaster response* (3rd ed.). Geneva, Switzerland: Author. Retrieved from http://www.sphereproject.org/

Spiegel, P. B. (2004). HIV/AIDS among conflict-affected and displaced populations: Dispelling myths and taking action. *Disasters, 28*(3), 322–339.

Spiegel, P. B., Bennedsen, A. R., Claass, J., Bruns, L., Patterson, N., Yiweza, D., & Schilperoord, M. (2007). Prevalence of HIV infection in conflict-affected and displaced people in seven sub-Saharan African countries: A systematic review. *The Lancet, 369*, 2187–2195.

Spiegel, P. B., Checchi, F., Colombo, S., & Paik, E. (2010). Health-care needs of people affected by conflict: Future trends and changing frameworks. *The Lancet, 375*, 341–345.

Spirer, H. F., & Spirer, L. (1993). *Data analysis for monitoring human rights.* Washington, DC: American Association for the Advancement of Science.

Stark, L., & Ager, A. (2011). A systematic review of prevalence studies of gender-based violence in complex emergencies. *Trauma, Violence, & Abuse, 12*(3), 127–134.

Stark, L., Asghar, K., Yu, G., Bora, C., Baysa, A. A., & Falb, K. L. (2017). Prevalence and associated risk factors of violence

against conflict-affected female adolescents: A multi-country, cross-sectional study. *Journal of Global Health, 7*(1), 010416. doi:10.7189/jogh.07.010416

Syrian Centre for Policy Research, American University. (2016). Forced dispersion: A demographic report on human status in Syria. Beirut, Lebanon: Author.

Toole, M. J., & Bhatia, R. (1992). A case study of Somali refugees in Hartisheik A camp, eastern Ethiopia: Health and nutrition profile, July 1988–June 1989. *Journal of Refugee Studies, 5,* 313–326.

Toole, M. J., & Waldman, R. J. (1990). Prevention of excess mortality in refugee and displaced populations in developing countries. *Journal of the American Medical Association, 263,* 3296–3302.

Toole, M. J., & Waldman, R. J. (1993). Refugees and displaced persons: War, hunger, and public health. *Journal of the American Medical Association, 270,* 600–605.

UNICEF. (2017). UNICEF Yemen humanitarian situation report (January 2017). Retrieved from http://reliefweb.int/report/yemen/unicef-yemen-humanitarian-situation-report-january-2017-enar

United Nations Administrative Committee on Coordination, Sub-Committee on Nutrition (ACC/SCN). (1994a). *Refugee nutrition information system* (Vol. 6). Geneva, Switzerland: Author.

United Nations Administrative Committee on Coordination, Sub-Committee on Nutrition (ACC/SCN). (1994b). *Refugee nutrition information system* (Vol. 7). Geneva, Switzerland: Author.

United Nations Administrative Committee on Coordination, Sub-Committee on Nutrition (ACC/SCN). (1995, July). *Refugee nutrition information system* (Vol. 11). Geneva, Switzerland: Author.

United Nations Administrative Committee on Coordination, Sub-Committee on Nutrition (ACC/SCN). (1996, April). *Refugee nutrition information system* (Vol. 15). Geneva, Switzerland: Author.

United Nations Administrative Committee on Coordination, Sub-Committee on Nutrition (ACC/SCN). (1998, June 15). *Refugee nutrition information system* (Vol. 24). Geneva, Switzerland: Author.

United Nations Administrative Committee on Coordination, Sub-Committee on Nutrition (ACC/SCN). (1999, March). *Refugee nutrition information system* (Vol. 26). Geneva, Switzerland: Author.

United Nations Administrative Committee on Coordination, Sub-Committee on Nutrition (ACC/SCN). (2002a, April). *Refugee nutrition information system* (No. 36/37). Geneva, Switzerland: Author.

United Nations Administrative Committee on Coordination, Sub-Committee on Nutrition (ACC/SCN). (2002b, October). *Refugee nutrition information system* (No. 39). Geneva, Switzerland: Author.

United Nations Administrative Committee on Coordination, Sub-Committee on Nutrition (ACC/SCN). (2003, November). *Refugee nutrition information system* (No. 43). Geneva, Switzerland: Author.

United Nations High Commissioner for Refugees (UNHCR). (1995a). *Reproductive health in refugee situations: An inter-agency field manual.* Geneva, Switzerland: Author.

United Nations High Commissioner for Refugees (UNHCR). (1995b). *Sexual violence against refugees: Guidelines on prevention and response.* Geneva, Switzerland: Author.

United Nations High Commissioner for Refugees (UNHCR). (2017a). *Global trends: Forced displacement in 2016.* Geneva, Switzerland: Author.

United Nations High Commissioner for Refugees (UNHCR). (2017b). Syrian regional refugee response update. Retrieved from http://data.unhcr.org/syrianrefugees/regional.php

United Nations High Commissioner for Refugees (UNHCR), & World Food Programme. (2011, January). *Guidelines for selective feeding: The management of malnutrition in emergencies.* Geneva, Switzerland: Author.

United Nations Standing Committee on Nutrition. (2004a). *Nutrition information in crisis situations. Report Number I.* Geneva, Switzerland: Author.

United Nations Standing Committee on Nutrition. (2004b). *Nutrition information in crisis situations. Report Number II.* Geneva, Switzerland: Author.

United Nations Standing Committee on Nutrition. (2004d). *Nutrition information in crisis situations. Report Number IV.* Geneva, Switzerland: Author.

United Nations Standing Committee on Nutrition. (2009a). *Nutrition information in crisis situations. Report Number XXI.* Geneva, Switzerland: Author.

United Nations Standing Committee on Nutrition. (2009b). *Nutrition information in crisis situations. Report Number XXII.* Geneva, Switzerland: Author.

United Nations Standing Committee on Nutrition. (2010). *Nutrition information in crisis situations. Report Number XXIII.* Geneva, Switzerland: Author.

United States Agency for International Development. (2017, March 3). South Sudan crisis fact sheet. Retrieved from https://www.usaid.gov/sites/default/files/documents/1866/south_sudan_crisis_fs05_03-03-2017.pdf

United States Committee for Refugees. (2005). USCRI to assist Burmese refugee tsunami victims in Southern Thailand. Retrieved from https://reliefweb.int/report/myanmar/uscri-assist-burmese-refugee-tsunami-victims-southern-thailand/

Uppsala Conflict Data Program. (2016). Retrieved from http://ucdp.uu.se/

Vu, A., Adam, A., Wirtz, A., Pham, K., Rubenstein, L., Glass, N., & Singh, S. (2014). The prevalence of sexual violence among female refugees in complex humanitarian emergencies: A systematic review and meta-analysis. *PLoS Currents, 18,* 6.

Walt, G. (1988). CHWs: Are national programs in crisis? *Health Policy and Planning, 3,* 1–21.

Working Group for Mortality Estimation in Emergencies. (2007, December). Wanted: Studies on mortality estimation methods for humanitarian emergencies, suggestions for future research. *Emerging Themes in Epidemiology, 4,* 9.

World Health Organization (WHO). (1999a). *Scurvy and its prevention and control in major emergencies.* WHO/NHD/99.11. Geneva, Switzerland: Author.

World Health Organization (WHO). (1999b). *Thiamine deficiency and its prevention and control in major emergencies.* WHO/NHD/99.13. Geneva, Switzerland: Author.

World Health Organization (WHO). (1999c). *Management of severe malnutrition: A manual for physicians and other senior health workers.* Geneva, Switzerland: Author.

World Health Organization (WHO). (1999d). Health Situation Report West Timor 15 Dec 1999. Source: ReliefWeb: https://reliefweb.int/report/indonesia/who-health-situation-report-west-timor-15-dec-1999

World Health Organization (WHO). (2000a). *Pellagra and its prevention and control in major emergencies.* WHO/NHD/00.10. Geneva, Switzerland: Author.

World Health Organization (WHO). (2000b). *The management of nutrition in major emergencies.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2003). *Communicable disease control in emergencies: A field manual* (M. A. Connolly, Ed.). Geneva, Switzerland: Author.

World Health Organization (WHO). (2009). *WHO child growth standards and the identification of severe acute malnutrition in infants and children. A joint statement by the World Health Organization and the United Nations Children's Fund.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2017a). South Sudan: Integrated Disease Surveillance and Response (IDSR): Epidemiological update W31 2017 (Jul 31–Aug 06). Retrieved from http://www.who.int/hac/crises/ssd/south-sudan-cholera-report-6august2017.pdf?ua=1

World Health Organization (WHO). (2017b). WHO scales up response in Somalia as drought-affected populations face difficult situation. Retrieved from http://www.emro.who.int/media/news/who-scales-up-response-in-somalia-as-drought-affected-populations-face-difficult-situation.html

World Health Organization (WHO). (2017c). Cholera outbreak response. Retrieved from http://www.emro.who.int/yem/yemeninfocus/situation-reports.html

World Health Organization (WHO). (2017d). Situation Report #119 on cholera in South Sudan 5 May 2017. Retrieved from http://www.who.int/hac/crises/ssd/sitreps/south-sudan-cholera-update-5may2017.pdf?ua=1

World Health Organization (WHO). (2017e). *Vaccination in acute humanitarian emergencies: A framework for decision-making.* WHO/IVB/17.03. Geneva, Switzerland: Author.

World Health Organization (WHO). (2017f). Cholera vaccines: WHO position paper. Retrieved from http://apps.who.int/iris/bitstream/10665/258764/1/WER9234-477-498.pdf

World Health Organization (WHO). (2017g). Antimalarial drug resistance in the Greater Mekong Subregion: How concerned should we be? Q&A with Dr Pedro Alonso, Director of the Global Malaria Programme. Retrieved from http://www.who.int/malaria/media/drug-resistance-greater-mekong-qa/en/

World Health Organization (WHO), & United Nations High Commissioner for Refugees (UNHCR). (2007). *Tuberculosis care and control in refugee and displaced populations: An interagency field manual* (2nd ed.) (M. A. Connolly, M. Gayer, & S. Ottmani, Eds.). Geneva, Switzerland: Author.

Yip, R., & Sharp, T. W. (1993). Acute malnutrition and high childhood mortality related to diarrhea. *Journal of the American Medical Association. 270,* 587–590.

Zwi, A., Ugalde, A., & Richards, P. (1999). The effects of war and political violence on health services. In L. Kurtz (Ed.), *Encyclopaedia of violence, peace and conflict* (pp. 679–690). San Diego: Academic Press.

CHAPTER 13

The Design of Health Systems

Anne J. Mills, Melisa Martinez-Álvarez, and M. Kent Ranson

Health systems are the means whereby many of the programs and interventions discussed in other chapters in this text are planned and delivered. They are a crucial influence on the extent to which countries are able to address their disease burden and improve overall levels of health and the health of particular groups in the population.

A health system comprises all organizations, institutions, and resources that produce actions whose primary purpose is to improve health (World Health Organization [WHO], 2000). The healthcare system consists of those organizations, institutions, and resources that deliver health care to individuals. Health systems are examples of "complex" systems, wherein the overall system contains numerous subsystems, each with its own characteristics, meaning that the outcome of various interacting subsystems is unpredictable (Adam & De Savigny, 2012).

Health systems vary greatly from country to country. Unlike in the study of disease, there is only limited standardized terminology or methodology for studying and understanding health systems. Each country's health system is the product of a complex range of factors, especially its historical patterns of development and the power of different interest groups (Balabanova et al., 2013). Nonetheless, it is possible to identify common features, and knowledge is increasing regarding which design features are associated with which outcomes, thereby facilitating cross-country learning.

It is extremely important to study and understand how health systems function and how they can be changed, given their importance in global and national economies. Health tends to consume a growing share of a country's gross domestic product (GDP) as income increases. On a worldwide basis, total expenditures on health care grew from 3% of world GDP in 1948 to 8.7% in 2006 (WHO, 2009). More recent trends for the period 2010 to 2015 show the same indicator increasing for high-income countries from 6.5% to 7.83%, and for low- and middle-income countries (LMICs) from 5.4% to 6.4% (WHO, 2017). Nevertheless, countries at similar income levels differ greatly in how effectively they look after the health of their populations. The health-related differences between countries of similar income can be enormous. For example, Bangladesh, with a 2016 GDP per capita of $1,030, has an under-5 mortality rate of 34 deaths per 1,000 live births, whereas Cote d'Ivoire has a 2016 GDP per capita of $1,553 and under-5 mortality rate of 92 deaths per 1,000 live births (World Bank Data, 2018a, 2018b). By comparison, in two somewhat richer countries, Thailand, with a GDP per capita of $5,902, has an under-5 mortality rate of 12 deaths, whereas South Africa, with only a slightly lower GDP per capita of $7,489, has a mortality rate of 43 deaths. It is challenging to disentangle the contributions of the health system to health outcomes relative to other determinants such as income, education, and diet (see the *Understanding and Acting on Social Determinants of Health and Health Equity* chapter), but it is clear that the health system is an important determinant of health.

The global attention given to achieving the health-related Millennium Development Goals (MDGs) by

2015 focused attention on the state of the health systems of many LMICs. Disease- and program-specific initiatives such as the GAVI Alliance (supporting immunization) and the Global Fund to Fight AIDs, TB and Malaria realized that the achievement and maintenance of their disease-focused goals was not possible in countries with weak and fragile health systems. Research pointed to the multiple problems affecting the health systems of low-income countries and many middle-income countries, and analyzed the multiple constraints they face in scaling up needed interventions to high levels of coverage (Mills & Hanson, 2003). There has recently been increasing recognition of the need for health systems resilience, so that such countries can prepare for and respond to crises effectively (Gilson et al., 2017; Kruk, Myers, Varpilah, & Dahn, 2015).

In the current age of the Sustainable Development Goals (SDGs), health systems remain a key theme, but are now strongly linked to the target of universal health coverage (UHC). Based on a review of evidence, McIntyre, Meheus, and Rottingen (2017) estimated that countries progressing toward UHC should aim at a target of government spending on health that amounts to at least 5% of GDP, and a per capita target of $86, to promote universal access to primary care in low-income countries.

Understanding health systems and how they can be changed is an endeavor that can benefit from the insights of a number of disciplines—most notably, economics, sociology, anthropology, history, political science, and management science. In the recent past, and not least because concerns regarding resource scarcity, cost inflation, and efficiency have been uppermost in policy makers' minds, the discipline of economics has dominated the study of health systems. However, with the emergence of health policy and systems research as an important area of inquiry (Gilson & WHO, 2013), other disciplines have been making growing contributions to effort, especially political science (contributing to understanding of the political and policy process) and the behavioral sciences (insights into the behavior of people and organizations). The importance of embracing other perspectives on health systems is also apparent in the report of the health systems knowledge network of the Commission on Social Determinants of Health (Gilson, Doherty, Loewenson, & Francis, 2007). This chapter draws strongly on economics to review key features of the design of health systems, but brings in learnings from other disciplines where they add critical insights.

The section that follows provides an overview of the health system and its key elements, and presents criteria for evaluating health systems. A brief review of the historical development of health systems is followed by a section addressing the fundamental and controversial question of the role of the state. Subsequent sections then consider each of the key functions of health systems in turn and recent reform trends: regulation, financing, resource allocation, and provision. Current topical issues in health systems are then reviewed. Throughout, the discussion is illustrated by country examples, with a core set of illustrations drawn from Tanzania, Sierra Leone, India, Mexico, and Thailand being chosen to illuminate key differences in health systems across the world.

It should be noted that the *Public Health Infrastructure* chapter focuses on the public health branch of a health system, including both its structure and its capabilities. This chapter addresses generic issues concerning the financing and organization of the system as a whole, as well as the more specific issues relevant to what is often called the healthcare system, referring to that part of the health system concerned largely with personal health services, which consume the great majority of health system funding.

Understanding the Health System

Since the seminal study of Kohn and White (1976), an expanding body of literature has been attempting to systematize the discussion of the various elements of health systems, categorize health systems into a limited number of types, develop performance indicators, and look at health systems less as economic entities and more as social institutions. These four issues are discussed in turn in this section.

Elements of Health Systems

Roemer (1991) identified five major categories that facilitate a comprehensive description of a country's health system:

- Production of resources (trained staff, commodities such as drugs, facilities, and knowledge)
- Organization of programs (by government ministries, private providers, and voluntary agencies)
- Economic support mechanisms (sources of funds, such as tax, insurance, and user fees)
- Management methods (planning, administration, regulation, and legislation)
- Delivery of services (preventive and curative personal health services; primary, secondary, and tertiary services; public health services; and services for specific population groups, such as children, or for specific conditions, such as mental illness)

More recently, WHO (2007) has developed a system for describing and categorizing health systems based on six "building blocks":

- Service delivery
- Health workforce
- Information
- Medical products, vaccines and technologies
- Financing
- Leadership and governance (stewardship)

These types of categorization are helpful for describing health systems; indeed, Roemer (1991, 1993) applied his approach to a large number of countries. However, such categorizations are less helpful for understanding how well health systems perform. Making this determination requires much more detailed subcategories and greater elaboration of the relationships, not just within each category but particularly between categories (e.g., between economic support mechanisms and organization of programs).

Typologies of Health Systems

To make comparisons based on how different types of health systems perform, it is necessary to group countries and their health systems into distinct types. There have been various attempts to do this. Countries can be classified according to the following criteria:

- The dominant method of financing (e.g., tax, social insurance, private insurance, out-of-pocket payments)
- The underlying political philosophy (e.g., capitalist, socialist)
- The nature of state intervention (e.g., to cover the whole population or only the poor)
- The level of gross national product (GNP) (e.g., low, middle, high)
- Historical or cultural attributes (e.g., industrialized, nonindustrialized, transitional)

A key difficulty, however, is that countries do not fit neatly into these categories. In particular, the health systems of LMICs tend to be fragmented, with different arrangements for different population groups (McPake & Machray, 1997). As an illustration of these differences, **EXHIBIT 13-1** summarizes the structure of the health systems in five countries.

The Organization for Economic Cooperation and Development (OECD) has developed a typology that is helpful for categorizing not only the economic

EXHIBIT 13-1 Illustrations of the Structure of Health Systems

Sierra Leone (GNI per capita $490; low-income country)

The public sector plays a leading role in Sierra Leone's health system, along with a large informal sector as well as high-tech Western-style private hospitals, faith-based and nongovernmental providers, independent private biomedical providers, drug sellers, and traditional healers (F. Martineau, personal communication, July 31, 2017). The primary health sector is managed at the district level by district health management teams, with decentralized funding provided through district or city councils. This organization was implemented as part of a broader push toward decentralization following the civil war (which ended in 2002) and has had mixed results. In particular, although health funding is supposed to be ring-fenced, in practice it can be extremely difficult for district medical officers or hospital superintendents to access funds. Secondary care is managed by hospital superintendents, who report directly to the country's Ministry of Health. There are four tertiary facilities focusing on different specialties, all located in the capital. The most important recent health system reform was the national rollout of free health care for children younger than 5 years, pregnant women, and lactating women. People living with disabilities, as well as Ebola survivors, are also legally entitled to free government care. To complement the free health initiative, performance-based financing was introduced for those sectors that deal most commonly with children and with pregnant and lactating women, including extra salary for staff members and discretionary funds to improve facilities.

Tanzania (GNI per capita $900; low-income country)

The public sector plays a leading role in Tanzania's health system, owning approximately 64% of all health facilities and covering all levels of care from primary to tertiary (Mtei et al., 2007). Relative to neighboring countries, the public system is quite decentralized, with public services managed at the district level. Mission (church) services are an important source of care outside the main towns, providing the same number of hospitals as the government, and are subsidized by the state. Private doctors practice in the main cities, and there is a large informal sector of traditional practitioners and drug sellers. The majority of health expenditures flow through the government; most of the remainder consists of out-of-pocket expenditures. The two largest health insurance arrangements are the National Health Insurance Fund (NHIF),

(continues)

EXHIBIT 13-1 Illustrations of the Structure of Health Systems *(continued)*

a mandatory scheme offering comprehensive benefits to the formal sector, and the Community Health Fund (CHF), now managed by the NHIF, which is a voluntary scheme for the informal sector in rural areas, offering limited benefits in lower-level public facilities. In 2016, total coverage of insurance schemes (including a small private insurance sector) was 25.8% of the total population (Health Policy Project, 2016).

India (GNI per capita $1,680; lower-middle-income country)

India's public sector is large in absolute terms, providing all levels of care. Health care is in general a state-level function, with central government involved mainly in overall policy and specific disease-control programs. There is a large formal private sector, providing both ambulatory and inpatient care, and an even larger informal sector consisting of unlicensed and unqualified practitioners and drug sellers. There is a limited formal interaction between public and private sectors. A compulsory state insurance system covers lower-paid, formal-sector workers, while another scheme covers senior government officers. An additional health insurance scheme was launched by the government of India in 2008, targeted at the 300 million people who live below the poverty line. As of 2017, this scheme covered some 41 million informal-sector workers and their families (RSBY, 2017). In addition to these measures, many states have introduced financial protection schemes for specific types of tertiary care such as cardiac surgery, cancer therapy, and neurosurgery.

Thailand (GNI per capita $5,640; upper-middle-income country)

Both the public and private health sectors are quite large in Thailand, providing all levels of care. There is widespread use of private-sector resources, especially for outpatient care. Compulsory social insurance covers those employees with formal private employment and finances care provided by public and private health facilities (chosen by the insured). Civil servants have their own medical benefit scheme, which pays for care mostly at public hospitals. Since 2001, the remainder of the population—who formerly could get a low-income card exempting them from fees if they were poor, purchase a voluntary health insurance card, or pay out of pocket for care—has been covered by a universal coverage arrangement in which individuals register at a local facility and can then access a wide range of free healthcare benefits (Towse, Mills, & Tangcharoensathien, 2004).

Mexico (GNI per capita $9,040; upper-middle-income country)

Both the public and private sectors play an important role in financing and provision of healthcare services in Mexico. Formal-sector employees are covered by various social insurance institutions. The poor receive care through government facilities or private providers (allopathic and traditional). There is little interaction between the public and private sectors, in the form of either regulation or contracting for service delivery. Concern about duplication and waste of resources within the three subsystems (i.e., social security, other government, and private) and lack of protection against the burden of healthcare costs, especially for the poor, have led to reforms based on decentralization, managed market principles, and a voluntary insurance scheme called *Seguro Popular*. Between *Seguro Popular* and its social security institutions, Mexico is reaching universal health coverage (i.e. 100% of the population covered by health insurance) (Knaul et al., 2012), although some estimate that more than 20% of the Mexican population remains unaffiliated with an insurance plan (CONEVAL, 2012).

Note: GNI (gross national income), calculated using the World Bank Atlas method, is in 2016 prices.

dimensions of health systems in OECD countries, but also the directions in which reforms are taking them (OECD, 1992). The key categories are as follows:

- Whether the prime funding source consists of payments that are made voluntarily (as in private insurance or payment of user fees) or are compulsory (as in taxation or social insurance)
- Whether services are provided by direct ownership (termed the integrated pattern, in which a ministry of health or social insurance agency provides services itself), by contractual arrangements (in which a ministry of health or social insurance agency contracts with providers to deliver services), or simply by private providers (paid by direct out-of-pocket payments)
- How services are paid for (prospectively, where financial risk is transferred to providers, or retrospectively, where the cost of care is reimbursed)
- The basic insurance package (population covered, services included, and degree of cost sharing)
- The degree of reliance on centralized "command-and-control" systems to steer demand and supply of health services, versus market mechanisms, which rely on private provision financed on a fee-for-service basis, competition among private providers driven by user choice, and private insurance
- The existence of a gatekeeper and user choice

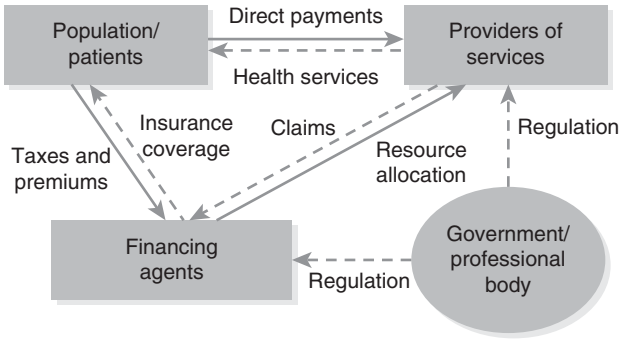


FIGURE 13-1 A map of the health system.

Modified from World Health Organization (WHO). (1993). *Evaluation of recent changes in the financing of health services*. Geneva, Switzerland: Author. Retrieved from http://apps.who.int/iris/bitstream/10665/41528/1/WHO_TRS_829.pdf. Copyright ©1993.

More recent OECD analysis has organized the various dimensions into six different types of health system (OECD, 2011). This categorization is most relevant to countries at a relatively advanced stage of development, where the coverage of the health system is universal or consists of a limited number of arrangements, as opposed to poorer countries, where the health system is fragmented.

Given the lack of an agreed typology for LMICs, the content of this chapter is based on a simple framework (shown in **FIGURE 13-1**) that identifies four key actors and four key functions required in any health system. The actors are as follows:

- The government and professional bodies that govern, structure, and regulate the system
- The population, including patients, who as individuals and households ultimately pay for the health system and receive services
- Financing agents, who collect funds and allocate them to providers or purchase services at national or lower levels
- The providers of health services, who themselves can be categorized in various ways, such as by level (primary, secondary, tertiary), service type (curative, preventive), ownership (public; private, for-profit; private, not-for-profit), degree of organization (formal, informal), or medical system (allopathic, ayurvedic)

The functions are as follows:

- Governance and regulation
- Financing (through taxes, premiums, and direct payments)
- Resource allocation
- Providing services

Evaluation of Health Systems

WHO (2000) has argued that the health system has three main objectives, which are intrinsically valuable:

- Good health (both its absolute level and its distribution across the population)
- Fairness in financial contribution
- Responsiveness to peoples' expectations (both level and distribution)

In assessing the actual performance of health systems, criteria of efficiency and equity are frequently applied, so it is important to understand their various meanings. Efficiency has a number of different dimensions:

- Macroeconomic efficiency refers to the total costs of the health system in relation to overall health status; countries differ in how efficiently their health systems convert resources used into health gains.
- Microeconomic efficiency refers to the scope for achieving greater efficiency from existing resources. It is of two types:
  - Allocative efficiency: Devoting resources to the mix of activities that will have the greatest impact on health (i.e., is most cost-effective).
  - Technical efficiency: Using only the minimum necessary resources to finance, purchase, and deliver a particular activity or set of activities (i.e., avoiding waste).

Equity refers to the fair distribution of the costs of health services and the benefits obtained from their use among different groups in the population. It is inherently a question of values, with views differing as to what constitutes fairness of financing or access to health care. However, indicators of who pays for health services and who receives benefits provide evidence that serves as the basis for judgments about the degree of equity achieved by particular health systems. Equity is commonly expressed in two different ways.

- *Horizontal equity* refers to the equal treatment of equals. With respect to financing and resource allocation, this type of equity implies that the charge levied by all agents or providers for a particular good or service should be the same for households with equal ability to pay (regardless of gender, marital status, and so on). Horizontal equity is therefore assessed by the extent to which contribution levels are similar between those with similar ability to pay. With respect to provision of services, horizontal equity means that individuals with the same health condition should have equal access to health services.
- *Vertical equity* is based on the principle that individuals who are unequal in society should be treated differently. Vertical equity in the financing and purchasing of health services means that consumers should be charged for the same good or service according to their ability to pay.

TABLE 13-1 demonstrates how equity and efficiency criteria can be used to guide the financing, allocation of resources, and provision of health services and to evaluate performance. In the literature, Mills et al. (2012) provide an example of financing and benefit incidence analysis to evaluate the equity of health systems in Ghana, South Africa, and Tanzania.

In WHO's (2000) Health Systems: Improving Performance study, a conceptual framework, based on the goals and functions listed earlier, was applied to country data to assess and understand country health system performance. Countries were ranked in relation to their attainment of the individual goals, in relation to overall goal attainment, and in terms of performance on level of health and the functioning of the overall health system. Although the conceptual framework and new databases have proved useful, the ranking was extremely controversial (e.g., see Almeida et al., 2001; Coyne & Hilsenrath, 2002) and has not been continued in subsequent World Health Reports.

Health Systems as Social Institutions

A number of health systems frameworks have been put forward since 1945 that build on the notion of health systems as social institutions, especially frameworks drawing on political theories. A review by van Olmen et al. (2012) found that the evolution of thinking about health systems over time is not a result of accumulation of knowledge, but rather a function of the policy context and dominant actors. In the 1950s, WHO and UNICEF were the dominant actors, and the scope of global health was increased from a medical focus to one of Health for All at the Alma Ata declaration. In the decades that followed, the World Bank

TABLE 13-1 Equity and Efficiency Criteria Applied to Financing, Resource Allocation, and Provision of Health Care

	Efficiency		Equity	
Functions	**Allocative**	**Technical**	**Horizontal**	**Vertical**
Financing		Maximize the proportion of resources raised that are actually available for purchasing health care (e.g., reduce the overhead costs of collecting taxes)	Equal payment by those with equal ability to pay (e.g., same insurance premium for same income group)	Payment in relation to ability to pay (e.g., progressive income tax rates)
Allocating resources	Purchase the mix of interventions that provides the greatest health gains	Maximize the proportion of resources spent by agents that are actually available for providing health care	Services purchased for similar groups (e.g., the elderly) should be the same in different geographic areas	Services purchased should reflect the different needs of different groups (e.g., the elderly versus children)
Providing services	Provide those interventions that return the greatest value for money (e.g., in a poor country, antenatal care should be provided before radiotherapy for cancer)	Make the best use of resources in providing interventions deemed worthwhile (e.g., have nurses as opposed to doctors provide most antenatal care)	Equal access for equal need (e.g., equal waiting time for treatment for patients with similar conditions)	Unequal treatment for unequal need (e.g., unequal treatment of those with trivial versus serious conditions)

Data from Donaldson, C., & Gerard, K. (1993). *Economics of health care financing: The visible hand.* London, UK: Macmillan Press.

became a more active player in health, and the scope of frameworks was narrowed to a special package of essential services that was deemed more realistic. Since 2000, health systems thinking has been shaped by a change in the number and type of global health actors (with the introduction of global health initiatives and private philanthropies), WHO's attention to performance of health systems, and the recognition of the complexity of health systems made up of dynamic relationships.

Gilson (2003) has called for health systems to be conceptualized beyond "delivery points for biomedical interventions"—that is, as complex geopolitical institutions, underlined by different sets of relationships, including relationships between patient and provider, between employer and employee, and between different providers. A key influence in these relationships is the concept of trust, whose formation and maintenance require adequate managerial and organizational practices and political processes, and a focus on equity and justice.

Building on the view of health systems as social institutions, a number of recent studies have conceptualized health systems as a demonstration of social solidarity. For instance, Kruk et al. (2010) view investment in health systems as a state-building activity in post-conflict settings; they suggest that such investment can promote social cohesion, restore accountability, strengthen the social contract, and strengthen government capacity. Social solidarity has also been argued to underpin European models of health system financing.

Backman et al. (2008) developed a set of indicators to assess the health systems of 194 countries from a human rights perspective. These indicators included recognition of the right to health, the availability of a national health plan, the degree of participation, discrimination, health information, access to health services, and distribution of human resources, among others. Building on this work, Gruskin et al. (2012) identified four priority areas for incorporating human rights in health systems frameworks: (1) including client interaction at all stages of health system design and implementation, (2) promoting equality and nondiscrimination, (3) systematically addressing laws and policies negatively affecting health and health service utilization, and (4) strengthening accountability.

A historical perspective on the development of health systems, provided in the next section, adds to the understanding of health systems as complex institutions that have developed over time in particular ways in response to local and external influences.

Historical Development of Health Systems

As indicated by archaeological evidence, medicine has had a role in all cultures and civilizations (e.g., in ancient Mesopotamia and Egypt). It has also been a concern of the state throughout the ages. For example, the law code of Hammurabi (1792–1750 BC) specified the fees for an operation to be paid to a healer on a sliding scale, depending on the status of the patient, and also specified penalties for failure. The Romans built hospitals for domestic slaves and soldiers in permanent forts in occupied territories such as England. Despite these efforts, the origins of the modern hospital lie primarily in the spread of Christianity and of ideas of Christian charity and caring for all who might be in need after the conversion of Constantine (died 337 AD) made Christianity an official imperial religion (Porter, 1996a). Hospitals were founded in the main cities of the Christian world, often associated with churches or monasteries. The Islamic world also developed hospitals, and by the 11th century there were large hospitals in every major Muslim town. Such hospitals were intended for sick individuals who lacked families or servants to care for them—for example, the poor, travelers, and those working away from home.

In Europe by the Middle Ages, a multiplicity of institutions and organizations had developed with pretensions to authority over medicine: the church, guilds, medical colleges, town councils, and powerful individuals. In Brussels, for example, a board of clergy, doctors, and midwives licensed midwives in the 15th century. The arrival of the plague—the Black Death—which in its first wave killed approximately 25% of Europe's population, stimulated growing state involvement in protection of health through measures such as imposing quarantines and isolating the sick.

From the early 19th century onward, the scientific basis of Western medicine was increasingly established, with scientific training becoming essential for the practice of medicine (Porter, 1996b). The 18th and 19th centuries saw a vast expansion of hospitals in Europe and the United States, supported both by philanthropy and by public funds, especially hospitals for persons with infectious diseases and the mentally ill (Porter, 1996a). The charitable and voluntary basis for the funding of hospitals in many countries faced a crisis, however, as medicine became more elaborate and expensive. The enormous increase in the number of surgical procedures performed and

the development of technology led to both much larger numbers of patients and much higher costs per patient. In the face of these trends, voluntary hospitals ran into financial difficulties. In the United States, hospitals developed business strategies based on insurance that could attract more affluent patients. In the United Kingdom, where insurance was much less well developed, hospitals were eventually brought into public ownership. In Scandinavia, local authorities had had responsibility for providing hospital services since the late 19th century, so these facilities developed largely as a public service.

The development of the discipline of public health in the 19th century was a response to the disease hazards of the urban environment. In England and Germany, public health measures focused on safe water supply and drains.

Over the 18th and 19th centuries, modern forms of medical regulation developed. Trends in countries in which medicine was dominated by free markets (such as the United States) converged with trends in countries in Europe with strong state control (such as Germany) to produce the closely regulated medical markets that exist today. However, the degree of state involvement in the provision of health services varied enormously between countries, as remains true today.

A key development was the increase in collective arrangements for funding health services. State services developed in all Western countries to provide health services for those persons who could not afford to purchase them themselves. In addition, mutual insurance schemes emerged in Europe and the United States as means to protect workers against financial losses, with these plans often including medical care. Such mutual insurance schemes were encouraged by the German states and ultimately developed into a national program of health insurance. In the United Kingdom, in contrast, they were nationalized as part of the expansion of state welfare. Other European countries also saw the development and expansion of compulsory financial arrangements for health services, whether through the extension of insurance arrangements for medical care (the Bismarck model, named after the German chancellor who introduced the first compulsory insurance scheme) or through general taxation (the Beveridge model, named after the British minister of health who is regarded as the founder of the U.K. National Health Service). Since World War II, all high-income countries—except the United States, where progress has been slower (Gaffney & McCormack, 2017)—have extended mechanisms for protection against the financial risks of ill health, to the point where they can be said to have achieved universal coverage.

Another key development in the creation of the modern health system was the development of organized systems of medical care, as opposed to fragmented and competing individual doctors and hospitals. World War I marked a turning point in Europe, when the need to organize medical care on a massive scale highlighted the advantages of a large, coordinated system. In the United Kingdom, the Dawson report in 1920 designed a system of district health services based on general practitioners and health centers, with referral upward to hospitals (WHO, 1999). The later development of the philosophy of national economic plans and of a strong government role in many sectors of the economy also supported the development of organized health systems. Even in those countries with less of a tradition of a strong state role in health services, cost escalation in recent decades has forced greater state involvement in this realm.

In the late 19th century Western medicine spread around the world, often as part of the process of colonial expansion (Zwi & Mills, 1995). Medicine acted in part as an agency of Western imperialism, with organized health services being a notable component of British, French, German, and Belgian colonization. These services were initially intended for the military, settler, and civil service communities, but it rapidly became apparent that protecting the health of expatriates required addressing health needs among the colonized peoples. In addition, health services were introduced by commercial interests if they believed that these services would improve their economic returns, and by the church as part of missionary activities. To a much greater extent than was the case in the home countries of the colonizers, the provision of health services in the colonies became associated with the state. This was accentuated in the postcolonial era by the prevalent ideologies of state-led growth and state responsibility for the welfare of all inhabitants. In the most extreme form of these ideologies, socialist countries such as Tanzania, China, and Vietnam banned private practice. More generally in Africa and Asia, the attention of policy makers focused on extending publicly financed and provided services to the whole population, while neglecting the often large private sector and the even larger traditional medicine sector.

Developments in Latin America were somewhat different. As in Africa, the earliest Western health services were developed by the colonists, especially for the armed forces and the police. Major employers provided health services, particularly where enterprises were remote from urban centers. Some religious

hospitals were built to care for the poor. These hospitals were later supplemented by government hospitals and clinics, especially in areas lacking charitable hospitals. A key difference between Latin America and most of Africa and Asia was the early development of compulsory insurance arrangements for workers in the formal sector. Because medical care infrastructure in this region was lacking, the insurance agencies often built and ran their own services, thereby contributing to the evolution of the parallel health systems still seen today in many Latin American countries.

The historical development of health services in many countries resulted in a health infrastructure that was biased toward hospitals. Attempts to reorient services culminated in 1978 in the Declaration of Alma Ata, which emphasized the importance of primary health care, involving the delivery of curative and preventive services at the community level. This perspective encouraged strong emphasis on the building up of integrated health services, involving community-based health workers. Nonetheless, a rival approach argued in favor of selective primary health care, to include those interventions that addressed the greatest disease burden and were most cost-effective (Walsh & Warren, 1979). Services for children were a key priority in this approach, and, together with the emphasis on family planning that resulted from the preoccupation of many donors with world population growth, meant that peripheral health services in many low-income countries were targeted primarily at women and children. Only recently has there been greater emphasis on achieving a more integrated approach to the delivery of health services at peripheral levels, and on addressing the health needs of adolescents and adults. The need to develop services to support human immunodeficiency virus (HIV) treatment, and then to deal with the growing incidence of noncommunicable diseases, has further highlighted the importance of continuity of care (Beaglehole et al., 2008).

A marked development in the 1980s and 1990s was greater questioning of the government's role in health systems. The most radical changes occurred in countries formerly under strict communist rule, where a market economy was introduced and market forces were allowed to influence health services (see Saltman & Figueras, 1997, for a discussion of such developments in eastern and central Europe). Social insurance arrangements were introduced for those in formal employment, health professionals permitted to have private practices, the development of private markets in pharmaceuticals was encouraged, and much greater costs fell directly on household budgets. In China, for example, approximately 71% of the population—including 48% of the rural population—had some insurance protection in 1981; by 1993, the overall insurance coverage in that country had dropped to 21%, with only 7% coverage of the rural population (WHO, 1999). Although changes were less dramatic in other parts of the world, governments were forced by economic crises in the 1980s and 1990s to consider how they could best prioritize what they do and ration services to those most in need; many introduced revenue generation schemes, such as user fees. This reconsideration was also forced by rapidly growing private markets in medical care.

Most recently, in the 21st century, there has been a marked increase in development assistance for health, in part associated with the movement surrounding the MDGs but also driven by the international response to the HIV pandemic and the entry of new sources of funding such as the Bill and Melinda Gates Foundation and new funding mechanisms such as the Global Fund for HIV, TB and Malaria. Funding streams became increasingly stratified by disease or health programs, in marked contrast to the primary healthcare ethos of earlier decades. By 2009, this shift in perspective was producing a strong reaction, with analyses highlighting how the success of disease-specific programs was hampered by overall weak health systems (Travis et al., 2004), and calling for much greater investment in health system strengthening. This trend culminated with the inclusion of universal health coverage as a target in the SDGs.

The Role of the State

The preceding historical review indicates that one of the key issues in the design of health systems is the role assigned to the state. This section examines the economic arguments commonly put forward to specify the state's role in health and then considers other explanations for the roles observed in practice. Health economics textbooks set out the arguments in more detail; see, for example, Chapter 7 in McPake, Normand, and Smith (2013), and Chapter 3 in Donaldson and Gerard (2005).

The first main economic justification lies in explanations of market failure. The efficient outcomes of private markets depend on a number of conditions being met. Because of the particular characteristics of health and health services, this may not be the case in the health sector.

First, the presence of externalities means that the optimal amount of health services may not be produced or consumed. Externalities are costs or benefits that are not taken into account in the transactions of producers or consumers. For example, an individual's decision on whether to be immunized will be related to the value of the protection to that individual, rather than the value of such protection that may be accorded to others and that reduces the pool of susceptible individuals.

Second, for goods that are public goods, the market may fail to produce them at all. Such goods are those where consumption is nonrival (i.e., consumption by one person does not reduce the consumption of another) and nonexcludable (i.e., a consumer cannot be prevented from benefiting from the good—for example, through requiring payment). Control of mosquito breeding sites to reduce malaria transmission is an obvious example: All people living in the area will benefit from such interventions regardless of whether they have paid for them. Information can also be seen to be a public good because it is nonrival; it is not nonexcludable, but the cost of providing information to extra people is often low. This applies to knowledge gained through research, for example.

Third, monopoly power can lead to market failure, because it enables the provider to charge more than if the market were competitive. Monopoly power may be held by a hospital in a particular geographic area, by a pharmaceutical firm, or even by a profession as a whole (such as the medical profession).

These arguments provide a rather weak justification for state intervention in the entire health system, since the range of services they apply to is quite limited. The arguments are most relevant to public health services and preventive care, and less relevant to the bulk of curative services. Moreover, problems such as monopoly are not unique to health and are commonly dealt with by regulation rather than by state provision of services. A more powerful argument for a large state role in health systems relates to the asymmetry of information between provider and consumer. Medical consultations are often sought precisely because patients do not know what is wrong with them: They are therefore ill informed, in contrast to the normal assumption in economics of perfectly informed consumers. Hence, in medical care, providers are in an unusually strong position. Although they may act as perfect agents for the consumer, it is also possible, especially when the consumer's income level is related to the care provided, that the personal interests of the providers may enter into decisions made on treatment. The poor and less educated are particularly vulnerable to unscrupulous profit seeking by private providers.

Another characteristic of health care is its uncertain nature, and the potential for high costs. This makes health care an obvious candidate for insurance, but it is generally accepted that private insurance markets do not work well in this realm (McPake et al., 2013). Individuals who purchase insurance may indulge in activities that put their health more at risk than if they were not insured, or once ill may consume more health care. This phenomenon, which is known as moral hazard, tends to raise the cost of insurance, making it unaffordable for some. Another problem is that those persons who are at greatest risk of needing care will be more likely to seek insurance, but due to asymmetries of information between insurer and insured, it is often difficult for the insurer to tailor the premium charged to the nature of the risk. This process, known as adverse selection, means that the insurer ends up with a more costly risk pool, premiums rise, and the healthier individuals opt out. In addition, insurance becomes more expensive, so that people who cannot afford the increased premiums are excluded from the market. Although the result is clearly inequitable, it is also inefficient because there will be people unprotected who would be willing and able to purchase insurance if the market worked well.

Other arguments in favor of state involvement are distinct from the arguments related to market failure. For example, one might argue that some types of health services are merit goods—that is, goods that society believes should be provided, but that individuals, if left to themselves, might underconsume because they are not the best judge of what is in their own or the public's interest. This argument is strongest for health services for children and the mentally ill.

Another argument is founded on equity principles: Namely, even with perfectly operating private markets for health services and health insurance, there will inevitably be individuals who are too poor to afford to access them. Although it could be argued that this problem could be addressed by income redistribution policies, equitable access to health services remains a concern; hence, it can be argued that providing benefits in kind is appropriate.

Although these are the standard arguments used to explore the appropriate role of the state in health services, the judgment of their significance differs greatly among economists, leading to radically different policy prescriptions. Even though much of this debate has focused on the relative merits of the U.S.

health system versus the Canadian or British health systems, it has also influenced the nature of the debate concerning the reform of health systems in LMICs, as noted later. Underlying this debate are alternative views on the ethical basis of a health system. One view sees access to health services as similar to access to other goods and services, and dependent on an individual's success in gaining or inheriting income. The other perspective sees access to health services as a right of citizenship that should not depend on individual income or wealth. According to the first view, the state's role in health should be confined to regulation of the market, public health measures, and public welfare for the poor to provide a minimum acceptable level of services, but nothing like the level of services available to those who are financially better off. According to the second view, the role of the state should be to ensure equal access to health services, meaning that access does not differ depending on economic or social status.

Although economic arguments provide justification for state involvement, they provide little guidance on the precise nature of intervention. In particular, they do not necessarily imply that the state should itself provide health services (as opposed to purchasing services from others). A key change in recent decades in thinking about public management has been the recognition that the state need not provide services itself directly, but instead could play an enabling role (Walsh, 1995).

An important influence on this position is a recognition that in many countries the state has failed in its policies to provide good-quality public services, including health services, for everyone. These arguments derive from a number of strands of economic thinking, especially public choice theory and property rights theories. The former is concerned with the nature of decision making in government. That is, public choice theorists argue that government officials are no different from anyone else in pursuing their own interests. Thus, politicians will be concerned with maximizing their chances of being reelected, and bureaucrats will serve their own interests (e.g., maximizing their budgets because their own rewards [salary, status] are related to that). The result is that the public sector is wasteful because politicians and bureaucrats have no incentive to promote allocative or technical efficiency.

In contrast, property rights theorists argue that the source of inefficiency in the public sector is the weakening of property rights. In the private sector, it can be argued, entrepreneurs or shareholders have a strong interest in the efficient use of resources. In contrast, in the public sector, there is little obvious threat to an enterprise if staff perform poorly; hence, incentives for efficient performance are weak.

These theories underlie what has been termed the "new public management," which seeks to expose public services to market pressures, without necessarily privatizing them (Walsh, 1995). Such approaches change the nature of state involvement, with policies of opening up services to competitive bids or putting services out to contract on a competitive basis; introducing internal markets where public providers have to compete for contracts from public purchasers; devolving financial control to organizations such as individual hospitals; spinning off parts of government into separate public agencies, such as an agency to manage government health services; and increasing the choice and influence of consumers by giving them resources in cash or kind, such as vouchers for treating sexually transmitted disease). This last approach, termed demand-side financing, enables consumers to purchase healthcare services from their chosen (usually accredited) provider.

Although theories justifying particular roles of the state feature prominently in writings on health systems, it is clear in practice that the actual role of the state in any particular country is shaped by a wide variety of influences. Most notable are the history of state involvement in health services and the rationale for its involvement over time, the extent to which private providers and insurers developed early in the history of the health system and thus were able to play a prominent role, and the attitude of the medical profession toward an increased state role (Mills et al., 2001). One key issue has been the extent to which the state took on itself the responsibility for providing services to the whole population, or instead concerned itself only with the poor and indigent.

EXHIBIT 13-2 summarizes the role of the state in five countries' health systems. In these countries, the government's financial contribution to health can be assessed against the commonly suggested benchmark of contributing 5% of GDP (McIntyre et al., 2017), and for Africa, the pledge of the African Union member states to set a target of allocating at least 15% of their annual budgets to improve the healthcare sector (WHO, 2011). In Tanzania, India, and Sierra Leone, the state's aims have historically been to provide services free at the point of use to the whole population. In contrast, both Mexico and Thailand have established specific schemes that cater to the poor and indigent, although in 2001 Thailand extended financial protection to the whole population through its "universal coverage" policy (Towse et al., 2004).

EXHIBIT 13-2 Illustrations of the Role of the State

Sierra Leone

The public health sector consumes 1.9% of GDP in Sierra Leone, and general government expenditures on health amount to 10.8% of total government expenditures and 7% of total health expenditures. The private sector in Sierra Leone is underdeveloped, with private facilities operating on a fee-for-service basis. Data on the private sector are scant.

Tanzania

Health services for the whole population have traditionally been seen as the responsibility of the state in Tanzania. Government expenditures on health accounts for 2.6% of GDP (and general government expenditures on health are 12.3% of total government expenditures), and 47% of total health expenditures are contributed by the government. More than 95% of hospital beds are owned either by government or by churches, while fewer than 5% are owned by the private sector (United Republic of Tanzania, Ministry of Health and Social Welfare, 2015). Independence brought in a government that was anti-private sector; even so, for at least the last 10 years, private-sector engagement has been an integral part of the country's health policy. Public-sector reforms have decentralized public management and introduced fees for public services. The poor are reliant mainly on public and church services and the informal private sector.

India

The historical emphasis in India was on a strong state role in health services, but resources were never provided to make this a reality. Public expenditures account for 1.4% of GDP in India, and general government expenditures on health amount to 5.0% of total government expenditures and 30% of total health expenditures. Approximately 65% of total hospital beds are public. Public services are often of poor quality and are adversely affected by the private practices of government doctors. People in general distrust the state. The policies related to charging for public facilities vary by state, but income from fees is small. There is widespread use of the private healthcare services sector by all sections of the population, including the poor.

Thailand

The public health sector amounts to 3.2% of GDP in Thailand, and general government expenditures on health represent 13.3% of total government expenditures and 77.8% of total health expenditures. In 2015, 79% of total hospital beds were found in public-sector facilities (Asia Pacific Observatory on Health Systems and Policies, 2015). Thailand has had a tradition of strong central government and laissez-faire economic policies. Historically, government policies have encouraged the private sector to grow through tax exemptions and public funding for private care for specific groups. With the introduction of the recent universal coverage policy, the government strengthened its commitment to inclusive social policies. Public services are of generally good quality; their main problem is considered to be lack of a consumer orientation.

Mexico

In Mexico, the public health sector makes up 3.3% of GDP, and general government expenditures on health account for 11.6% of total government expenditures and 51.8% of total health expenditures. Approximately 66% of hospital beds are found in public facilities (OECD, 2016). The government's primary role is that of owner of social security institutions, which account for some 55% of general government expenditures on health. The federal and state health ministries provide health care to those not covered by social security institutions, with state health ministries operating *Seguro Popular*, a program that provides free access to almost 300 healthcare interventions. There is dissatisfaction with the quality of public services in general, with those who can afford it preferring to use private services.

Note: Where not otherwise cited, expenditure data are from http://www.who.int/nha/country/en/.

Governance and Regulation

The governance function for health systems encompasses a set of rules guiding the roles and responsibilities of political, economic, and institutional processes with the aim of achieving health-sector objectives (Brinkerhoff & Bossert, 2008). Health governance involves three sets of actors: state actors (health and nonhealth, including health ministry, health insurance and pharmaceutical procurement and distribution agents, ministry of finance, the judiciary, regulatory bodies), health service providers (public-, private-, and voluntary-sector hospitals, laboratories, pharmaceutical companies, equipment suppliers, and education institutions), and beneficiaries (service users and the general public). Governance rules should ensure the accountability of key actors to the general public for

access to health services that achieve good health outcomes, satisfaction for patients and the general public, and equity. Accountability mechanisms can include judiciary and regulatory processes, advisory committees, and transparency of information. Governance also involves a policy process wherein actors with competing priorities have a "level playing field," with the aim of reducing unfair lobbying practices and corruption and increasing the voice of underserved populations.

States need to have sufficient capacity, power, and legitimacy to ensure effective governance. Yet ministries of health often lack the capacity to ensure a fair health policy process, which in turn undermines their legitimacy. This problem is particularly acute in fragile states (Brinkerhoff, 2007). A study in Nepal, Burma, and Afghanistan, for example, showed the link between service delivery and state fragility (Berry & Igboemeka, 2004). Civil society also needs to have sufficient capacity and power to engage in the health policy process, in its role as a representative of the broader public and as a watchdog of health services provision. In addition, an uneven distribution of power and external factors (either global or domestic) can hinder effective governance.

Within the broader concept of governance, regulation is traditionally defined as the imposition of rules by government, backed by the use of penalties, that are intended to modify the behavior of individuals and firms in the private sector. With the creation of autonomous entities within the public sector (e.g., self-managing hospitals), regulation commonly now extends to those organizations. Regulation is a role that all governments must carry out, regardless of their degree of involvement in health services provision. The traditional rationale for regulation relates to the arguments of market failure outlined previously as well as to the desire of governments to meet other social objectives, such as equity. Market failure creates the need either to regulate to make the market work better (e.g., to limit the control any one pharmaceutical firm may have over the market) or to prevent harmful effects (e.g., to ensure minimum quality standards for private clinics).

Although regulation is often thought of as action involving control, sanctions, and penalties, it can also take the form of incentives to encourage appropriate behavior. In health services, where outcomes are difficult to observe (i.e., it can be difficult to relate a treatment to a change in health), it can be argued that incentives are particularly appropriate, although they can be complicated to design and predicting response can be difficult (Baldwin & Cave, 1999). In LMICs where capacity and resources for enforcement are often weak, it has been argued that the concept of regulation should be extended to encompass incentives in combination with other tools such as consumer information, contracting, and accreditation, and working through other actors such as consumer groups or nongovernmental organizations (NGOs) can offer policy makers a wider and softer approach to enforcement (Ensor & Weinzierl, 2007). Indeed, the prime motivation for examining regulation in LMIC health systems has usually been determining how best to engage with the private healthcare sector. In this context, "prohibit, constrain, encourage, or purchase" (Montagu & Goodman, 2016) have all been strategies considered. The first two of these approaches are considered here, and the remaining two in subsequent sections.

Regulatory action seeks primarily to influence the following factors:

- Market entry and exit
- Remuneration of providers
- Quality and distribution of services
- Standards and quality

Key mechanisms used in the health system to regulate the provision of health services are summarized in **TABLE 13-2**. Controls over market entry and exit are not shown separately because they also serve to influence quantities and quality.

Licensing of professionals to provide services is one of the key forms of regulation, with professional councils usually being empowered to carry out this function. As new professions arise or become more important, eventually they are brought within the scope of laws. Although such laws dictate entry into the market, their prime rationale from a government perspective is to maintain quality and protect the consumer. However, actual experience demonstrates that licensing on its own is not adequate to ensure quality, since other factors affect professional behavior.

A second key type of regulation is licensing or registration of facilities, which is required before they can open. Legislation often specifies the requirements that particular categories of facility should meet, covering such aspects as trained staff, availability of equipment and supplies, and buildings. Because of the costliness of high technology, some countries have an approval process for the purchase of major items of equipment. Entry to medical school may also be controlled with the same aim: Controlling costs by limiting supply. In a normal market, such action might be expected to raise costs, but it is considered justified in health because of the power physicians have to generate their own income or to prevail on the government to employ greater numbers of physicians than the country can really afford.

Often countries are concerned about the geographic distribution of providers, and controls and

TABLE 13-2 Examples of Regulatory Mechanisms for Healthcare Provision

Variable	Mechanism	Examples from LMICs
Quantities/ distribution	■ Bans on provision	■ Abortion services (many countries)
	■ Licensing of providers	■ Universal for main professional groups
	■ Licensing of facilities	■ Increasingly common for hospitals and clinics
	■ Controls on number and size of medical schools	■ Common (e.g., Latin America)
	■ Controls on location of facilities and technology	■ South Africa: private hospitals; Tanzania pharmacies
	■ Incentives to practice in underserved areas and specialties	■ Many countries for doctors, often in the form of compulsory rural service
	■ Requiring capitation or case-based payment to control the supply of services	■ Social insurance in Korea (case payment) and Thailand (capitation)
Prices	■ Negotiation of salary scales	■ Zimbabwe: nursing salaries; Argentina: doctors
	■ Fixing of charges (e.g., for lab tests; drugs markup)	■ South Africa: drugs markup for medical schemes' reimbursement; Niger: wholesale and retail drugs markups
	■ Negotiation of reimbursement rates	■ Many social insurance schemes (e.g., Chile)
Quality	■ Licensing of practitioners	■ Universal for main professions
	■ Registration of facilities	■ Increasingly common; specifies structural standards
	■ Control on who provides which services	■ Restrictions on drug sales by unqualified staff (many countries); on drug dispensing by general practitioners (e.g., Zimbabwe); on a range of procedures (clinical officers, Kenya)
	■ Accessibility	■ Hospitals legally obliged to provide emergency care irrespective of patient financial status (Thailand, Malaysia)
	■ Required complaints procedures	■ Consumer laws applicable (India)
	■ Required provision of information for monitoring quality	■ Many countries
	■ Control of training curricula	■ Many countries
	■ Requirements for continuing education of physicians	■ Increasingly being introduced
	■ Accreditation	■ Increasingly being introduced (e.g., schemes in Taiwan, Kenya, and Thailand)

Data from Bennett, S., & Ngalande-Bamda, E. (1994). *Public and private roles in health: A review and analysis of experience in sub-Saharan Africa.* Geneva, Switzerland: World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/handle/10665/62173/WHO_ARA_CC_97.6.pdf?sequence=1&isAllowed=y

incentives are used to influence where new providers can set up. For example, South African provinces have the authority to license the creation and expansion of private hospitals, depending on "necessity" (Doherty, 2015). Certificate-of-need legislation has been used for many years to control the construction of new buildings and investment in new equipment in the United States (Certificate of Need State Laws, 2017).

Control of prices and reimbursement levels may have several purposes: To restrict incomes in the private sector so that remuneration differences between public and private professionals do not get too great; to ensure that health services remain affordable for the not-so-wealthy; and to restrict the financial burden placed on risk-pooling arrangements, such as social insurance or employer medical benefit schemes. However, given the power of the medical profession, there is a risk that price control will operate more in the interests of the profession than in the interest of the public; in addition, this practice can be difficult to enforce or monitor.

Control of quality is one of the prime concerns of regulation. Licensing and registration have this aim, as well as control of quantity. Doherty (2015) provides a mapping of these regulations affecting providers in seven sub-Saharan African countries. In Kenya, for

example, a private clinic must be kept in good order and state of repair, and must have a spacious consultation room, a sheltered waiting room, a treatment room, access to toilets, and transport.

Regulations often seek to control the nature of services provided, both to ensure that services are within the competence of a particular type of provider and to limit the scope for excessive service provision. Regulations usually specify which type of health professional can prescribe which type of drug, limiting, for example, the range of drugs that can be given by low-level health workers. It is quite common for private practitioners to be allowed to dispense medicines only if there is no pharmacy nearby. Where this rule does not exist, drug dispensing is often a major source of private physicians' income, leading to predictable concerns about overprescribing.

Control of training curricula is fundamental to ensuring quality, and is often one of the functions given by law to professional bodies. A trend in high-income countries, which is also becoming apparent elsewhere, is the requirement for professionals to receive regular refresher training to merit continued licensing. Such a provision is demanding on regulatory bodies, because it requires the introduction of monitoring procedures, training programs, and relicensing arrangements.

Accreditation is a process of certifying that a facility meets a certain standard. It is usually applied as a self-regulating procedure that is voluntary and managed by an independent body. In practice, it may act more as a regulatory device than as a peer review process, especially when accreditation is required for hospitals to be eligible for reimbursement from a social insurance scheme, as in Taiwan (Lu & Chiang, 2011).

Governments also regulate other markets with considerable relevance to health services, including the health insurance and pharmaceutical markets. Regulation of both of these markets has unique features. In the case of insurance, regulations may impose a particular approach to risk pooling (e.g., requiring schemes to give lifetime coverage or to use community risk rather than risk rating). In the case of pharmaceuticals, regulations may establish which drugs can be imported and which can be sold over the counter or require a doctor's prescription; they may also specify quality control procedures for imported and locally made drugs.

Separate consideration may be given in regulatory structures to not-for-profit providers. On the one hand, they may be treated more strictly. For example, their fee structures may be regulated, they may be required to provide a certain amount of free care to the poor, and requirements to provide information may be stricter. On the other hand, they may benefit from their not-for-profit status. For example, tax exemptions are often available to these organizations.

In practice, regulation encounters a number of key problems (Doherty, 2015; Mills et al., 2001; Sheikh, Saligram, & Hort, 2013). **EXHIBIT 13-3** highlights these issues for the five exemplar countries. One problem is that laws are frequently outdated and are difficult to change. For example, many LMICs have laws they inherited from colonial regimes that have not kept pace with the development of the private sector, resulting in whole categories of facilities that may be completely unregulated. Private laboratories are often a case in point.

EXHIBIT 13-3 Key Regulatory Problems in Selected Countries

Sierra Leone (Frederick Martineau, personal communication, July 31, 2017)

- Informal payments for care that should be free remain common.
- Large numbers of volunteer health workers work in the government sector for many years before being officially recognized with a salary.
- Drug procurement and distribution are particularly problematic, with frequent stock-outs requiring patients to visit private pharmacies.
- Informal practitioners are largely unregulated and unknown to the state.

Tanzania (Mtei et al., 2007)

- The country relies on legal instruments rather than complementing them with other approaches to regulation.
- There is a limited capacity to enforce regulations, especially with respect to drug prescribing and sales.
- The regulations covering private providers permit the health minister to set maximum prices for medical treatment, but on a practical level private hospitals are free to charge what they like.
- The registrar of private hospitals has a mandate to inspect private facilities, but there is little evidence such inspections are done.

(continues)

EXHIBIT 13-3 Key Regulatory Problems in Selected Countries *(continued)*

India (Mills et al., 2001)

- Few states require any registration and inspection of private hospitals.
- Practice by unqualified personnel is widespread.
- Unethical practices (e.g., payments between hospitals and general practitioners to encourage referrals) are widespread.
- No database of private providers exists.
- There is little ability to enforce regulations.
- Regulatory bodies lack resources.

Thailand (Teerawattananon, Tangcharoensathien, Tantivess, & Mills, 2003)

- The regulatory framework is largely complete, but its application is weak.
- Regulatory bodies lack resources for enforcement.
- There is insufficient information on activity in the private sector.
- It is difficult to control unethical practices (e.g., turning away emergency cases).
- Professional council regulation is largely ineffective.

Mexico (David Lugo Palacios, personal communication, July 31, 2017)

- Private hospitals are not subject to a strict process of accreditation that verifies their capacity to provide an acceptable standard of care.
- Public facilities are used by physicians for treating their own private patients.
- The private sector resists providing epidemiologic and other information.
- Some people are covered by two, or even three, public health insurance schemes, which in principle is not allowed by the *Seguro Popular* regulation. For example, more than 14% of *Seguro Popular* beneficiaries are also covered by social security institutions (FUNSALUD, 2012).

Another problem is that regulation requires substantial knowledge on the part of the regulatory bodies. However, it is common in LMICs for even basic information, such as lists of providers and facilities, to be incomplete. Moreover, the poorer the country, the greater the proportion of providers that are small and informally organized, making it difficult to require any regular provision of information by those organizations.

A third problem is that of regulatory capture: The body meant to be doing the regulating in practice operates in the interests of those being regulated, rather than in the public interest. This is a common problem in the case of regulation of professional groups, which is often done by the profession itself, leading to slow processing of complaints and concerns related to professional negligence. Moreover, the overlap of public and private interests can make it extremely difficult to introduce new regulations or change existing laws. In India, as in a number of countries, it is common to find government-employed doctors with private practices—with or without legal sanction—and senior Ministry of Health officials, as well as politicians, having financial interests in private-sector health services (Mills et al., 2001). Thus, a clear distinction between the regulators and those being regulated is lacking. In India, the consumer rights law provides an alternative channel for pursuing complaints, but consumer forum officials interviewed by Sheikh et al. (2013) reported that it was difficult to rule against doctors because of the subjective nature of medical cases and a perception that doctors were doing "fundamentally noble work" (p. 51).

A fourth problem is the inadequacy of the resources provided to the regulatory bodies to apply the laws effectively. Quality monitoring, in particular, requires regular inspection to ensure laws are being followed. This places great demands on the limited staff capacity of regulatory agencies, especially in regard to monitoring of drugs and clinics, whose outlets are numerous and widely dispersed. A further problem may be that low-paid staff seek illicit payments instead of carrying out their jobs effectively. An extreme form of this failure to monitor was found in China: Because government subsidies to public health activities were severely cut, environmental health units were dependent on revenue generation for much of their income and, therefore, tended to inspect those firms that were more able to pay their fees (Liu & Mills, 2002). Those that were less profitable, and in turn likely to have worse safety and hygiene practices, went uninspected.

A final problem is a lack of institutional structures to back up the regulatory process. Strong consumer groups, media, professional associations, and insurance agencies all have important roles to play in ensuring regulations are respected and enforced (Brinkerhoff & Bossert, 2008). The consumer role is particularly important, because consumers can identify problems through complaints procedures and legal action and also levy pressure more broadly through consumer groups. However, the common imbalance in power and access to resources between consumers and professionals suggests that complementary pressures are also important. One source of this type of influence can be purchasing agencies, which are considered in depth later in this chapter.

As part of efforts to strengthen their health systems, many countries have amended out-of-date legislation and sought to ensure that new private activities are brought within the scope of the law. Countries that previously banned or strictly controlled the private healthcare sector, such as countries in eastern and central Europe, some countries in Africa, and Vietnam and China, now allow and in some cases even encourage private-sector services through tax subsidies. However, much of the expansion of the private sector has taken place without deliberate planning, and often with little regulation. The challenges of both designing and implementing regulations have encouraged an emphasis on interventions that incentivize private providers to improve the quality and coverage of their healthcare services (Montagu & Goodman, 2016), including social marketing and vouchers (found to be successful) and accreditation and contracting (less evidence of success).

Financing

This section lays out a conceptual model for the financing of health services, and describes and evaluates the major sources of health financing in LMICs.

Conceptual Model

Financing refers here to the raising or collection of revenue to pay for the operation of the health system. This definition is narrower than that used by WHO, for example; WHO considers the financing building block to encompass raising money, pooling funds, and paying providers. Financing agents are those entities that collect money to pay providers on behalf of consumers. Financing agents may be publicly or privately owned, and may provide health services directly (e.g., the ministry of health through public hospitals and health centers) or purchase health services from providers (e.g., a private insurer may purchase inpatient care from a variety of hospitals).

There is some variation in the literature as to the definition of "sources" of financing. Sources may be defined as the entities that provide funds to financing agents. In this sense, individuals and firms can be thought of as the primary sources of funds. Individuals generate income in the form of wages or salaries, while businesses may earn a profit on capital investments or rent on properties owned. Resources may pass through several levels of sources before reaching the agents. For example, the ministry of finance can be thought of as a secondary-level source insofar as it generates funds by taxing the incomes of households and businesses and then transfers these resources to other government agencies to purchase health services. A single entity may act both as a source and as an agent of financing. For example, households commonly pay for health services both indirectly (through taxation, contributions to social and private insurance, donations to charities, and so on) and directly (through out-of-pocket payments).

More often, however, the term "source" is applied to the method whereby an agent mobilizes or collects resources. For example, the sources of financing for the ministry of health include personal and business taxes, and donations, loans, and grants from domestic and foreign agencies. The sources of financing for private insurance agencies are premiums paid by the enrollees in these schemes. In this chapter, the term "source" is used with this definition in mind unless otherwise indicated.

Description and Evaluation of Predominant Sources

This section defines the most commonly used sources of financing and briefly discusses the efficiency, equity, and revenue-generating ability of each source. **TABLE 13-3** summarizes the relative merits of each source.

Efficiency with respect to a source of financing involves a number of elements, including administrative (or technical) efficiency, stability, and flexibility. Administrative efficiency relates to the cost of the system's management and is the difference between gross yield (all funds that are collected) and net yield (that portion of the gross yield that is actually available for the purposes of health services delivery). This difference results from the costs of revenue collection, allocation, and distribution; advertising and promotion; and funds lost to corruption and fraud as well as the

TABLE 13-3 Evaluation of Health Financing Sources

	Efficiency			Equity		
	Administrative Efficiency	Stability	Flexibility	Horizontal	Vertical	Revenue Generation
Public Sources						
General tax revenues	Highı	Low	Low	High	Progressive	High
Retail sales taxes	High	High	Low	High	Regressive	Low
Lotteries and betting	Low	High	Low	High	Regressive	Low
Deficit financing	Low	Low	Low	Depends	Depends	Depends
External grants	Low	Low	Low	High	Progressive	Low
Social insurance	Low	High	Low	High	Regressive	Depends on size of formal sector
Private Sources						
Households	Low	High	High	Low	Regressive	High
Employers	Low to medium	High	Variable	Low	Depends	Low
Private insurance	Low	High	High	Low	Regressive	Low
Voluntary organizations	High	Variable	Variable	High	Progressive	Medium

cost of fighting that corruption and fraud. The stability of an agent is determined by the degree to which revenue raising varies with changes in economic or political conditions. Finally, for a financing agent to be efficient, there must be flexibility in terms of the allocation of funds to different expenditure categories. The least flexible sources of financing are those pledged to a specific activity. Public-sector sources tend to be less flexible than private-sector sources due to the rules and regulations that are often applied to government spending, as well as the political constraints on reallocation.

The concepts of horizontal and vertical equity of financing were introduced earlier. With respect to vertical equity, a progressive system is one in which lower-income groups pay a smaller proportion of their income to fund services compared to higher-income groups. A regressive system is one in which lower-income groups pay a higher share of their income for this purpose than do higher-income groups. A proportional or neutral system is one in which all income groups pay the same percentage of their income for financing purposes.

Apart from problems of inefficiency and inequity, health systems in many LMICs face the difficulty of simply not being able to generate sufficient funds to ensure that the entire population has access to a minimal package of health services. Thus, a goal of the

financing function of health systems is to increase the availability of funds for the purchase and provision of health services. As countries become richer and the demand for high-technology, hospital-based interventions increases, the goal generally shifts from generating funds to constraining the financial flow through the health system (i.e., cost containment).

Public Sources of Financing

Direct taxes are paid directly by individuals or organizations to government and include personal income tax, property and land taxes, taxes on domestic business transactions and profits, duties on imports and exports, and property taxes. Some portion of these resources may then be allocated to the annual budget for health services. The most commonly quoted example of general tax financing for health services is the United Kingdom's National Health Service.

Direct tax revenues should have relatively high net yields, though this will depend on the overhead costs of the government bureaucracy needed to collect, allocate, and disburse them. They may not be a particularly reliable or stable source of financing, because the health sector must compete directly with other social and economic programs for a portion of the government's budget; as such, this source may fluctuate depending on the economic and political climate. Furthermore, this source of financing is likely to be inflexible because it is controlled by public-sector agents that are constrained by rules and regulations and the political feasibility of reallocations.

Direct taxation achieves horizontal equity insofar as taxes on individuals are usually not related to characteristics other than income. Income tax is generally a progressive form of revenue raising across the world, including in LMICs (Asante, Price, Hayen, Jan, & Wiseman, 2016), because income tax rates usually rise as a person's taxable income increases. Direct taxes are typically progressive, as in Bangladesh, the Philippines, Sri Lanka, and Thailand, where direct taxes are almost exclusively paid by the better-off members of society (O'Donnell et al., 2008). Notably, however, the ability of taxation to redistribute resources from the rich to the poor is hindered when the wealthy are able to evade the payment of taxes.

Ability to mobilize resources is another strength of direct taxation. Although most developing countries are restricted in their ability to collect income taxes and indirect taxes (due to limited infrastructure and small formal sectors), the government has many other options for generating tax revenue, including property, business, and import and export taxes. For example, taxes on international trade in 2016 accounted for 24% of central government revenues in Bangladesh, 13% in Russia, and 33% in Namibia (World Bank, 2017), but overall these proportions have been declining over time.

Indirect taxes pass through an intermediary en route to government coffers. Such taxes are incorporated into the selling price of a good or service; they include sales and value-added taxes (taxes on a broad variety of items) and excise duties (imposed on the sale of specific items, such as tobacco products, beer, and liquor). Revenues generated in this manner are often allocated to finance specific programs. Taxes that are pledged to a specific sector or activity are described as "hypothecated," and the practice is known as earmarking.

As with direct sales taxes, the net yield of indirect taxes will vary depending on the efficiency of the government agency responsible for collecting them. Indirect taxes are likely to be reliable when they are earmarked for the health sector, or even specific projects within the health sector. The flexibility of this source may be constrained by the government rules and regulations that guide revenue allocation.

Indirect taxation, and excise duties in particular, is generally considered to be regressive, because poorer households often spend a higher percentage of their income on the goods being taxed (e.g., alcohol and cigarettes). As in higher-income countries, this has been found to be the case in LMICs, especially for value-added taxes (VAT) (Asante et al., 2016). In an exception to this pattern, Mills et al. (2012) found that VAT was progressive in Ghana and Tanzania, reflecting the types of goods and services liable for VAT and the local consumption patterns in poorer groups.

Lotteries and betting may also serve as sources of earmarked income for health services, although these methods are not often used. They have low net yields because they are costly to administer. As with indirect taxes, the resulting revenues are likely to be reliable because they are earmarked, but inflexible because they are administered by public agents. Like retail sales taxes, lotteries and betting tend to take a particularly heavy toll on the earnings of the poor, because of their popularity among lower-income groups.

National authorities can augment general tax revenues through domestic and international deficit financing (loans) and through grants. Deficit financing means that funds are borrowed for a specific project or activity and have to be paid back to the source over some future period of time. Domestic deficit financing is usually achieved through the issuance of debt certificates, or bonds, with guaranteed interest rates. International borrowing typically takes the form of loans from bilateral and multilateral

organizations. External grants are transfers to governments made in cash, goods, or services by foreign governments or organizations; they do not have to be repaid.

In 2008, the World Bank noted that there were more than 100 international entities providing development assistance for health. Total official development assistance to health reached nearly $24 billion in 2013, up from $7.6 billion in 2003 (Grollman et al., 2017). According to the same source, only 9.6% of all development assistance for health was spent on health system strengthening and general health-sector support. Major sources of grants include bilateral donor agencies (such as Britain's Department for International Development [DFID] and the U.S. Agency for International Development [USAID]), the Bill and Melinda Gates Foundation, and the Global Fund to Fight AIDS, Tuberculosis and Malaria. One of the largest financiers of health services in LMICs is the World Bank, which provides two main types of loans. The first type is intended for developing countries that are able to pay near-market interest rates. The second type of loan goes to the poorest countries in the world, many of which are in Africa. These loans carry little or no interest, and are long term (generally 40 years, including 10 years' grace). Loans that bear an interest rate substantially below market interest rates are termed soft loans.

The costs of processing and administering donor assistance in the health sector can be quite high, particularly when the aid to a country is spread among a large number of donor projects. Fragmentation is common: According to a recent study, a typical recipient country received aid for reproductive, maternal, newborn, and child health from an average of 15 donors between 2008 and 2013, although some countries had as many as 30 donors. In addition, it results in officials spending a large amount of their time meeting donors' requirements (Martinez-Álvarez et al., 2017). The stability of this financing source is also limited, insofar as external funds are typically of short duration, with no guarantee of renewal.

The flexibility of loan and grant financing is variable. At one time, it was common for external funds to be made available for specific, free-standing health projects; decisions on expenditures were usually made prior to the funds' disbursement. Later, encouraged by high-level declarations on aid effectiveness in the early 2000s (Rome, Paris, and Busan), donors increasingly chose to disburse aid as program-based approaches. These have the characteristics of being led by the recipient country, with a single program and budget framework; they promote donor coordination and use of country public financial management systems (Martinez-Álvarez, 2014). Program-based approaches include aid delivered as general or sector budgetary support, sector-wide approaches, and basket funds (where donors pool their funds and jointly agree with governments on the priorities and management procedures). More recently, disillusioned with a lack of progress and under increasing pressure to show results, donors have favored vertical project-based funds over program-based approaches.

The equity of deficit financing will depend entirely on how the loan is ultimately paid back. If, for example, funds generated through direct taxation are used to pay back a loan, then their impact may be progressive. External grants should be equitable since they are provided by wealthy nations and should be used to establish and run projects in remote or underserved areas (though this is by no means always the case). However, the extent to which they actually achieve a progressive redistribution of resources will depend on the extent to which the government shifts its spending as the result of the grants.

The displacement effect that one source of finance may have on others is termed fungibility, and the introduction or expansion of one source of financing may have an impact on the efficiency, equity, or revenue-generating ability of other sources. This may be especially important when large external grants or loans are introduced into a health sector. For example, it might seem that a large external grant earmarked for the treatment of tuberculosis (TB) among the rural poor would be equitable. However, in response to such a grant, the ministry of health may shift resources that would have been spent on this activity toward high-technology, hospital-based services. Thus, the positive impact that the new source of funding has on equity might be counterbalanced by the negative equity impact of displacement.

Despite debates about the existence and impact of fungibility, in reality little is known about the actual extent to which this practice occurs (Martinez-Álvarez, 2014), partly due to data limitations. Current estimates of health-sector fungibility range from a decrease of $0.27– $1.65 to a $1.50 increase in response to a $1 increase in development assistance for health. A country case study in Tanzania found some evidence of substitution of development assistance for health from the health sector to the education sector, perhaps reflecting a budget allocation process operating under conditions of scarcity and rational behavior by the government (Martinez Álvarez, Borghi, Acharya, & Vassall, 2016). Fungibility is one of the rationales for program-based approaches.

Only limited resources are generated through external loans and grants except in the poorest countries: Low-income countries spend, on average, $37 per capita on health; of this amount, 28% comes from development assistance (WHO, 2017). In specific countries, however, external funding can represent a much higher share of total health expenditures—for example, 49% in Mozambique.

Social insurance premiums are mandatory insurance payments made by employers and employees in the formal sector, usually as a percentage of wages; hence, they are often called payroll taxes. Social insurance payments can have relatively low net yields due to the cost of processing claims. According to WHO (2010a), over the period 2000–2007, administrative costs in high-income country social insurance funds amounted to approximately 4%. Data on LMICs are sparse, but the share attributable to administrative costs has been found to range from 1% to 27%, with an average of 23% noted in middle-income countries. The monies generated are likely to be stable because they are earmarked for the health sector; however, they may not be flexible, again related to government restrictions and requirements.

While the value of social insurance premiums is generally based on income, social insurance contributions tend to be regressive in countries with universal systems, such as France, Japan, the Netherlands, Spain, and South Korea (O'Donnell et al., 2008; Wagstaff et al., 1999). This situation arises because contributions are typically subject to a ceiling. Social insurance premiums may be progressive if ceilings can be eliminated or if low-income groups are exempt from contributions. In LMICs, where social insurance tends to include only wealthier, formal-sector workers, contributions are progressive. This is the case, for example, in China, Indonesia, and the Philippines (O'Donnell et al., 2008). In these countries, however, the insurance coverage is largely restricted to the better-off segments of the population.

The coverage achieved by social insurance schemes in many LMICs has been limited, because premiums can be readily collected only from formal-sector employees. In addition, achieving high compliance of employers can be challenging—as found, for example, in a study in Laos (Alkenbrack, Hanson, & Lindelow, 2015). The limited coverage has given rise to much criticism of the equity of social insurance arrangements, since a relatively small proportion of the population has access to better services than much of the rest of the population, and sometimes also benefit from government subsidies to the social insurance program. One response to this problem has been to use income from general taxes to expand coverage of health services for those people outside the formal sector. For example, this has been the policy in Thailand (Tangcharoensathien et al., 2011). Over time, funding schemes for different population groups can be made more comparable, with some countries, such as Costa Rica, the Republic of Korea, and Taiwan, achieving universal or near-universal coverage by combining funds from social insurance contributions with general tax revenues in a single national scheme, so as to ensure that all population groups have access to the same types and levels of care.

Another response, for countries at an earlier stage of health system development, has been to allocate public subsidies to voluntary health insurance in an effort to expand enrollment. As explained later, community-based health insurance schemes have generally stayed small, but a few countries—notably Rwanda, Ghana, and China—have seen considerable gains in population coverage by subsidizing local schemes and bringing them into a national set of arrangements. Rwanda rapidly scaled up coverage of its "Mutuelles de Santé," China reached universal coverage of basic medical insurance in 2011, and Ghana has scaled up district-based health insurance schemes as part of the National Health Insurance Scheme, to the point that approximately 40% of the total Ghanaian population held valid membership cards in 2016 (Alhassan, Nketiah-Amponsah, & Arhinful, 2016). India has employed a different insurance approach: The introduction of heavily subsidized state schemes largely intended for those persons below the poverty line, through contracts with insurance companies that arrange the provision of care and payment (RSBY, 2017). In general, increases in population coverage have proved easier to achieve than reductions in out-of-pocket payments, due to limited benefit packages, continued exposure to fees, and in some cases quality of care concerns.

Private Sources of Financing

Direct household expenditures include all out-of-pocket payments or user fees paid by the consumer of health services directly to the provider (including private practitioners, traditional healers, and private pharmacists). Even services provided by the government or an insurance program may include some element of copayment, which may take a variety of forms. Coinsurance means that the consumer is responsible for paying for a certain percentage of the costs of all services received. Limited indemnity means that the insurer covers health service costs only up to a prespecified absolute amount (or ceiling), above which

the consumer is responsible for paying. A deductible is a specific amount that must be paid by the consumer, after which reimbursement starts.

The administrative efficiency of direct household expenditures is low, due to the labor-intensive task of collecting fees from individuals. The stability of household spending on health will vary according to household income, but it is likely to be fairly stable unless economic crisis causes widespread poverty. Household spending is extremely flexible and will be allocated to the most pressing health needs as they are perceived by members of the household. Such spending is horizontally inequitable, since it varies according to factors such as distance lived from health facilities and individual preferences. It also demonstrates vertical inequity: For example, Asante et al. (2016) found that out-of-pocket payment was regressive according to all studies included in their systematic review, with the exception of Zambia, where it was proportional.

There are, however, ways in which user fees can be made more equitable. For example:

- Full-cost charges can be levied on patients who bypass the referral system by going directly to hospitals without being referred or having a genuine emergency.
- Charges can be levied for private rooms at government hospitals that charge for the full costs.
- Clinics at government facilities can be opened outside normal working hours for those who are willing to pay and want to avoid queues.

Direct employer financing occurs when firms pay for, or directly provide, health services for their employees. Employers as agents are likely to be more efficient than households, though less efficient than compulsory purchasers of care (such as social insurance schemes) due to the former's fragmentation. While employer financing may be reliable (in the absence of economic crisis), it is relatively inflexible because employers are biased toward specific types of health services (e.g., curative care). Direct payment by employers contributes to horizontal inequity in the health system as a whole because employed workers are disproportionately young and healthy. Because the benefits of employer financing are generally restricted to employees within the formal sector, this method of financing is likely to have little impact on the redistribution of resources among different income groups. The quantity of resources mobilized by employers can be significant in some countries; for example, in Zambia, financing provided by private employers accounted for roughly 9% of the country's total health expenditures (World Bank, 2012). As governments implement social insurance schemes, these tend to replace employer financing of health services.

Voluntary health insurance consists of two main types. Private commercial insurance involves payment of regular premiums to private insurance companies in return for coverage of prespecified health service costs. This insurance typically does not cover the costs of frequent, predictable events (such as pregnancy). Experience rating means that premiums are based on an individual's actuarially determined likelihood of illness. Community rating means that the premium is based on the pooled risk of a defined group of people (e.g., inhabitants of a geographic area or employees of a firm).

Private commercial insurance tends not to be an efficient method of mobilizing funds for the health sector (Donaldson & Gerard, 2005). Its net yield is low because of the costs involved in assessing risk, setting premiums, designing benefit packages, distributing the insurance, marketing, processing claims, reinsuring, and detecting fraud. Indeed, the administrative costs of private insurers in high-income countries were found to be, on average, 12%—three times the corresponding costs of social insurance schemes (WHO, 2010a). In unregulated markets, administrative costs plus profits may account for a substantial share of the premiums. Private health insurance is a stable agent of financing because it is not subject to political allocation processes, and it must be flexible to respond to consumers' needs and to attract clientele.

Private commercial insurance premiums are an example of a horizontally inequitable source. Experience rating means that premiums will vary according to factors that are considered by the insurer to be related to risk of illness, such as age, sex, and occupation. Private insurance tends to be regressive because rates are adjusted for risk, and the poor are at highest risk of falling ill. The ability to mobilize resources through private commercial insurance is limited in poorer countries because these schemes are targeted to a small (although well-off) segment of the population—less than 2% of the population in most cases.

The second type of voluntary health insurance consists of community-based schemes targeting the informal, usually rural, sector. Although a number of such schemes exist in various countries, they have struggled to grow in size and scope. Historically, they were an important stage in the development of universal coverage of health services in Europe and Japan. Today, however, few LMICs have schemes that are sufficiently developed to offer this sort of promise in the near future—with the exception of those discussed earlier, which have developed as part of social

insurance arrangements. Voluntary health insurance contributions for those in the informal sector are generally found to be regressive (Asante et al., 2015).

Charitable contributions are contributions made in cash or in kind. Examples include cash contributions from wealthy families, business enterprises, or religious organizations; community labor for construction and maintenance of local health facilities, including clinics and environmental sanitation projects; and local help in specific disease eradication campaigns. Voluntary organizations or NGOs have high net yields, although their ability to generate funds may be unpredictable. The flexibility will vary from one voluntary organization to another, but generally these sources prefer to fund specific types of health services, not necessarily those most suitable for the community being served. Voluntary organization funding should be progressive, in that such organizations raise revenue from the better-off segments of society, although their ability to mobilize resources is limited.

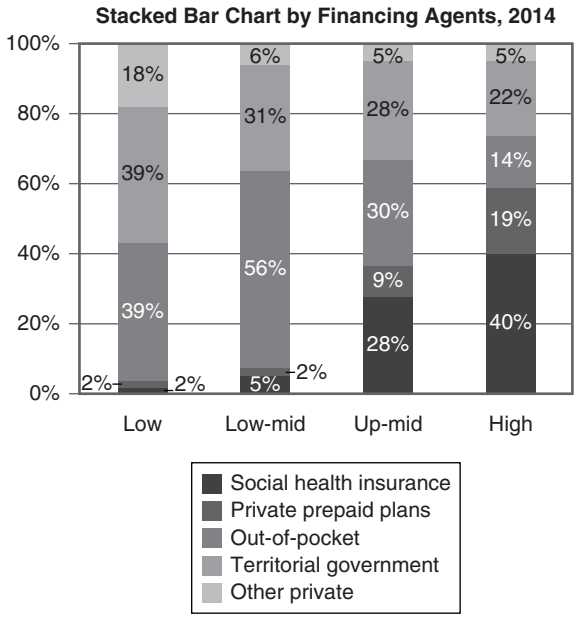


FIGURE 13-2 Health expenditures by financing agents.

Reproduced from http://apps.who.int/nha/database/Comparison_Report/Index/en accessed August 17, 2017. Data from Health expenditure series. Geneva: World Health Organization (latest updates and more information on countries are available at: http://apps.who.int/nha/database/Select/Indicators/en). For methodological notes on analysis see http://apps.who.int/nha/database/DocumentationCentre/Index/en

Mixes of Financing Sources

No one source of financing stands out as being superior in terms of all the outcomes considered, nor is there an optimal mix of sources that can be prescribed for all countries. Countries vary in terms of the number of different financing agents and methods that are used, and the mix can change over time within a country. The mix of sources used will depend in part on the relative importance that policy makers place on the various objectives described previously and in part on the mix of sources historically used in that country. Most LMICs have more pluralistic health financing structures than are found in high-income countries (**FIGURE 13-2**). Low-income countries typically finance the bulk of their healthcare expenditures through (1) direct household expenditures, (2) general taxation, and (3) external financing. The most important financing agents are typically the ministry of health (and other government agencies), households, and firms. As countries grow richer, external concessionary financing disappears, an increasing share of health expenditures is pooled, and a decreasing share comes from out-of-pocket payments.

Although using a variety of sources may increase the resources available for health care and allow for better adaptation to the diverse social and economic conditions that may exist within a country, it makes the pursuit of health policy goals more complex than in a system where one source of financing dominates. The greatest risks in a pluralistic health financing system are those of duplication, overlap, and inconsistency in functions and coverage. These problems can adversely affect both efficiency, because of duplication and competition for scarce human resources between different arrangements, and equity, because of limited risk pooling and lack of cross-subsidies among various income groups.

Reform of health financing has been high on the policy agenda of many countries for many years. In the 1990s, the introduction (or raising) of user fees was advocated, on the grounds that they offered a significant source of additional revenue, with their harmful effects on access being countered by exemption policies. Expectations are now less sanguine and opposition to user fees is widespread: Administrative difficulties of fee collection have been significant; ability to pay has been a significant barrier to use, and exemption schemes have not been effective; and even if there is ability to pay, people have been unwilling to pay if service quality does not improve. As a consequence, attention has turned to other ways of seeking to ensure that public funds are targeted to the poorest segments of the population (a topic considered in the next section, in the discussion of resource allocation).

The introduction of social health insurance was a core element of health system reforms in central and eastern Europe following the collapse of the Soviet Union, and has also been a key financing strategy in the rapidly industrializing countries of Southeast Asia. Social health insurance has also been of considerable interest in a number of African countries and in the

Caribbean. Most recently, the focus has shifted to the financing strategies appropriate for making progress toward universal health coverage, with continuing debate over the relative merits of insurance contributions and general tax funding as means to incorporate the informal sector into pooled funding arrangements (Tangcharoensathien et al., 2011).

Resource Allocation

The significance of processes of resource allocation in a health system has been clearly acknowledged only in the past two decades or so. Financing agents, such as government bodies at various levels (ministries of health and provincial offices), social insurance agencies, and private insurers, have always been one of the components of a health system, channeling funds from sources (taxpayers, payers of insurance premiums) to health services providers. A more recent trend is the emphasis on their role as active purchasers of services rather than as passive allocators of funds (Figueras, Robinson, & Jabowski, 2005; RESYST, 2014). For government bodies, such as ministries of health, this shift has required the creation of a clear distinction between their role as purchasers of health services and their role as providers of health services. For other financing agents, this separation of responsibilities requires them to take a much more active role in decisions about which services should be paid for and how they should be paid for.

The Purchasing Role

An emphasis on the purchasing role has arisen for a number of reasons:

- Integrated systems of purchasing and provision, such as those created in 1948 in the United Kingdom's National Health Service and existing in many similar public systems across the world, are thought to operate more in the interests of providers than in the interests of the general public and to lack incentives to be technically efficient. Highlighting purchasing as a separate function helps redress the balance and strengthen the power of managers.
- Patients, who in markets for other goods and services would be the purchasers, are believed—for the reasons laid out earlier—to be in a weak position to be active purchasers.
- In health systems such as that in the United States, where there has been considerable third-party payment (i.e., insurance agencies pay providers), cost escalation has been a major problem because patients choose their health provider, who is then reimbursed by the insurer. Although the insurers fund a large volume of services, traditionally they did not use this power to influence either the quantity or the price of services provided. The development of the managed care approach has led to an emphasis on the role of the purchaser as managing the provision of health services and ensuring that there are incentives for efficiency and cost containment (and also for equity, where the purchaser has public responsibilities). "Strategic purchasing" has included, for example, strict control over utilization, especially of specialists and of inpatient care; controlling drug costs by creating drug formularies and using generic medications; and disease prevention and management programs.

Purchasers may be of different types and sizes. The most limited role would be that assumed by an insurance agency, whose concerns are solely the patient group it cares for and ensuring it maintains its position in the marketplace. The most extensive role occurs when a purchaser has responsibilities for the health services of the population of a defined geographic area, and hence can plan services on a population basis. This is the position of public purchasers in health systems where the roles of purchaser and provider are distinguished. Agreements with providers define the services needed; these agreements may be formal legal contracts (and need to be, where providers are private bodies) or, alternatively, may simply be management agreements.

An important issue in the design of a purchasing role is whether there should be a single purchaser or a number of purchasers that compete with each other for clients. Private insurance agencies compete for clients, and it can be argued that this competition creates pressures for efficiency and services that meet client preferences. It remains an open question for tax-funded and social insurance–funded health systems whether a single purchaser or multiple purchasers is desirable. Reasons for caution about encouraging competition among purchasers include the following:

- If the system of allocating funds does not provide adequate compensation for meeting the costs of the health risks covered, the purchaser will have an incentive to avoid enrolling those individuals who are considered more expensive risks; this problem is known as cream skimming. In reality, it is difficult to predict the health service needs of a given population and compensate purchasers appropriately.
- It can be difficult for individuals to choose between competing purchasers: The more superficial

aspects of the promised package of health services may influence them, rather than technical quality, which is more difficult to judge. This would then lead purchasers to compete on the superficial aspects.
- If there are economies of scale in purchasing, these may be lost.
- Transactions costs—the costs of agreeing on arrangements for purchasing services—may be higher.

Whatever the number of purchasers, information systems are crucial in enabling purchasers to carry out their role. Government health systems traditionally have poor information on cost and quality, so new systems have to be developed to underpin the purchasing function. Insurance agencies often have much information since they are paying for services, but rarely exploit these data to monitor providers. Purchasers need information to act as active purchasers, including information on provider performance (e.g., waiting times, specific health outcomes where this can clearly be related to services provided) and adherence to standard treatment protocols (e.g., with respect to use of antibiotics).

Payment Mechanisms

A key influence on the success with which the purchasing role is carried out is the method chosen to pay providers, whether these be bodies responsible for services for specific populations, individual providers, or specific facilities. Although financial remuneration is not the only influence on provider performance, it is certainly a powerful one.

Authorities responsible for the provision of services to a specific population can be assigned a global budget. A key issue, then, is how that budget is determined. There has been increasing interest in using a resource allocation formula to calculate the budget appropriate to the populations of different geographic areas. The most well-known example is found in the United Kingdom, where funds have been allocated from the national level to regions using a formula that reflects the need for health services of each region, as proxied by indicators such as size of population and standardized mortality ratios (Department of Health and Social Security, 1976). Similar formulas have been used in South Africa and Zambia to allocate funds geographically. This approach is combined with various methods of paying providers and facilities within each geographic area, and can be the means to ensure funding distribution reflects equity goals.

Providers can be paid by salary, fee for service, capitation, or bundled payment (Porter & Kaplan, 2016). Salary represents a fixed annual payment unrelated to workload. Salary scales allow an individual's remuneration over time to be increased. Although in theory pay raises can be given on the basis of performance, in practice it is common for years of service alone to determine pay increases and promotions, thus undermining providers' incentives to work hard. Another problem with salary payment that is not inherent in the method is the level of the salary. In low-income countries, healthcare providers often have low salaries, further weakening their incentive to work hard or to work the required number of hours, and encouraging staff members to find additional ways of generating income, such as demanding informal payments from patients.

Additional elements can be added to salary payments to encourage good performance. These can be financial, as when an element of performance-related pay is included (e.g., a pay increase or end-of-year bonus can be based on a performance assessment), or nonfinancial (e.g., award of certificates or other ways of giving recognition to high performers).

Salary is the basis for remuneration in public systems, especially in hospitals and in privately owned facilities for nonmedical staff and even for physicians in some high-income countries. Although evidence is scanty, it is probably not uncommon in LMICs for at least some of the physicians in private hospitals to be salaried.

Fee for service has traditionally been the payment method for general physicians and specialists in a number of countries in continental Europe and in North America and Japan. It has also been adopted in some new social insurance schemes in Asia and eastern Europe. Physicians usually prefer this method of payment because it does not involve an employer/employee relationship—a factor that explains its persistence despite known problems with this payment mechanism. Where a financial agent pays the bill, there will be agreement on the fee schedule, which is usually negotiated with the medical association.

From a patient's perspective—particularly a patient covered by insurance—the fee-for-service payment scheme is attractive. It readily permits free choice of doctors, since payment can follow the patient. It encourages the doctor to be responsive to the patient, and there is no incentive to underprovide. However, from a purchaser's perspective, fee-for-service payment may encourage doctors to provide more consultations and more expensive procedures, especially where an ethical framework for medical practice is not strongly established. Administrative costs can be

higher because of the need to monitor claims, and some degree of fraudulent claims is inevitable.

Some adjustments can be made to fee-for-service methods of payment to address some of these problems. For example, the overall budget can be fixed, as in Germany, so that volume of services in excess of that budgeted for will cause the fee per item of service to fall. Alternatively, copayments can be required from patients, although there is little good evidence that these act as a constraint on physicians' behavior; indeed, copayments may simply render care unaffordable for lower-income groups in the population.

Capitation is the most common payment strategy for primary care services and involves a fixed payment per year per person registered with the provider. It may differ depending on the nature of the patient—for example, a higher payment may be provided for the elderly and children, reflecting their greater needs for health services. Capitation payment has been the traditional form of payment for general physicians in much of Europe, sometimes supplemented by extra payments to encourage particular aspects of primary care services (such as primary care teams) or priority services (such as preventive care).

This payment system supports continuity of care and an emphasis on preventive services, not just curative care; it makes the general physician a gatekeeper for hospital care, thereby encouraging the provision of services at the lowest possible level. The capitation payment scheme leaves the doctor substantially free to practice medicine and to organize the primary care service with little interference and with minimum administration. From the patient's point of view, capitation can ensure a personal relationship with a doctor and a personal medical record, although changing doctors may be difficult. However, because payment does not depend on the number of times a patient is seen, it can also encourage doctors to minimize the volume of services given to patients and to refer patients to the hospital unnecessarily, subject only to the need to keep patients sufficiently satisfied that they do not change doctors. Doctors may also try to avoid registering the more demanding and expensive patients. The extent to which this is a problem will depend on how much of the cost of services is covered by the capitation fee (e.g., whether drugs are paid for separately) and to what degree the capitation fee is adjusted for the risks of specific patient groups, such as the elderly.

Hospitals may be paid a fixed annual budget (often called a global budget) or in a variety of ways that reflect workload. Fixed budgets have traditionally been paid to public hospitals, but some countries have introduced them even for private hospitals that are paid by an insurance fund. Budgets are a highly effective means of cost containment and can provide a manager with great flexibility if discretion is allowed on line-item expenditures. However, there can be little incentive to have a high turnover of patients (since this increases costs) or to provide good-quality service. These problems can be addressed by good monitoring systems.

Payments to hospitals that reflect workload can be based on itemized bill, daily rate (including all recurrent costs), average cost per patient, case adjusted for diagnosis, types of services, or block and volume contracts. These payment models range from the most detailed—and therefore the most demanding and costly to administer—to the least detailed and the simplest to administer. All have their own particular advantages and disadvantages.

Payment by itemized bill has the same inherent problems as fee for service for doctors: It can encourage excessive procedures and hospital stays. Per diem rates discourage excessive procedures but encourage unnecessary stays. Average cost per patient discourages long length of stay but does not encourage technical efficiency. Case-based payment, particularly as developed in the United States with the diagnosis-related group approach, reimburses hospitals for the "average" case but provides an incentive to classify patients in more expensive groups or to shift patients out of the hospital earlier. Case-based payment is also information intensive and hence costly to administer. Bundled payment has developed to address the issues of case-based payment, notably by "bundling" all care needed for a particular episode in one payment, and making payment contingent on patient outcomes.

The introduction of contracts into the United Kingdom health system was accompanied in the early years mainly by block and volume contracts. In the former, a provider is contracted to deliver one or more services (e.g., acute hospital services) to a given population for a fixed sum—it is, in effect, a capitation payment. A block contract transfers financial risk to the provider, so it may encourage cream skimming and the provision of too few services. This type of payment, which is often risk adjusted to a certain extent, is increasingly being used in managed care arrangements and between social insurance agencies and providers, as in Thailand. In contrast, volume contracts are appropriate where the purchaser wishes to limit the number of procedures paid for (e.g., elective surgery). Purchasers can use these contracts to take advantage of economies of scale that can be achieved by larger units with high volume.

These many different payment methods provide a confusing range of options. The key issues relate to the risks of undesirable consequences; they are summarized in **TABLE 13-4**. In particular:

- What are the incentives to overprovide or underprovide services to patients within the facility? A key issue is whether the purchaser or the provider bears the financial risk.
- What are the incentives to be technically efficient?
- What are the incentives to exclude altogether certain types of patients?
- Given these incentives, what are the administrative costs of the payment system together with the monitoring required to prevent abuse?

This provides a rather crude basis for evaluating methods—not least because financial remuneration, although important, is not the only factor affecting the behavior of providers. The effects of a method in practice will also depend on the system and context within which it is introduced—for example, the extent to which the following characteristics are present:

- Purchasers or patients can change providers.
- Strict ethical standards are adhered to and monitored by the medical profession, thereby limiting cream skimming and both under-provision and over-provision of services.
- The media are active in publicizing cases of medical negligence.
- Consumers are well informed and able to exercise their right to choice effectively.

Because each payment approach has its advantages and disadvantages, which also depend on the context in which it is introduced, it is difficult to be prescriptive. In practice, some of the problems with any one method may be addressed by combining methods (e.g., using capitation payment but with additional fees for certain procedures). Perhaps the strongest conclusion that can be drawn is that fee-for-service payment should be avoided as the main payment approach. Ensor et al. (1997) concluded that providing care on a fee-for-service basis costs one-third more than using capitation, without substantial differences in health

TABLE 13-4 Key Payment Methods and Their Consequences

Payment Method	Unit of Services Paid For	Key Undesirable Consequences
Salary	Usually one month's work	Restrict number of patients and services provided
Fee for service	Individual acts or visits	Expand number of cases seen and service intensity; provide more expensive services and drugs
Capitation/block contract	All relevant services (e.g., primary care; hospital care) for a patient in a given time period	Attract more healthy registered patients and avoid less healthy ones; minimize contacts per patient and service intensity
Fixed budget	All services provided by a facility in a given time period	Reduce number of patients and services provided; keep patients in hospital longer
Daily rate	Patient-day	Expand number of bed-days (through longer stays or more admissions)
Case payment	Cases of different types	Expand the less serious cases seen; decrease service intensity; provide less expensive services
Bundled payment linked to patient outcome	Single payment, contingent on outcomes, for all care required to treat condition, regardless of who provides care and where	Lack of clarity over episodes of care for chronic conditions; discourage necessary specialty care

Data from Bennett, S. (1997). Health-care markets: Defining characteristics. In S. Bennett, B. McPake, & A. Mills (Eds.), *Private health providers in developing countries* (pp. 85–101). London, UK: Zed Books.

outcomes, and that fee-for-service payment for outpatient care is associated with 11% higher expenditures in OECD countries. Other points are that no payment method will work well if providers think they are seriously underpaid, and the more complex the payment method, the greater the need is for a sophisticated information system that organizes payment and monitors provider behavior.

Innovations in Resource Allocation

Many countries have introduced changes to traditional approaches to allocating resources and paying providers. One of the most common features of health system strengthening programs has been the introduction of a purchaser–provider split in public systems that were previously integrated. For example, in Thailand, the social security office acts as a purchaser on behalf of the insureds and accredits and monitors the hospitals that they use. Along with the specification of the purchaser's role has come an emphasis on contractual arrangements between purchasers and providers. At one extreme, these arrangements may be seen as no more than the formalization of a management relationship, as where annual contracts are agreed upon between the ministry of health and a regional health authority (Trinidad), or between a province and a district (South Africa, Zambia, and Zimbabwe). At the other extreme, the contracts may be awarded on a competitive basis and may be legally binding. Despite the popularity of this type of reform, its value has not been well established, and reforms have experienced considerable implementation difficulties, especially where they were accompanied by a policy to decentralize employment contracts of health workers.

Another development has been the use of outsourcing or contracts—usually with NGOs, though sometimes with commercial providers—to improve or expand service provision in contexts where governments function poorly or have limited capacity (Liu, Hotchkiss, & Bose, 2008). The experiments with contracting district health services in Cambodia are widely quoted, and similar arrangements have been tried in Afghanistan and Rwanda, for example. Contracting with NGOs has also been used widely to expand availability of disease-specific services such as HIV prevention and TB treatment.

Health system reforms have commonly involved changes to traditional modes of payment, although the nature of the reforms has depended on the starting point and can be severely constrained by powerful interest groups; hence, a great variety of reforms are apparent in practice. Although, in general, fee-for-service methods of payment are seen to be undesirable, countries such as those in eastern and central Europe that previously provided salaries to primary care providers have sought to raise remuneration levels and encourage greater productivity by using a mix of salary, fee for service, and capitation. Indeed, capitation appears to be one of the areas where there is the most experimentation worldwide. Traditionally an approach for paying primary care providers, capitation has been extended to pay for hospital services as well (e.g., in Thailand) and is the basis for payment in many managed care arrangements.

Global budgets for hospitals have been a feature in Europe under different funding regimes. Concern that they do not provide incentives for efficiency has led to the introduction of mechanisms that relate payment to measures of hospital activity, such as bed-days, cases, or even specific services. Some similar reform trends can be seen elsewhere; Thailand, Taiwan, and Korea have all tried case-based payments, with Thailand adopting the approach nationwide.

"Pay for performance" (also known as P4P and as results-based financing) to both individual workers and facilities has also attracted considerable interest. It can be distinguished from fee-for-service or case-based payment by its aim to use measures of performance rather than merely activity, though in practice this distinction may not be clear. Many individual studies have been done of various P4P schemes, and various systematic reviews are available as well. For example, Oxman and Fretheim (2009) concluded that there is good evidence that financial incentives targeting recipients and providers can be effective in the short term in achieving well-defined behavioral goals, such as increased provision of specific primary care services. There is less evidence that they can sustain long-term change. There is also concern, and some evidence, of unintended consequences, including reducing the provision of non-incentivized services and gaming of the system (e.g., false reporting).

Concern that the poorest members of a population often do not benefit from public funding has led to the use of various approaches to improve targeting, including geographic targeting through resource allocation formulae, focusing additional funding on specific programs targeted at populations with the least access, equity funds that reimburse health facilities for exemptions, social funds that often involve local communities in decisions on spending (World Bank, 2003, 2004), and vouchers or cash payments that channel purchasing power directly to those in need, such as vouchers for pregnant women that reduce the price of an insecticide-treated mosquito net, or

cash to pregnant women to help cover the costs of accessing safe delivery services. An evaluation of one of the world's largest financial incentive programs (Powell-Jackson, Mazumdar, & Mills, 2015) found that such cash payments to women in India were associated with increased use of maternity services, but not with reduced neonatal deaths, possibly because women predominantly sought care from faculties that were not equipped to handle complex cases.

In those countries introducing compulsory insurance and those engaged in reforming their existing systems, a key—and controversial—issue has been whether to encourage competition between insurance funds and to encourage choice of insurer. In Colombia and Argentina, for example, there is competition among insurers to enroll individuals in social insurance, with insurers compensated by a risk-adjusted capitation payment (Londoño & Frenk, 1997). A study in Colombia found some evidence of sickness fund selection based on health status, confirming fears that cream skimming can be a problem in the presence of competitive pressures on insurers (Trujillo & McCalla, 2004). In India, the huge RSBY scheme uses competitive bidding to identify the insurer that will manage the scheme at the district level; this insurer is then required to enroll eligible households within the district (RSBY, 2017).

A trend that is likely to accelerate in the future is the increasing prominence of healthcare companies interested in breaking into LMIC markets. Options for involvement include contracting with governments or social insurance agencies to manage health programs for particular population groups, covering groups that are allowed to opt out of social insurance arrangements, or simply taking advantage of the growing demand for private hospital care, especially in South and East Asia. These developments are addressed in more detail in the *International Trade and Health* chapter.

The specification of basic or essential packages of interventions, which purchasers require providers to make available, was first proposed in the World Bank's World Development Report 1993. This approach involves analyzing the burden of disease and the cost-effectiveness of interventions, so as to identify high-priority interventions that governments should ensure are universally available. It has been followed by the Disease Control Priorities Project, which in 2017 completed its third iteration. This project has introduced several new features, most notably systematic assessment of priorities for intersectoral interventions as well as health-sector interventions, and for the latter, allocation to five platforms—population-based care, community-level care, health center, first-level hospital, and referral hospital (Jamison et al., 2018). A "model" concept of essential universal health coverage has 244 interventions that provides a starting point for country-specific analysis of priorities. Of these interventions, a total of 104 are considered to constitute a model highest-priority package.

While the global analyses have been replicated in a number of countries, shifting allocation of public subsidies away from lower-priority services such as tertiary hospital care has proved politically difficult, and there is as yet inadequate evidence of the successful implementation of this approach to priority setting. Perhaps as a consequence, attention has turned to the institutions needed by countries to support the priority-setting process and embed health technology assessment processes in routine public decision making. Both NICE (the U.K. National Institute for Health and Care Excellence) and HiTAP (the Thai Health Intervention and Technology Assessment Program) have provided examples of how this aim might be achieved (Tantivess, Teerawattananon, & Mills, 2009).

Provision of Services

Health service providers in LMICs can be categorized into seven main groups:

1. Government-run health services for the general public (including the services of the ministry of health and those services coming from other government ministries, such as local government and education).
2. Services run by social insurance agencies for the insured and their dependents.
3. NGO services, including those run by church organizations and charitable groups.
4. Occupational healthcare providers (medical services provided by employers for their employees). This includes services for the armed forces and police, which come under government ministries. Universities may also provide services for their own staff.
5. Private, for-profit allopathic providers, both individuals and facilities.
6. Traditional systems of medicine, such as Ayurveda, homeopathy, and Chinese medicine; and traditional healers of various types, including traditional birth attendants.
7. The informal sector of drug peddlers and unqualified practitioners.

In general, the richer a country, the more organized and structured and the less diverse the system of health services provision. For example, over time

government-run services may be brought under the aegis of a single structure, as they were in the United Kingdom under successive rounds of reorganization; alternatively services of the ministry of health and social insurance agency may be amalgamated, as they were in Costa Rica. As government services and services for the insured improve, there is less reason for employers' health services to provide general medical care. Moreover, as regulatory structures strengthen, the informal sector becomes much smaller. Finally, as the availability and quality of prepaid health care improve, there is less reason for households to purchase care directly.

Information on the physical supply of health services in LMICs is sorely lacking, and indeed these data resources are much weaker than information on the demand side. At an aggregate level, for example, there is information on beds and health professionals. The number of hospital beds (acute and chronic) per 1,000 people ranges from 0.10 in Mali and Iran, to 13.7 in Japan, and the number of physicians per 1,000 people ranges from .01 in Liberia to 7.74 in Qatar (Index Mundi, 2017). But interpreting such data is difficult: The type and quality of health service inputs are varied, and there is no agreed benchmark for what constitutes under-provision or over-provision of services.

Another lens that can be used to examine the supply of health services is the level of use. This information is more readily available from country surveys, but difficult to summarize because health services comprise many different types of services aimed at different age groups. However, countries can be assessed by, for example, the extent to which groups with specific needs, such as women and children, have high levels of coverage of antenatal care, immunizations, and so on. The Countdown to 2015 (now the Countdown to 2030), for example, has systematically assessed country performance on health outcomes and services for women and children (Countdown to 2015, 2015), and WHO is now tracking universal coverage progress including basic hospital access and health worker density (ratios of beds and physicians to population).

One of the key issues in health services provision has been what is called the public–private mix—that is, the relative numbers and performance of providers in the public and private sectors and how the two sectors interact. McPake (1997) has suggested that there may be at least three different patterns. Among the lowest-income countries, the formal private, for-profit sector is small, but especially in Africa there is a rather larger NGO sector. The informal sector is large, consisting of unregistered allopathic providers, drug sellers, and a variety of traditional practitioners. Most formal health services are provided by the government, but health workers often have their own private practice.

In some LMICs, especially in Asia, the private, for-profit sector plays a much more important role in health services delivery; this constitutes the second pattern. For example, in India 70% of total health expenditures comes from private sources, and of this amount 89% takes the form of out-of-pocket payments (WHO, 2017). An estimated 93% of hospitals and 63% of hospital beds are privately owned, mainly by individual physicians (Peters et al., 2002). Large and concentrated populations may be one explanation for this pattern; another may be a longer history of Western health services and training of professionals, together with government health services that have never extended to provide coverage of the whole population. However, even in these countries, private provision is still concentrated at lower levels of the health system. For example, in India, private-sector providers are generally lacking at the tertiary care level, and the private sector provides 79% of outpatient care for those persons below the poverty line, much of it of low quality and provided by untrained providers including drug sellers.

McPake (1997) has also identified a third pattern of provision, in countries with rather higher per capita GNP. This pattern usually includes a major role for social insurance (funding services either through its own facilities, as is common in Latin America, or through mixed public and private providers, as is more common in Asia) and a private sector that is playing an increasingly important role in healthcare provision (certainly at the secondary level and sometimes also at the tertiary level).

Because the relative roles of public and private sectors have been a key policy question, there has been much interest in whether private providers are more efficient than public providers. Basu et al. (2012) undertook a systematic review of this issue in relation to the themes of accessibility and responsiveness; quality; outcomes; accountability, transparency, and regulation; fairness and equity; and efficiency. This review confirmed some existing views: Private providers had poorer medical standards of practice and patient outcomes, but patients valued their more speedy care and greater responsiveness. Efficiency tended to be lower in the private sector than in the public sector, in part because of incentives in the former to provide unnecessary tests and drugs. Public-sector services were more likely to suffer from lack of resources: equipment, drugs, and trained health workers. If the private sector is defined broadly, to include unlicensed

and uncertified providers, more care was obtained in the private sector than in the public sector; however, if unlicensed healthcare providers are excluded, the majority of people accessed public-sector care.

This review demonstrates the challenge of comparing the public sector with a private sector that is enormously diverse, both within a country and across countries. In India, for example, at one end of the spectrum are private hospitals delivering services of international standard; at the other end are unlicensed and unqualified practitioners using Western prescription drugs. In between are trained physicians and Western-style hospitals whose quality of care can be extremely poor (Bennett, 1997). Private clinics, although in theory run by doctors, in practice often depend on staff with little training because, in many countries, doctors will also have a public post where they spend at least some part of their time, as well as a stake in several clinics. Financial relationships between different types of facilities are also a concern in a number of countries; hospitals, laboratories, and diagnostic centers may pay general practitioners to refer patients to them or may pass a share of their fee to the referring physician.

While public-sector services show less extremes of performance, they also commonly demonstrate problems of inefficiency. While absolute lack of resources can limit their costliness, there is considerable evidence of poor resource use, such as low staff productivity and waste of drugs and supplies (WHO, 2010b). As in the case of private facilities, a variety of explanations can be posited for observed public provider inefficiencies, some of which do not relate to ownership per se. In particular, decision making may be centralized and staff at hospital level given little power to control resource use. Low salaries also contribute to poor performance, because they reduce staff motivation, and staff may need to spend time generating income in other ways to ensure an income adequate for survival.

Innovations in Service Provision

Initiatives to strengthen service provision have included decentralization, increasing competition and diversity of ownership, strengthening primary care services, and quality improvement.

Decentralization of management of services has been an almost universal theme for many years. At the national level, decentralization has taken the form of restructuring the role and functions of the ministry of health. For example, some countries have created executive agencies to take over management responsibility at the national level (e.g., Ghana), leaving the ministry to concentrate on regulation, policy, and monitoring. Some form of decentralization to intermediate and local levels is also a common theme. Most countries taking this tack have chosen to decentralize services within a hierarchical structure with the ministry of health at the top, although there are some notable examples of devolution of health to local government (e.g., the Philippines). Finally, decentralization to the hospital level is a common trend in countries with centrally funded, public systems of health provision (Mills et al., 2001). In the lowest-income countries, because of limits on management capacity, this reform may be confined to teaching hospitals, but in some countries, such as Indonesia, a wider range of hospitals have been made "autonomous" or "corporatized."

Competition between providers is also being widely promoted as a means to encourage efficiency, albeit with rather more caution and doubts about its effects than in the case of decentralization. Likewise, diversity of ownership is promoted as a means to increase competitive pressures, especially within poorly performing public systems of provision. African governments have been urged to give greater emphasis to NGOs. In a number of Southeast Asian countries, tax incentives have been provided for the construction of private hospitals (e.g., India, Philippines, Thailand), although from the perspective of the health sector as a whole there are considerable doubts about the desirability of this policy (Mills et al., 2001).

Strengthening the role of primary care has long been a theme in reforms in many countries, although often without substantial changes in resource allocation patterns. Comparative historical studies of countries that have seen considerable improvements in the health status of their populations, while spending only a relatively modest amount per capita, have highlighted the role that an effective primary care system played in these health gains (Balabanova et al., 2013). The recent emphasis on universal coverage has also focused attention on the critical importance of ensuing everyone has access at least to good-quality primary care and referral to a local hospital. The United Kingdom's reforms of the 1990s, which gave funds to primary care doctors to purchase other services, have aroused much interest, although efforts to create high-quality primary care services, with a gatekeeper role for the primary care provider, do not as yet appear to feature strongly in reforms in most LMICs, with the exception of Thailand. Policies haves focused on more incremental changes, such as integrating disease-specific services into primary care.

Included in many country plans have been measures to improve the quality of the services provided. In high-income countries, a range of approaches is being used, including evidence-based medicine, technology assessment, clinical guidelines, medical audits, quality assurance methods, and payment incentives. These are beginning to appear in less wealthy countries as well, but in the lowest-income countries more basic problems have been addressed, such as availability of drugs and supplies, improvement of staff skills, and maintenance of equipment.

Increasingly, human resources are being highlighted as an area where improvements are most needed. In many of the poorest countries, the number of trained health providers is grossly inadequate to provide even a basic package of services to the whole population. The international brain drain has also had a devastating effect on some countries (Cometto, Tulenko, Muula, & Krech, 2013). Brain drain is one symptom of a widespread problem in many countries whose health systems are characterized by low financial remuneration of health workers, low motivation, poor working conditions, and poor performance. To remedy this problem, it has to be tackled by action on a number of fronts. Promising approaches include greater use of nonphysician clinicians and community-based workers; improved working and living conditions; and support for circular migration, in which migrants are encouraged to return to their home country having gained valuable skills and experience elsewhere.

Countries have also sought to increase the influence of users and communities over health providers—to encourage patient-centered care, and to hold providers accountable for good performance—although there are few well-evaluated examples. In many of the poorest countries, local government structures are weak, ruling them out as an immediate means for ensuring local representation. Some reforms include the introduction of health or hospital boards with citizen representation. Often community involvement is seen to occur via NGOs rather than government health services. Little attention has been paid to patient rights, and patients tend to appear as the object of reforms, rather than as the subject.

Performance of Different Types of Systems

The previous sections have demonstrated that the design of health systems varies greatly between countries, particularly with respect to the following:

- The sources of funding (e.g., balance between general taxes, insurance contributions, and out-of-pocket payments)
- The number and degree of integration of financing agents and providers (are there large numbers of financing agents or one major one, such as a ministry of health or single social insurance agency; are financing agents and providers integrated or separated?)
- The ownership of providers (public; private, not-for-profit; private, for-profit)
- The extent to which the whole population of a country has access to the same services, or different groups in the population have different entitlements and use different providers

Perhaps not surprisingly, these marked differences have led to intense debate over whether any one design can be shown to perform better—in terms of criteria such as efficiency and equity—than any other design. Attention has particularly focused on differences between the U.S., Canadian, and British systems. The United States has historically relied heavily on voluntary insurance organized largely through employers, plus publicly funded programs for low-income patients and older adults. Canada has a compulsory national insurance system, whereas the United Kingdom relies largely on general tax revenues to finance its health system. Whereas most services in the United Kingdom are publicly owned, privately owned services play an important role in Canada and especially the United States.

TABLE 13-5 shows some key comparative indicators for the United States and Canada for 2009 or a similar time frame (Folland, Goodman, & Stano, 2013). As seen in the table, Canada had substantially lower per capita spending on health and GDP share than the United States. The United States spent 82% more on health care per capita than Canada, even though 19% of its population had no health insurance in 2009. Health status indicators such as life expectancy were better in Canada. Studies of waiting times and physician practice patterns show that in some instances Canadians got less health care or had to wait longer for services, but there were few observable effects on mortality and other outcome indicators.

Explanations as to why Canada spent less highlights two key differences between the U.S. and Canadian health systems. First, physician fees and hospital costs were significantly lower in Canada, no doubt because these charges are regulated. Physician fees are negotiated between physician associations and provinces, and hospital budgets are set by the provinces.

TABLE 13-5 Comparative Data on U.S. and Canadian Health Systems

Indicators	Canada	United States
Population, 2010 (in millions)	33.6	309.1
GDP per capita, 2010 (2010 $US)	39,400	47,200
Per capita health expenditure, 2009 ($US PPP)	4,363	7,960
Health spending, 2009 (% of GDP)	11.4	17.4
Percentage of health spending, 2008:		
Public expenditures	70.2	46.5
Inpatient care	27.3	24.5
Outpatient care	25.4	44.2
Pharmaceuticals	17.2	11.9
Acute care inpatient beds/1,000 population, 2007	1.8	2.7
Average length of stay, 2007 (acute care days)	7.5	5.5
Uninsured population, 2009 (%)	0.0	18.7
Out-of-pocket payments per capita, 2009 ($)	636	976
Private insurance expenditures on health, 2009 (%)	13.4	34.4
Life expectancy (years) at birth: females, 2005	83.0	80.4
Life expectancy (years) at birth: males, 2005	78.3	75.4

Note: Financial data are given in U.S. dollars.

Modified from Folland, S., Goodman, A. C., & Stano, M. (2013). *The economics of health and health care* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Second, the Canadian system had substantially lower administrative costs. Administrative costs accounted for an estimated 31% of total healthcare spending in the United States in 1999, as compared to 16.7% in Canada. Studies comparing expenditures in a larger number of rich countries have found that countries in which health services are financed primarily by private payments have the highest expenditures, but there is no evidence that this higher spending translates into better health status. Health systems where there is comprehensive risk pooling based on compulsory insurance or tax finance and covering the whole population appear to be more cost-effective (McPake et al., 2013).

Similar, detailed comparisons have not been done for LMICs. They can take two key lessons from richer countries:

- A significant public share in financing enables greater control over expenditures, meaning that higher population coverage can be achieved at lower cost.
- The greater the fragmentation of the health system and the greater the reliance on private insurance, the greater the proportion of total health expenditures consumed by administrative costs.

However, rich countries demonstrate that there are a variety of ways in which a strong public role and

coordinated health system can be achieved, and that the traditional model prevalent in many LMICs of an integrated public system financed from general taxation is only one approach.

Universal Health Coverage and Global Health Security

Two themes are currently dominating health systems concerns in LMICs, and are considered here because they concern the system as a whole rather than the individual functions that were considered separately earlier in this chapter. These themes are universal health coverage and global health security.

In 2010, the World Health Organization urged its members to aim to provide affordable universal coverage and access, defined as ensuring that all people can use the promotive, preventive, curative, and rehabilitative health services they need, of sufficient quality to be effective, while ensuring that the use of these services does not expose the user to financial hardship. This call to action accelerated a new global movement, which gradually became endorsed by all the major global players in global health. Universal health coverage was recognized in the SDGs, with goal 3, target 3.8, specifically assessing progress toward it. Existing initiatives relating to donor coordination and health system strengthening were built on to establish the International Health Partnership for UHC 2030. The WHO and World Bank support the global coordination of UHC 2030.

Universal health coverage requires countries to address health system weaknesses and limitations on a number of fronts. In terms of governance and regulation, two aspects are critical. First, governments need to commit to support the development of the social solidarity that is required to underpin a universal health coverage goal. Although universality is most likely to be achieved via a combination of arrangements and through prioritizing a limited range of services, it still requires governments to demonstrate a commitment to equity that has often not been visible in the past. Second, given that resources need to be deployed across both the public and private sectors, strengthening the regulatory role of government will be critical in enabling resources currently in the private sector to be deployed to support universal health coverage.

Financing poses major challenges given the cost of even a limited basic package of services. Jamison et al. (2018) calculate that achieving full implementation in 2030 (assuming 80% effective coverage) of their highest-priority package would cost $25–$45 (in 2012 dollars) per capita per year on top of what low-income countries are already spending; implementing the essential package would cost $57–$103 per capita. Even the high-priority package would require low-income countries to spend an additional 3% to 5.5% of their current gross national income on health care. In lower-middle-income countries, this increase would be 1.3% to 2.1%, a more affordable level.

While economic growth should help generate some of the resources, Ly et al. (2017) have identified a major challenge by analyzing the changes in the composition of health expenditures in sub-Saharan Africa from 1995—namely, the increase in government expenditures on health since 1995 has not been sufficient to compensate for the reduction of development assistance for health as countries "graduate" from low-income to middle-income status. Indeed, Ly et al. found that out-of-pocket expenditures increased during this transition. These authors argue for an approach to universal health coverage based on "progressive pragmatism," consisting of governments and donors agreeing on feasible targets based on the economic situation of the countries and donor finances, adapted to the sociopolitical realities at the national and global levels. Their list of recommendations for different countries demonstrates the changes also required in the resource allocation and provision functions to underpin universal health coverage, as well as in the broader economy:

- Middle-income countries that can afford the resource package (such as South Africa) need to focus on governance, service quality, reduction of income inequality, and strengthening of tax systems to reduce dependency on natural resources.
- Low-income countries that will soon graduate to middle-income status (such as Kenya) face the challenge of decreasing donor support before they have developed sufficient domestic resource mobilization strategies. Apart from agreeing on transitional amounts of external funding and improving domestic resource mobilization, these countries need to address issues of insurance pool fragmentation, management capacity, and the delivery of cost-effective, high-quality services.
- Low-income countries that will not be able to afford an essential package of services by 2030 (such as Madagascar and Sierra Leone) will still rely on donor financing, but should focus their attention on economic growth.

The second theme is that of global health security. During the 64th General Assembly in 2011, WHO

adopted a resolution to strengthen national health emergency capacities and health system resilience. Threats to health include emerging infectious diseases, human-made and natural disasters, and conflicts. There is increasing recognition that to respond to health emergencies effectively, a health system needs to be resilient (Kruk et al., 2015). In a review of health system resilience, Bayntun et al. (2012) found that each of the health system building blocks includes elements to build resilience, but also important aspects that have not been considered. For instance, while there has been recognition of the importance of training health workers on disaster management, less attention has been paid to establishing the commitment and ensuring the welfare of health workers during such events. Health workers are at high risk of infection in pandemics, for example, and need to be reassured that they will have access to the right equipment, prophylaxis, and treatment. Other areas that need improvement include incorporating ethical decision making into leadership and governance structures, improving the use of information systems and communication, carefully considering the suitability of a mass vaccination campaign, having clear financing policies as part of disaster preparedness plans, investing in community preparedness capabilities, and safeguarding medical records.

Despite these commitments and needs assessments, the health systems of the three countries most affected by the 2013–2016 Ebola outbreak (Sierra Leone, Liberia, and Guinea) were weak and underfunded (Kieny, Evans, Schmets, & Kadandale, 2014). Most external funding was given to support HIV, malaria, TB, and maternal health, with insufficient investment in health system capacity. Their health systems' weaknesses had devastating consequences not just for the three countries' capacity to diagnose and treat infected individuals, but also to cope with routine work such as deliveries, malaria, and chronic conditions.

A group of experts, reflecting on the aftermath of the Ebola outbreak in West Africa, conceptualized health security as both collective security (epidemics crossing national borders) and individual security (access to safe and effective health services, products, and technologies) (Heymann et al., 2015). In particular, the authors pointed to the importance of addressing the quality of drugs, better allocation of resources, and management of migration, and called for global action in achieving universal health coverage. Health security and concepts of resilience are addressed in greater detail in the *Public Health Infrastructure* chapter.

Conclusion

This chapter has outlined the main components of health systems in LMICs. It has focused mainly on the healthcare system, with little attention (due to lack of space) given to the sectors that supply resources to the health system (such as education and training institutions, and the pharmaceutical, medical supplies, and medical equipment industries) or to broader influences on health, such as government policies on smoking, alcohol consumption, and transport—all of which have major impacts on health. Other chapters of this text provide complementary material that addresses those topics.

The key functions of health systems are governance and regulation, financing, resource allocation, and health services provision, and the key actors as the government, populations, financing agents, and providers. Although these core elements can be identified in all systems, the number of bodies involved in each one, and the way they relate to each other, differs enormously in practice, making it difficult to draw clear conclusions on whether any one arrangement of functions and actors is better than any other.

The key problems of health systems in LMICs can be summarized as follows:

- In many low-income countries, resources and funds are grossly inadequate to provide even a basic level of care for the population, with a number of sub-Saharan African governments spending less than $20 per capita on health services.
- Levels of health produced by health systems are often lower than what is technically possible.
- Many activities funded by the public sector are not cost-effective, and coverage of interventions that are highly cost-effective is inadequate (Jamison, 2015). Conversely, too high a share of the budget is spent on hospital care, especially higher levels of hospital care.
- Health systems operate with low or variable levels of technical efficiency. Studies of facility costs invariably show great variation in costs across similar types of services, to an extent that is not easily explained by differences in quality but is more likely to be due to problems of managing resources.
- Quality of services is poor in public facilities, especially in the poorer countries where funds are limited, health service inputs are in short supply, equipment is poorly maintained, and staff are poorly paid and hence poorly motivated; staff are often criticized for their lack of courtesy to

patients and lack of responsiveness to their needs. Many private services are also often of low quality and almost completely unregulated.
- The possibilities for health interventions created by technological developments place ever-increasing demands on limited funds, accentuating the need for governments to consider ways of setting priorities and defining limits on which services can be provided.
- New problems, such as the growing importance of chronic diseases, are putting even greater pressure on health systems.
- Numbers of healthcare workers are grossly inadequate in most countries, and their unequal distribution means that rural services have even worse staffing levels.
- Most LMICs fare poorly in terms of both equity of access to health services and equity of payment for services, with a few notable exceptions. Poorer households commonly use health services less frequently, especially in rural areas where access is more difficult and expensive, and spend a much higher proportion of their income on health services, partly because public subsidies do not meet their needs well.

In recent years, health systems have been characterized by the following trends:

- An emphasis on the regulatory and enabling role of the state
- An emphasis on increasing public control over sources of funding and increasing risk pooling
- More explicit prioritization of which services can be financed and provided, especially by the public sector and in terms of cost-effectiveness criteria
- A desire to improve targeting of public subsidies to those persons most in need of healthcare services
- A greater role of the private sector, especially in the provision of health services
- Greater decentralization of the management of provision within the public sector
- Creation of arrangements that encourage competition between providers, in an attempt to improve efficiency and quality
- Greater emphasis on the role of consumers and patients, as informed purchasers, in patient-centered care, and as citizens to whom providers should be accountable

A 2011 study by the OECD concluded that "there is no single type of healthcare system that performs systematically better in delivering cost-effective health care, as both market-based and more centralised command-and-control systems have strengths and weaknesses. It seems to be less the type of system that matters, but rather how it is managed" (p. 221). However, there is a tendency for ideology rather than good empirical evidence to drive reform policies. While values do and should play an important role in decisions relating to health systems, the relatively limited empirical evidence on how best to improve the performance of health systems makes it challenging to resolve ideological arguments by reference to evidence. Moreover, reforms affect the position of powerful interest groups, which are likely to mobilize to block change. Far greater understanding is needed of the best ways both to introduce reforms and manage these various interests.

Discussion Questions

1. How can the concepts of efficiency and equity be used to assess the performance of a health system?
2. Access to health care can be viewed as similar to access to other goods and services—that is, as dependent on an individual's success in gaining or inheriting income, or as a right of citizenship that should not depend on individual income or wealth. Debate the relative merits of these two positions.
3. What should be the respective roles of the government and the private sector in the health system of a LMIC of your choice?
4. What are the relative strengths and weaknesses of the main financing sources in a LMIC of your choice?
5. Outline the pros and cons of pay-for-performance mechanisms.

References

Adam, T., & De Savigny, D. (2012). Systems thinking for strengthening health systems in LMICs: Need for a paradigm shift. *Health Policy and Planning, 27*(suppl. 4), iv1–iv3.

Alhassan, R. K., Nketiah-Amponsah, E., & Arhinful, D. K. (2016). A review of the national health insurance scheme in Ghana: What are the sustainability threats and prospects? *PLoS One, 11*(11), e0165151.

Alkenbrack, S., Hanson, K., & Lindelow, M. (2015). Evasion of "mandatory" social health insurance for the formal sector: Evidence from Lao PDR. *BMC Health Services Research, 15*, 473.

Almeida, C., Braveman, P., Gold, M. R., Szwarcwald, C. L, Mendes Ribeiro, J., Miglionico, A., . . . Viacava, F. (2001). Methodological concerns and recommendations on policy

consequences of the World Health Report 2000. *The Lancet*, *357*(9269), 1692–1697.

Asante, A., Price, J., Hayen, A., Jan, S., & Wiseman, V. (2016). Equity in health care financing in low-and middle-income countries: A systematic review of evidence from studies using benefit and financing incidence analyses. *PLoS One*, *11*(4), p.e0152866.

Asia Pacific Observatory on Health Systems and Policies. (2015). The kingdom of Thailand. health system review. *Health Systems in Transition, 5*(5).

Backman, G., Hunt, P., Khosla, R., Jaramillo-Strouss, C., Fikre, B. M., Rumble, C., . . . Vladescu, C. (2008). Health systems and the right to health: An assessment of 194 countries. *The Lancet, 372*(9655), 2047–2085.

Balabanova, D., Mills, A., Conteh, L., Akkazieva, B., Banteyera, H., Dash, U., . . . McKee, M. (2013). Good health at low cost 25 years on: Lessons for the future of health systems strengthening. *The Lancet, 381*(9883), 2118–2133.

Baldwin, R., & Cave, M. (1999). *Understanding regulation: Theory, strategy, and practice.* Oxford, UK: Oxford University Press.

Basu, S., Andrews, J., Kishore, S., Panjabi, R., & Stuckler, D. (2012). Comparative performance of private and public healthcare systems in low- and middle-income countries: A systematic review. *PLoS Medicine, 9*(6), e1001244.

Bayntun, C., Rockenschaub, G., & Murray, V. (2012). Developing a health system approach to disaster management: A qualitative analysis of the core literature to complement the WHO Toolkit for assessing health-system capacity for crisis management. *PLoS Currents: Disasters, 1*(4), e5028b6037259a.

Beaglehole, R., Epping-Jordan, J-A., Patel V., Chopra, M., Ebrahim, S., Kidd, M., & Haines, A. (2008). Improving the prevention and management of chronic disease in low-income and middle-income countries: A priority for primary health care. *The Lancet, 372*(9642), 940–949.

Bennett, S. (1997). Health-care markets: Defining characteristics. In S. Bennett, B. McPake, & A. Mills (Eds.), *Private health providers in developing countries* (pp. 85–101). London, UK: Zed Books.

Berry, C., & Igboemeka, A. (2004). *Service provision in difficult environments: Issues arising from DFID support to health sector interventions in Burma, Afghanistan, and Nepal.* London, UK: Department for International Development.

Brinkerhoff, D. W. (2007). Introduction: Governance challenges in fragile states: Re-establishing security, rebuilding effectiveness, and reconstituting legitimacy. In D. W. Brinkerhoff (Ed.), *Governance in Post-conflict societies: Rebuilding fragile states* (pp. 1–22). London, UK: Routledge.

Brinkerhoff, D. W., & Bossert, T. J. (2008). *Health governance: Concepts, experience, and programming options, in health systems 20/20*. Washington, DC: USAID.

Certificate of Need State Laws. (2017). Retrieved from http://www.ncsl.org/research/health/con-certificate-of-need-state-laws.aspx

Cometto, G., Tulenko, K., Muula, A. S., & Krech, R. (2013). Health workforce brain drain: From denouncing the challenge to solving the problem. *PLoS Medicine, 10*(9), e1001514.

CONEVAL: Consejo Nacional de Evaluación de la Política de Desarrollo Social. (2012). La pobreza en la población indígena de México, 2012. Retrieved from http://www.coneval.org.mx/Medicion/Paginas/pobreza_indigena_2012.aspx

Countdown to 2015. (2015). *A decade of tracking progress for maternal, newborn and child survival.* New York, NY/Geneva, Switzerland: UNICEF & World Health Organization.

Coyne, S. C., & Hilsenrath, P. (2002). The World Health Report 2000: Can health care systems be compared using a single measure of performance? *American Journal of Public Health*, *92*(1), 30–33.

Department of Health and Social Security. (1976). *Sharing resources for health in England: Report of the Resource Allocation Working Party.* London, UK: Her Majesty's Stationery Office.

Doherty, J. E. (2015). Regulating the for-profit private health sector: Lessons from East and Southern Africa. *Health Policy and Planning, 30*(suppl. 1), i93–i102.

Donaldson, C., & Gerard, K. (2005). *Economics of health care financing: The visible hand.* London, UK: Macmillan Press.

Ensor, T., & Weinzierl, S. (2007). Regulating health care in low- and middle-income countries: Broadening the policy response in resource constrained environments. *Social Science & Medicine, 65*(2), 355–366.

Ensor, T., Witter, S., & Sheiman, I. (1997). Methods of payment to medical care providers. In S. Witter & T. Ensor (Eds.), *An introduction to health economics for eastern Europe and the former Soviet Union* (pp. 97–114). Chichester, UK: John Wiley & Sons.

Figueras. J., Robinson, R., & Jabowski, E. (2005). Purchasing to improve health systems performance: Drawing the lessons. In J. Figueras, R. Ronson, & E. Jakubowski (Eds.), *Purchasing to improve health systems performance* (pp. 44–80). Milton Keynes, UK: Open University Press.

Folland, S., Goodman, A. C., & Stano, M. (2013). *The economics of health and health care* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

FUNSALUD (Fundación Mexicana para la Salud). (2012). *Universalidad de los Servicios de Salud. Una propuesta de FUNSALUD.* Mexico, DF: Author.

Gaffney, A., & McCormick, D. (2017). The Affordable Care Act: Implications for health-care equity. *The Lancet, 389*(10077), 1442–1452.

Gilson, L. (2003). Trust and the development of health care as a social institution. *Social Science & Medicine, 56*(7), 1453–1468.

Gilson, L., Barasa, E., Nxumalo, N., Cleary, S., Goudge, J., Molyneux, S., . . . Lehmann, U. (2017). Everyday resilience in district health systems: Emerging insights from the front lines in Kenya and South Africa. *BMJ Global Health, 2*(2), e000224.

Gilson, L., Doherty, J., Loewenson, R., & Francis, V. (2007). *Challenging inequity through health systems. Final Report: Knowledge Network on Health Systems.* WHO Commission on Social Determinants of Health. Geneva, Switzerland: World Health Organization.

Gilson, L., & World Health Organization (WHO). (2013). *Health policy and system research: A methodology reader: The abridged version*. Geneva, Switzerland: World Health Organization.

Grollman, C., Arregoces, L., Martínez-Álvarez, M., Pitt, C., Mills, A., & Borghi, J. (2017). 11 years of tracking aid to reproductive, maternal, newborn, and child health: Estimates and analysis for 2003–13 from the Countdown to 2015. *The Lancet Global Health, 5*(1), e104–e114.

Gruskin, S., Ahmed, S., Bogecho, D., Ferguson, L., Hanefeld, J., MacCarthy, S., . . . Steiner R. (2012). Human rights in health systems frameworks: What is there, what is missing and why does it matter? *Global Public Health, 7*(4), 337–351.

Health Policy Project. (2016). *Health financing profile. Tanzania*. Washington, DC: Health Policy Project. Retrieved from https://www.healthpolicyproject.com/pubs/7887/Tanzania_HFP.pdf

Heymann, D. L., Chen, L., Takemi, K., Fidler, D. P., Tappero, J. W., Thomas, M. J., . . . Rannan-Eliya, R. P. (2015). Global health

security: The wider lessons from the west African Ebola virus disease epidemic. *The Lancet, 385*(9980), 1884–1901.

Index Mundi. (2017). Health services. Retrieved from https://www.indexmundi.com/facts/topics/health#Health%20services

Jamison, D. T. (2015, February 5). Disease control priorities: Improving health and reducing poverty. *The Lancet*, *391*(10125), e11–e14. doi: 10.1016/S0140-6736(15)60097-6

Jamison, D. T., Alwan, A., Mock, C. N., Nugent, R., Watkins, D., Adeyi, O., . . . Zhao, K. (2018). Universal health coverage and intersectoral action for health: Key messages from Disease Control Priorities, 3rd edition. *The Lancet 391*(10125):1108–1120.

Kieny, M.-P., Evans, D. B., Schmets, G., & Kadandale, S. (2014). Health-system resilience: Reflections on the Ebola crisis in western Africa. *Bulletin of the World Health Organization, 92*(12), 850.

Knaul, F. M., González-Pier, E., Gómez-Dantés, O., García-Junco, D., Arreola-Ornelas, H., Barraza-Lloréns, M., . . . Frenk, J. (2012). The quest for universal health coverage: Achieving social protection for all in Mexico. *The Lancet, 380*(9849), 1259–1279.

Kohn, R., & White, K. L. (1976). *Health care: An international study.* London, UK: Oxford University Press.

Kruk, M. E., Freedman, L. P., Anglin, G. A., & Waldman, R. J. (2010). Rebuilding health systems to improve health and promote statebuilding in post-conflict countries: A theoretical framework and research agenda. *Social Science & Medicine, 70*(1), 89–97.

Kruk, M. E., Myers, M., Varpilah, S. T., & Dahn, B. T. (2015). What is a resilient health system? Lessons from Ebola. *The Lancet, 385*(9980), 1910–1912.

Liu, X., Hotchkiss, D. R., & Bose, S. (2008). The effectiveness of contracting-out primary health care services in developing countries: A review of the evidence. *Health Policy and Planning, 23*(1), 1–13.

Liu, X., & Mills, A. (2002). Financing reforms of public health services in China: Lessons for other nations. *Social Science and Medicine*, *54*(11), 1691–1698.

Londoño, J. L., & Frenk, J. (1997). Structured pluralism: Towards an innovative model for health system reform in Latin America. *Health Policy, 41*(1), 1–36.

Lu, J.-F. R., & Chiang, T. L. (2011). Evolution of Taiwan's health care system. *Health Economics, Policy and Law, 6*(1), 85–107.

Ly, C., Eozenou, P., Nandakumar, A., Pablos-Mendez, A., Evans, T., & Adeyi, O. (2017). The economic transition of health in Africa: A call for progressive pragmatism to shape the future of health financing, *Health Systems & Reform, 3*(4), 290–300.

Martinez Álvarez, M. (2014). *Development assistance for health in Tanzania: Has the sector wide approach achieved the principles of aid effectiveness?* [PhD dissertation]. London, UK: London School of Hygiene & Tropical Medicine. Retrieved from http://researchonline.lshtm.ac.uk/2095790/

Martínez-Álvarez, M., Acharya, A., Arregoces, L., Brearley, L., Pitt, C., Grollman, C., & Borghi, J. (2017). Trends in The Alignment And Harmonization Of Reproductive, Maternal, Newborn, And Child Health Funding, 2008–13. *Health Affairs*, *36*(11), 1876-1886.

Martinez Álvarez, M., Borghi, J., Acharya, A., & Vassall, A. (2016). Is development assistance for health fungible? Findings from a mixed methods case study in Tanzania. *Social Science & Medicine, 159,* 161–169.

McIntyre, D., Meheus, F., & Rottingen, J.-A. (2017). What level of domestic government health expenditure should we aspire to for universal health coverage? *Health Economics, Policy and Law, 12*(2), 125–137.

McPake, B. (1997). The role of the private sector in health service provision. In S. Bennett, B. McPake, & A. Mills (Eds.), *Private health providers in developing countries* (pp. 21–39). London, UK: Zed Books.

McPake, B., & Machray, C. (1997). International comparisons of health sector reform: Towards a comparative framework for developing countries. *Journal of International Development, 9*(4), 621–629.

McPake, B., Normand, C., & Smith, S. (2013). *Health economics: An international perspective*. Abingdon, UK: Routledge.

Mills, A., Ataquba, J. E., Akazili, J., Borghi, J., Garshong, B., Makawia, S., . . . McIntyre, D. (2012). Equity in financing and use of health care in Ghana, South Africa, and Tanzania: Implications for paths to universal coverage. *The Lancet, 380*(9837), 126–133.

Mills, A., Bennett, S., Russell, S., Attanayake, N., Hongoro, C., Muraleedharan, V. R., & Smithson, P. (2001). *The challenge of health sector reform: What must governments do?* Basingstoke, UK: Macmillan.

Mills, A., & Hanson, K. (Eds.). (2003). Expanding access to health interventions in low and middle-income countries: Constraints and opportunities for scaling-up. *Journal of International Development, 15*(1), 1–131.

Montagu, D., & Goodman, C. (2016). Prohibit, constrain, encourage, or purchase: How should we engage with the private health-care sector? *The Lancet, 388*(10044), 613–621.

Mtei, G., Mulligan, J., Palmer, N., Kaumuzora, P., Ally, M., & Mills, A. (2007). *An assessment of the health financing system in Tanzania: Implications for equity and social health insurance.* Ifakara, Tanzania: Ifakara Health Research and Development Centre; Ministry of Health and Social Welfare.

O'Donnell, O., van Doorslaer, E., Rannan-Eliya, R. P., Somanathan, A., Adhikari, S. R., Akkazieva, B., . . . Zhao, Y. (2008). Who pays for health care in Asia? *Journal of Health Economics, 27*(2), 460–475.

Organization for Economic Cooperation and Development (OECD). (1992). Sub-systems of financing and delivery of health care. In *The reform of health care* (pp. 19–29). Paris, France: Author.

Organization for Economic Cooperation and Development (OECD). (2011). A new look at OECD health care systems: Typology, efficiency and policies. In *Economic policy reforms 2011. Going for growth sub-systems of financing and delivery of health care* (pp. 221–237). Paris, France: Author.

Organization for Economic Cooperation and Development (OECD). (2016). *OECD reviews of health systems: Mexico 2016.* Paris, France: OECD Publishing. Retrieved from http://www.oecd.org/publications/oecd-reviews-of-health-systems-mexico-2016-9789264230491-en.htm

Oxman, A. D., & Fretheim, A. (2009). Can paying for results help to achieve the Millennium Development Goals? Overview of the effectiveness of results-based financing. *Journal of Evidence-Based Medicine, 2*(2), 70–83.

Peters, D. H., Yazbeck, A. S., Sharma, R., Ramana, G. N. V., Pritchett, L., & Wagstaff, A. (2002). *Better health systems for India's poor: Findings, analysis, and options*. Washington, DC: World Bank.

Porter, M. E., & Kaplan, R. S. (2016). How to pay for health care. Retrieved from https://hbr.org/2016/07/how-to-pay-for-health-care

Porter, R. (1996a). Hospitals and surgery. In R. Porter (Ed.), *The Cambridge illustrated history of medicine* (pp. 202–205). Cambridge, UK: Cambridge University Press.

Porter, R. (1996b). Medical science. In R. Porter (Ed.), *The Cambridge illustrated history of medicine* (pp. 154–201). Cambridge, UK: Cambridge University Press.

Powell Jackson, T., Mazumdar, S., & Mills, A. (2015). Financial incentives in health: New evidence from India's Janani Suraksha Yojana. *Journal of Health Economics, 43,* 154–169.

RESYST. (2014). What is strategic purchasing for health? Retrieved from http://resyst.lshtm.ac.uk/sites/resyst.lshtm.ac.uk/files/docs/reseources/Purchasing%20brief.pdf

Roemer, M. I. (1991). *National health systems of the world: Vol. 1. The countries.* Oxford, UK: Oxford University Press.

Roemer, M. I. (1993). *National health systems of the world: Vol. 2. The issues.* Oxford, UK: Oxford University Press.

RSBY. (2017). About RSBY: Genesis of the RSBY. Retrieved from http://www.rsby.gov.in/about_rsby.aspx

Saltman, R. B., & Figueras, J. (1997). *European health care reform: Analysis of current strategies.* Copenhagen, Denmark: World Health Organization, Regional Office for Europe.

Sheikh, K., Saligram, P. S., & Hort, K. (2013). What explains regulatory failure? Analyzing the architecture of health care regulation in two Indian states. *Health Policy and Planning, 30*(1), 39–55.

Tangcharoensathien, V., Patcharanarumol, W., Ir, P., Aljunid, S. M., Mukti, A. G., Akkhavong, K., . . . Mills, A. (2011). Health-financing reforms in southeast Asia: Challenges in achieving universal coverage. *The Lancet, 377*(9768), 863–873.

Tantivess, S., Teerawattananon, Y., & Mills, A. (2009). Strengthening cost-effectiveness analysis in Thailand through the establishment of the Health Intervention and Technology Assessment Program. *PharmacoEconomics, 27*(11), 931–945.

Teerawattananon, Y., Tangcharoensathien, V., Tantivess, S., & Mills, A. (2003). Health sector regulation in Thailand: Recent progress and the future agenda. *Health Policy, 63*(3), 323–338.

Towse, A., Mills, A., & Tangcharoensathien, V. (2004). Learning from Thailand's health reforms. *British Medical Journal, 328*(7431), 103–105.

Travis, P., Bennett, S., Haines, A., Pang, T., Bhutta, Z., Hyder, A., . . . Evans, T. (2004). Overcoming health-systems constraints to achieve the Millennium Development Goals. *The Lancet, 364*(9437), 900–906.

Trujillo, A. J., & McCalla, D. C. (2004). Are Colombian sickness funds cream skimming enrollees? An analysis with suggestions for policy improvement. *Journal of Policy Analysis and Management, 23*(4), 873–888.

United Republic of Tanzania, Ministry of Health and Social Welfare. (2015). Health sector strategic plan, July 2015–June 2020. Retrieved from http://www.tzdpg.or.tz/fileadmin/documents/dpg_internal/dpg_working_groups_clusters/cluster_2/health/Key_Sector_Documents/Induction_Pack/Final_HSSP_IV_Vs1.0_260815.pdf

van Olmen, J., Marchal, B., Van Damme, W., Kegels, G., & Hill, P. S. (2012). Health systems frameworks in their political context: Framing divergent agendas. *BMC Public Health*, *12*(1), 774.

Wagstaff, A., van Doorslaer, E., van der Burg, H., Calogne, S., Christiansen, T., Citoni, G., . . . Hakinnen, U. (1999). Equity in the finance of health care: Some further international comparisons. *Journal of Health Economics, 18*(3), 263–290.

Walsh, J. A., & Warren, K. S. (1979). Selective public health care: An interim disease control strategy in developing countries. *New England Journal of Medicine, 301,* 967–974.

Walsh, K. (1995). *Public services and market mechanisms: Competition, contracting and the new public management.* London, UK: Macmillan Press.

World Bank. (2003). *World development report 2004: Making services work for poor people.* Oxford, UK: Oxford University Press.

World Bank. (2004). *The Millennium Development Goals for health: Rising to the challenges.* Washington, DC: Author.

World Bank. (2008). *World Bank: Global monitoring report.* Washington, DC: Author.

World Bank. (2012). *Health equity and financial protection report: Zambia.* Washington, DC: Author.

World Bank. (2017). Taxes on international trade. Retrieved from http://data.worldbank.org/indicator/GC.TAX.INTT.RV.ZS

World Bank Data. (2018a). GDP per capita (constant 2010 US$). Retrieved from https://data.worldbank.org/indicator/NY.GDP.PCAP.KD?view=chart

World Bank Data. (2018b). Mortality rate, under-5 (per 1,000 live births). Retrieved from https://data.worldbank.org/indicator/SH.DYN.MORT?view=chart

World Health Organization (WHO). (1999). *The world health report 1999: Making a difference.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2000). *The world health report 2000: Health systems—improving performance.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2007). *Everybody's business: Strengthening health systems to improve health outcomes: WHO's framework for action.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2009). WHO Statistics 2009, Health Expenditure. http://www.who.int/whosis/whostat/EN_WHS09_Table7.pdf

World Health Organization (WHO). (2010a). *Administrative costs of health insurance schemes: Exploring the reasons for their variability. HSS/HSF/DP.E.10.8.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2010b). *Health systems financing: The path to universal coverage.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2011). The Abuja Declaration: Ten years on. Retrieved from http://www.who.int/healthsystems/publications/abuja_declaration/en/

World Health Organization (WHO). (2017). Global Health Expenditure Database. Retrieved from http://apps.who.int/nha/database/Regional_Averages/Index/en

Zwi, A., & Mills, A. (1995). Health policy in less developed countries: Past trends and future directions. *Journal of International Development, 7*(3), 299–328.

CHAPTER 14

Public Health Infrastructure

James Whitworth, Siobhan Leir, Sakib Rokadiya, Osman Dar, and David L. Heymann

Introduction

Infrastructure is "the basic services of social capital of a country, or part of it, which make economic and social activities possible" (Rutherford, 1992). In the context of public health, infrastructure can be described as "formal and enduring structures that support public health, having both tangible and intangible aspects and existing inside and outside the government sector," comprising three main elements (Powles & Comim, 2003):

1. *Institutions* and institutional capacity (encompassing both state and nonstate actors) appropriate to predict, prevent, and respond to health problems:
   - Public health activities require legal and regulatory frameworks to enforce public health measures, coupled with the broader contribution of health-related civic organizations.
   - Public health institutions must have the capacity to monitor and respond to changing patterns and determinants of disease. For example, infectious disease surveillance requires appropriate laboratory facilities to identify pathogens.
2. *Knowledge*, as acquired and put to use both by the general population and by professionals:
   - The capacity of a country to take up and apply knowledge is a powerful predictor of mortality decline. Countries must devise systems to harness and effectively implement global technological and intellectual developments.
   - A capable and qualified health workforce is a crucial component of a strong public health infrastructure. The quality of vocational and research training in public health disciplines is a key determinant of staff competency. The ability of professionals to acquire and assimilate the latest knowledge and developments in public health—increasingly via the use of digital technologies—also determines their effectiveness.
   - A well-informed and well-educated general population brings health benefits by contributing to changes in behavior.
3. *Commodities:* Resources or physical infrastructure
   - For example, diagnostic tests, effective sewers, and safe water supplies.

Perhaps the most widely accepted definition of public health was provided by Donald Acheson in the 1988 report on Public Health in England: "Public health is the science and art of preventing disease, prolonging life and promoting health through organised efforts of society." This definition captures the hallmarks of modern public health theory and

practice—an emphasis on collective responsibility for health and the key role of the state, a population-level perspective, a concern for the underlying socioeconomic determinants of health and disease linked to an emphasis on prevention, and an underscoring of the notion of public health as a multidisciplinary subject (Beaglehole & Bonita, 2004). Successfully translating public health theory into practice requires the establishment of all three main elements of strong and reliable public health infrastructure as outlined previously.

A robust public health infrastructure provides part of the necessary foundation for a well-functioning health system. The World Health Organization (WHO, 2000) has defined health systems as encompassing all activities whose primary purpose is to promote, maintain, or restore health. This broad definition captures both formal *healthcare services* focused on personal healthcare and *public health capabilities* focused on promoting health and preventing sickness. The public health branch of a health system, which is the focus of this chapter, has been defined as "all public, private, and voluntary entities that contribute to the delivery of essential public health services within a jurisdiction" (Centers for Disease Control and Prevention [CDC], 2014). It is important to distinguish between the two components of a health system, yet one must also understand that both must work effectively and synergistically to give rise to a well-functioning health system. These two components are inextricably linked. To illustrate, in the case of an infectious disease outbreak, a healthcare system without the support of strong public health capabilities will lack the ability to monitor disease trends and predict impending threats and, therefore, will be unable to plan and mobilize the response required (Heymann et al., 2015). Conversely, a public health system without strong primary and secondary healthcare services lacks both the "radar screen" to pick up the initial cases and the delivery system to execute an effective response (Commission on a Global Health Risk Framework for the Future, 2016). Thus, both sets of services and capabilities are necessary to address public health threats. Indeed, healthcare systems require the flexibility to be scaled up in times of crisis. Meaningful investment in both health system branches is critical; however, in most countries, healthcare services—preventive, curative, and palliative interventions—account for the bulk of expenditure and activity (WHO, 2000). As discussed in the "Strengthening Public Health Systems" section of this chapter, it is generally agreed that there is a need to seek a better balance.

The WHO framework describing health systems in terms of six core components or "building blocks"—that is, service delivery, health workforce, health information systems, access to essential medicines, financing, and leadership/governance (WHO, 2010)—does not help ensure that public health receives the appropriate attention. Indeed, it obscures the dynamic links between each component. It also fails to acknowledge the differences between the healthcare and public health branches of a health system, suffers from the lack of a community focus, and does not address either the underlying social and economic determinants of health or the role of non-health-sector actors.

Adoption of the Sustainable Development Goals (SDGs) by the United Nations in 2015 presented an opportunity to develop an ambitious agenda for global health system reform grounded in intersectoral cooperation, resilience, human rights, and country ownership (Schäferhoff, Suzuki, Angelides, & Hoffman, 2015). The SDGs place people at the center, and one goal (SDG 3) relates explicitly to health: To "ensure healthy lives and promote wellbeing for all at all ages" (United Nations, 2015). Recognizing the broad determinants of health, attainment of the health targets contained within SDG 3 is dependent not only on actions within the health sector, but also on economic, social, educational, and environmental factors. The field of public health practice is well placed to address these complex, multidisciplinary challenges. Indeed, it can be argued that each of the 17 SDGs has significant public health implications; thus, the SDGs can be interpreted as an agenda for public health to create the conditions in which people can be healthy (Bellagio District Public Health Workshop Participants, 2016; Fried, 2015).

Chapter Outline

In this chapter, we take a systems approach first to define essential public health functions, and then to review current practice and initiatives to strengthen the public health capabilities of health systems in low- and middle-income countries (LMICs). This discussion includes universal health coverage; global, regional, and community-based initiatives; laboratory capacity; human resources; and consideration of responding to disasters and emergencies. We then examine the International Health Regulations, the Global Health Security Agenda, and Joint External Evaluations. These derive from a public health perspective of *health security*, whereas the early sections of the chapter assume a *human rights* perspective. Most of the new funding for public health in recent years has been focused on these health security initiatives, and this is where much

of the progress in strengthening public health infrastructure has been seen.

Essential Public Health Functions

Public health requires effective action by many different organizations and players at all levels of governance, whether local, national, or international. While the balance of responsibility will vary by context, the basic functions or services that must be assured in every locality remain broadly constant. The "10 Essential Public Health Functions," first described by the CDC in 1994, serve as benchmarks for the activities that all public health agencies should undertake, and are organized under the three core functions of public health: assessment, policy development, and assurance (CDC, 2014). These functions, which are listed in **EXHIBIT 14-1**, are discussed in more detail in the remainder of this section. The consensus is that countries and regions should adapt this list to produce a set of context-specific essential public health functions based on local stakeholder input (Bettcher, Sapirie, & Goon, 1998; Bishai et al., 2016).

It is imperative that public health services monitor and measure the health status of the community they serve. An accurate and periodic assessment of the population's health through surveillance of disease and risk factor trends underpins the promotion of health and prevention of sickness. This may involve the use of a variety of methods and technologies to collect, store, interpret, and communicate data. Effective monitoring and analysis of trends allow for the timely identification and investigation of potential health threats. Further, a functioning health information system allows for future health needs to be anticipated, thereby enabling preventive measures to be implemented. The ability to respond effectively to an emerging health problem rests upon the availability of quality diagnostic services, including laboratory capacity, and appropriate action plans to address major health threats. Mechanisms should also be in place to disseminate raw data and research results that may have value in combating the health threat. This sharing of information facilitates a coordinated response and helps to ensure that any actions taken are informed by the evidence.

Public health activities cannot function effectively without community participation and buy-in from an informed, educated, and empowered population. A key component of the public health infrastructure, this support may be established through the implementation of initiatives to build knowledge and skills and by promoting the adoption of behaviors associated with favorable health outcomes. It is also necessary to stimulate the community to take appropriate actions to address health issues and ensure conditions present in the community are conducive to good health. Communities can be engaged in this way through the mobilization of formal or informal community partnerships to support public health. Education at the individual and community levels is critical, as many public health problems are associated with human behaviors and lifestyles.

EXHIBIT 14-1 The 10 Essential Public Health Functions

Assessment

1. Monitor health status to identify and solve community health problems.
2. Diagnose and investigate health problems and health hazards in the community.

Policy Development

3. Inform, educate, and empower people about health issues.
4. Mobilize community partnerships and action to identify and solve health problems.
5. Develop policies and plans that support individual and community health efforts.

Assurance

6. Enforce laws and regulations that protect health and ensure safety.
7. Link people to needed personal health services and assure the provision of healthcare when otherwise unavailable.
8. Assure a competent public and personal healthcare workforce.
9. Evaluate the effectiveness, accessibility, and quality of personal and population-based health services.
10. Conduct research for new insights and innovative solutions to health problems.

Reproduced from Centers for Disease Control and Prevention. (1994). *The public health system and the 10 essential public health services.* Retrieved from https://www.cdc.gov/stltpublichealth/publichealthservices/essentialhealthservices.html

The formulation and enforcement of sound plans, policies, laws, and regulations are essential to support public health efforts, guide public health practice, and protect health. For example, these may include emergency response plans, policies mandating the reporting of highly transmissible diseases, or the regulation of environmental hazards. The review, evaluation, and revision of current regulations and laws should be conducted regularly as appropriate. Compliance efforts and enforcements should also be supported as needed. It is important to be aware that influencing politics, and thereby policy making, is an essential function of public health.

Actively linking people to health services and assuring the provision of healthcare when it might otherwise be unavailable or difficult to access are paramount given the health inequalities evident within every society. Achieving this goal requires identifying barriers to care, facilitating entry into a coordinated system of clinical care, assuring ongoing care management, and providing appropriate and targeted health information for at-risk population groups. As health disparities, particularly with regard to communicable diseases, can affect society as a whole, there is a pragmatic as well as ethical rationale for addressing health inequalities at national and international levels.

Central to the operation of an effective public health system is a competent public and personal health workforce. To assure such a workforce is available, health workers must be assessed regularly, standards maintained (e.g., through licensing requirements), and continual education provided, including supporting the development of leaders with the vision essential to ensure the health of society and the implementation of innovative public health measures. Linked to this element is the need for ongoing evaluation of the effectiveness, accessibility, and quality of health services.

Research is critical to the advancement of public health, as this is the way in which we refine current interventions and practices, and identify, monitor, and support the development of innovative solutions, incorporating cutting-edge technologies to address health issues. Given the explosive growth in the volume and complexity of data generated by health and other systems, new strategies and methods will be required to manage and analyze the massive amounts of data to glean insights of public health importance. It is therefore essential that public health practitioners continue to push the frontiers of science and technology to develop sophisticated research strategies. Subsequently, it is critical to ensure that such research knowledge is disseminated to and applied by public health practitioners and policy makers. Facilitating linkages between researchers and public health practitioners and policy makers is important to ensure the timely translation of research advances into health policy and practice.

It is worth noting that items 7, 8, and 9 in Exhibit 14-1 all refer to a healthcare system, recognizing the aforementioned interlinkage between public health activities and healthcare services. All 10 items link directly to, and are dependent on, the existence of an effective public health infrastructure.

Strengthening Public Health Systems

Currently, healthcare systems (medical services focused on treating the sick) are given priority over public health systems (services responsible for protecting people from becoming sick) all over the world (WHO, 2000). In other words, treatment tends to crowd out prevention. However, there is a broad consensus among public health experts that improving and maintaining population health cannot be achieved through a health system that focuses predominately on the sick. While a sickness-based system is necessary, it is not sufficient to identify and treat cases of high-burden conditions such as coronary heart disease, tuberculosis, and cancer. To anticipate evolving health threats and organize a timely response, human resources must be devoted to the monitoring and analysis of epidemiological data. In this sense, the public health system acts as the first line of defense against health crises, such as outbreaks of infectious diseases.

The Economic Case

Despite its crucial role, public health remains acutely under-resourced in most parts of the world. Being adequately prepared for future health threats has significant economic benefits as well as the obvious health benefits. To illustrate, the Commission on a Global Health Risk Framework for the Future (2016) estimates that an extra $4.5 billion per year is required to strengthen global and national preparedness and response efforts just to combat epidemic diseases. This figure includes expenditures for strengthening national public health systems, financing global coordination and contingency efforts, and funding for research and development.

To put this figure into perspective, the World Bank has estimated the cost of a severe pandemic (i.e., one on the scale of the 1918 influenza pandemic, which killed between 50 million and 100 million people in a

global population of less than 2 billion) at more than $3 trillion, or 4.8% of global gross domestic product (GDP) (Jonas, 2014). Even this figure could be an underestimate: Aggregate GDP losses for Sierra Leone, Guinea, and Liberia in 2014 and 2015 as a result of the Ebola epidemic are estimated at more than 10% of those countries' GDP (United Nations Development Group, 2015; World Bank, 2014). Indeed, scenario modeling conducted by the Commission on a Global Health Risk Framework for the Future suggests that during the 21st century, global pandemics could cost in excess of $6 trillion, with an estimated loss of more than $60 billion per year.

The economic impact of infectious disease outbreaks is expected to increase in tandem with greater human and economic connectivity, as transnational supply chains, increased travel, and access to communication technologies facilitate the spread of the pathogen itself, and of fear (Commission on a Global Health Risk Framework for the Future, 2016). The latter point is particularly important: Most of the economic impact of outbreaks stems not from mortality or morbidity, but rather from behavioral changes driven by fear of infection (Burns, van der Mensbrugghe, & Timmer, 2008). In turn, this fear is driven by a mix of heightened awareness and ignorance. For example, the 2015 outbreak of Middle East respiratory syndrome (MERS) in South Korea caused widespread panic and resulted in a substantial change in consumer behavior (World Bank, 2017a). The leisure and entertainment sectors were particularly hard hit: Cinema visits dropped by 52% on a year-on-year basis in the first two weeks of June, and tourism fell by 41%. Indeed, in June 2015, the consumer sentiment index compiled by the Bank of Korea dropped below the neutral 100 mark for the first time since 2012, signifying a deteriorating economic outlook. This highlights the need to strengthen public health systems: A healthcare system alone is not configured to address these social issues, whereas the public health system is well positioned and equipped to allay public concerns and address any misconceptions.

Estimating the economic costs of an outbreak and the savings accrued by investing in strong public health systems is fraught with challenges and uncertainties. However, investing in public health systems for outbreak preparedness is hardly wasted money: Strengthening the core foundations of public health and primary care systems are crucial to achieving many other health objectives (Commission on a Global Health Risk Framework for the Future, 2016). For example, a robust surveillance system and improved laboratory diagnostic capabilities will help health systems tackle endemic diseases such as malaria and tuberculosis, as well as noncommunicable diseases. Underinvestment in these core capabilities, in contrast, puts a country at perpetual risk of disaster and, because infectious diseases do not respect national borders, leaves everyone in the world vulnerable. There is therefore a strong economic case for investing in public health infrastructure at all levels of governance.

Universal Health Coverage

One of the targets contained in Goal 3 of the SDGs is to achieve universal health coverage (UHC) (United Nations, 2015). UHC is a broad concept that can be most simply defined as all people receiving the essential health services they need, without suffering financial hardship (WHO, 2013). Today, an estimated 400 million people lack access even to essential health services (WHO & World Bank, 2015). Every year, approximately 100 million people—more than 60% of whom reside in India alone (Berman, Ahuja, & Bhandari, 2010)—are forced into poverty as a result of healthcare-related fees, with a substantial number choosing to avoid seeking medical advice entirely because they do not possess the financial means to do so (Dar, Buckley, Rokadiva, Huda, & Abrahams, 2014; Mills, 2014). It is estimated that investments to expand services toward UHC could prevent 97 million premature deaths globally between now and 2030, and add as much as 8.4 years to life expectancies in some countries (WHO, 2017d). These health gains tend to be greater in less developed countries, and accrue predominantly to the poorest population groups.

Along with these health gains, the achievement of UHC can bring significant economic and political benefits. Politicians have recognized the role of UHC in maintaining social order and strengthening national security by improving the social solidarity of the population (Heymann et al., 2015; Nicholson, Yates, Warburton, & Fontana, 2015). Uncertainties associated with the financial consequences of poor health and inequitable access to care represent significant welfare losses to populations, sometimes giving rise to civil unrest. For example, the primary motivation for the Chinese public health reforms undertaken in 2009 was to ensure "a harmonious society" (Cheng, 2008).

On a practical level, UHC can be achieved only within the context of a strong health system. Thus, the pathway to UHC necessarily requires a fundamental strengthening of a country's health system, including the development of a skilled and motivated healthcare workforce and a supply of, and access to, medicines and equipment (Nicholson et al., 2015).

UHC is integral to the achievement of the goals of global health in general. Given the interconnectedness of the world today, the availability of accessible and universal health services in all countries is a crucial first line of defense against global health threats (Heymann et al., 2015). The emergence of many infectious diseases of international concern (e.g., Ebola and artemisinin-resistant malaria) have occurred in settings without UHC and, therefore, without health systems that are able to effectively perform the essential public health functions or deliver medical services. Given the weakness of their preexisting health infrastructure, such countries are ill prepared to cope with public health emergencies, which fuels the spread of disease and its resulting death toll. Inadequate health coverage also encourages people to travel across borders to seek care in better-serviced countries, thereby increasing the potential for infection to spread and posing social and economic burdens in recipient nations.

UHC is also integral to instilling community trust in health systems and subsequently influencing health-seeking behavior. If, for example, communities cannot access quality health services such as diabetes care or maternal services in normal, non-emergency situations, that failure erodes their faith and trust in the same systems when health authorities attempt to manage emerging health threats, such as Ebola. Ultimately, without the provision of horizontally integrated primary care—a key component of UHC—the health security threat of both communicable and noncommunicable diseases cannot be adequately mitigated.

UHC may also play a key role in the tackling of specific health issues—for example, antimicrobial resistance (AMR), which represents a growing global public health threat. Evidence has shown that out-of-pocket (OOP) health expenditures are strongly correlated with AMR in LMICs (Alsan et al., 2015). This association was found to be driven by countries that require copayments for antimicrobials in the public sector. It is suggested that this policy shifts demand from the public sector to the less well-regulated private sector, where supply-side incentives to overprescribe are greater and quality assurance less standardized. Such practices facilitate the development of AMR through overconsumption, inappropriate dosing, shortened therapy courses, and use of substandard or counterfeit drugs. Therefore, progress toward UHC—a key tenet of which is the implementation of robust financing structures to reduce OOP expenditures—may help to control AMR in LMICs. The *Pharmaceuticals* chapter covers AMR issues in greater depth.

Regional Initiatives to Strengthen Public Health Practice

Over the past two decades, WHO regional offices have led the majority of public health performance improvement efforts (Bellagio District Public Health Workshop Participants, 2016). Regions are generally characterized by countries with broadly similar epidemiologic features and health systems organization, and often share a common language, culture, and level of economic development. Consequently, it often makes sense to pool resources and expertise to design and implement regional-level health system strengthening initiatives.

Such initiatives have a number of advantages over global or national-level initiatives. First, through the development of collaborating centers and establishment of local task forces, initiatives can be tailored to regional priorities and, therefore, made more relevant than global initiatives. WHO regional offices also have the ability to convene international expertise and collaboration through regional international partnerships such as the Economic Community of West African States (ECOWAS) or the South Asian Association for Regional Cooperation (SAARC). This facilitates coordination and collaboration among the countries of that region, which is particularly important in the control of infectious diseases and for smaller or island nations that are unable to build sufficient capacity independently (World Bank, 2017a). The sharing of information, expertise, and specialist facilities is both more cost-effective (as it is less expensive to develop high-level expertise at the regional multicountry level rather than at the individual-country level) and strengthens preparedness and response activities. Examples of regional initiatives for preparedness include the Mekong Basin Disease Surveillance network and the Southern Africa Center for Infectious Disease Surveillance.

The development of regional partnerships also helps to build capacity in other areas. For example, Thailand is a regional leader in the field of workforce development, and delivers a number of training programs to its fellow Southeast Asian countries. Since 1998, Thailand has offered an International Field Epidemiology Training Program (FETP) to trainees from neighboring countries, and has facilitated the expansion of national FETPs throughout the region (Field Epidemiology Training Program, 2017).

In addition, regional technical meetings and conferences are regularly convened to share knowledge and experiences. Moreover, with regard to WHO operations, by implementing initiatives through regional offices and

by deploying local technical officers, member states may have more confidence in requesting technical assistance and may be more likely to take up the recommendations proposed (Bellagio District Public Health Workshop Participants, 2016).

For sustainability, it is critical to build institutional support for such regional initiatives and to ensure that regional platforms are well resourced—factors that have hampered past efforts (Bellagio District Public Health Workshop Participants, 2016; World Bank, 2017a). In addition, it is useful to promote collaboration and cooperation between regions. Many such collaborations are already in place; for example, efforts at public health practice improvement in Europe and the Americas have informed efforts to measure and improve practices in the Eastern Mediterranean (WHO Eastern Mediterranean Regional Office, 2017). However, cross-regional partnerships (and intraregional partnerships) could be facilitated with greater investment and support from global organizations such as WHO and the International Association of National Public Health Institutes (IANPHI) (Bellagio District Public Health Workshop Participants, 2016). This is particularly necessary in cases where political rivalries complicate or hamper collaborations between neighboring countries.

Strengthening Public Health Systems at the Community or District Level

Traditionally, centrally controlled national health systems and top-down vertical health programs have dominated the healthcare landscape. Consequently, efforts to strengthen public health systems have focused on central levels of government. More recently, the need to develop strong public health systems at the district or community level has been increasingly recognized.

While it is difficult to accurately measure outcomes associated with increased community engagement, many benefits from strengthening public health systems at the local level have been observed. By building community capacity and rural social capital, community participation enables more locally responsive health services to be developed, leading to improvements in health services quality, knowledge, and health-related outcomes, as well as providing local employment opportunities (Kenny et al., 2013; Kulig, 2011). Further, by marshaling local resources and local political will to address the root causes of poor population health, and by offering community members input and ownership of the decisions regarding their health, horizontally integrated public health programs often prove more resilient and sustainable than corresponding vertical initiatives (Alford, 2009; Sherry, Ghaffar, & Bishai, 2017). Increased resilience and sustainability of local health systems, together with a corresponding growth in rural economies and improved fiscal responsibility, is in the interests of the national government as well as the community itself (Alford, 2009).

Community perceptions, behavior, and compliance with control measures are key parameters in effective disease control (Heymann et al., 2015). The 2014–2016 Ebola epidemic clearly highlighted the importance of community engagement and empowerment in combating disease outbreaks, and community resistance to control efforts is considered to be a key factor responsible for fueling the spread of the virus (WHO, 2015c). For example, patients were often hidden from the authorities in their homes, contact tracing was complicated by lack of compliance, and strikes by local hospital staff and burial teams further impeded control efforts. This resistance stemmed, in part, from a lack of trust in the health authorities (as discussed later in Exhibit 14-3), coupled with inadequately resourced local health workers and a lack of consideration given to cultural beliefs and behavioral practices. Further, the management of the outbreak was largely taken out of the hands of the affected communities, with local faith- and community-based organizations—which often have a high reserve of trust and are able to address issues in a culturally sensitive manner—generally excluded from control efforts (Kennedy & Nisbett, 2015).

District health systems, which often serve as the first point of contact for patients, play a key role in infectious disease control. District-based surveillance systems that can rapidly detect an infectious disease outbreak at its source can make a critical difference in determining the course and scale of an outbreak (Bellagio District Public Health Workshop Participants, 2016). For example, patient zero of the Nigerian Ebola outbreak in 2015 was diagnosed within 24 hours of his entry at Lagos airport, and, serendipitously, the local health department was able to reassign a team of contact tracers that was already in place for the polio eradication program (Bellagio District Public Health Workshop Participants, 2016). As a result of this prompt response at the district level (and also an element of good luck), Nigeria's Ebola outbreak was quickly contained.

Another advantage associated with strengthening district- or community-level public health systems is, simply, due to the sheer number and variety of districts and communities in place in a country. In a

failed state, national-level strengthening efforts can be slow moving and ineffective. Nevertheless, successful communities may be found within failed states; thus, one may be able to accomplish at the subnational level what cannot be done at the national level (Sherry et al., 2017).

District-level public health systems face four key challenges (Bellagio District Public Health Workshop Participants, 2016):

- Lack of continuity of leadership
- Lack of visibility of public health work
- Lack of multisectoral involvement
- Lack of public health governance

A lack of continuity in leadership (both administrative and technical) is characteristic of centralized health systems or countries where political considerations dominate the health landscape. Central decision makers can transfer key district-level personnel with little or no consideration of the local situation and the interruptions to services that might result. Therefore, district-level health systems may be strengthened by decentralizing control, such that local public health units are managed and financed by an authority at the subnational level—for example, out of a governor's or mayor's office or by a local board of health. The central government may still provide guidance and technical support in terms of policies and strategic objectives, as well as supplementary funding for priority programs, but the basic operations should be under local control to assure continuity of leadership. Further, when public health units are part of the local governance structure, collaboration with local stakeholders is more easily facilitated.

To address the lack of multisectoral involvement in district-level health systems, an intersectoral local public health committee should be created (Bellagio District Public Health Workshop Participants, 2016). This provides a platform to link various sectors together, and facilitate collaborations in the planning and implementation of public health programs. Sectors involved may include agriculture, education, transportation services, and water boards, for example. Additionally, public health system strengthening efforts can be facilitated by linking public health practitioners to a community of practice, enabling the exchange of ideas and knowledge (Bellagio District Public Health Workshop Participants, 2016).

Community Public Health Platforms

Community platforms facilitate partnerships among health professionals, politicians, and community members, enabling information to be shared and local resources convened to support public health programs. The engagement of community members in health initiatives ensures that programs and data collection are aligned with the community's needs, that data are shared with the community to guide actions, and that such actions marshal all of a community's capital and human resources as well as public revenue. Community public health platforms have been defined as "the partnerships formed in order to assess and assure population health" (Sherry et al., 2017).

The ability of a community platform to improve health outcomes depends on a number of factors: engagement in power sharing, formation and maintenance of collaborative partnerships, deployment of multicultural- and gender-sensitive health workers for services delivery, incorporation of the voice and agency of beneficiary communities in research protocols, and facilitation of bidirectional learning (Cyril, Smith, Possamai-Inesedy, & Renzaho, 2015). Five key features integral to the development of effective community platforms have been identified (Sherry et al., 2017):

- Level of community engagement
- Health system context and role of the government
- Breadth of intersectoral partnerships
- Sustainability
- Leadership and platform structure that promotes integration across all partners

These, in turn, rely on access to data about health problems and health threats, the means and will to share data and control with community members, balance between delivering clinical services and preventing disease in whole populations, and advocacy to maintain community engagement against pressure to consolidate control (Sherry et al., 2017). **EXHIBIT 14-2** describes the environment and structures for community platforms in low-functioning (Haiti) and high-functioning (Indonesia) scenarios.

Laboratory Services and the National Reference Laboratory

A credible and accessible laboratory service (comprising both public health and clinical laboratories) capable of producing timely, reliable, and accurate results is a cornerstone of any country's health services capacity. Under the International Health Regulations, countries are required to develop the capacity to detect, investigate, and report to the international community public health emergencies, such as disease outbreaks (WHO, 2016b).

At present, the laboratory infrastructure of many LMICs remains weak, characterized by limited quality

EXHIBIT 14-2 The Functioning of Community Public Health Platforms in Haiti and Indonesia

Haiti

Haiti has suffered from decades of political, economic, and social instability, exacerbated by frequent natural disasters and disease outbreaks. Haiti remains the poorest country in the Americas, and is one of the poorest and most socioeconomically unequal countries in the world (World Bank, 2017d). It is dependent on external development assistance, with nongovernmental organizations (NGOs), faith-based organizations, and informal providers—all with varying agendas—dominating the healthcare arena. To illustrate the scale of external involvement, *prior* to the 2010 earthquake, an estimated 8,000 to 9,000 NGOs were working in the country (Zanotti, 2010). Health services are delivered in a fragmented and unregulated manner, with the Ministry of Public Health and Population relegated to a marginal role (Hill, Pavignani, Michael, Murru, & Beesley, 2014). The weakness of the state and reliance on development partners has created an environment in which external entities often drive priority setting and resource allocation (Sherry et al., 2017). Without effective government leadership, coordination among these entities and between sectors is poor. Lack of service integration and cooperation has led to further fragmentation and duplication of efforts (Hill et al., 2014). This has, in turn, made it difficult for communities both to engage meaningfully in the public health process and to hold the government accountable for health services quality. As such, communities often favor traditional medicine over "formal" healthcare providers (Hill et al., 2014).

Given that NGOs provide the majority of health-related services, a significant challenge for Haiti is the quest to create sustainable solutions. Addressing community needs requires a flexibility in agenda setting that not all NGOs possess. Nevertheless, several NGOs have been successful in building capacity in the areas of development, health, and education (Sherry et al., 2017). These NGOs had several factors in common (Zanotti, 2010):

- Local, Haitian origins
- Connected to a diverse international network of donors and not accountable to a single agency
- Focused on addressing local needs and the needs of the most vulnerable
- Sharing a vision that tied economy, politics, and human rights

For example, Partners in Health created an agricultural arm in 2002—Partners in Agriculture (Zanmi Agrikol [ZA])—which aims to tackle malnutrition (Zanotti, 2010). Along with providing emergency food supplies, ZA teaches Haitians agricultural best practices and provides farmers with tools and seed to cultivate their land. It also buys seeds back from them, thereby supplying families with the money needed to maintain a sustainable livelihood. In turn, excess seeds are saved and donated to other farmers for the following sowing season. This has created a positive expansion of access to locally produced food.

This case study illustrates how, in the absence of a strong and stable government, NGOs and donor agencies can promote (or hinder) the development of successful community platforms for health.

Indonesia

In 2005, the Lumajang district health office in Indonesia led the launch of the Gerbangmas movement—a platform for communities, public health authorities, and other government sectors to work collaboratively to achieve 21 indicators of concern. Only about one-third of these are traditional health indicators; the remainder address determinants of health such as poverty reduction, literacy, and waste management. This multisectoral movement has resulted in improvements in all indicators, and has proved to be sustainable even in the context of changing economic and political landscapes (Blas, Sommerfeld, & Kurup, 2011; Commission on Social Determinants of Health, 2008).

The development of the Gerbangmas movement stemmed from decentralization of the Indonesian health system, which allowed for peripheral innovation (Sherry et al., 2017). The district health office acted as the initial champion for the movement. It provided technical and financial support and coordination during the initial scale-up, through which stakeholders could collaborate around common goals, without competing for resources. As the initiative evolved, the district office relinquished leadership to an NGO.

The partnership structure of the initiative provided clear roles for each sector, with the creation of a common pool of funds that communities were able to draw on to invest in interventions in multiple sectors (e.g., the purchase of iodinated salt, food supplements for children younger than the age of 5 years, or family planning educational materials) (Blas et al., 2011). With the community at the center of the partnership (with power over program management and priority setting), a structure in which all sectors stood on an equal footing, and a common set of objectives to strive for, the Gerbangmas movement facilitated effective collaborations between sectors (Sherry et al., 2017). The district governor mandated that all community empowerment programs use Gerbangmas as an entry point, thereby reducing competition and preventing outside interests from affecting the success of the partnerships across sectors (Blas et al., 2011).

(continues)

EXHIBIT 14-2 The Functioning of Community Public Health Platforms in Haiti and Indonesia *(continued)*

The design of the Gerbangmas movement also ensured its sustainability. With resources, funding, and training provided by partnering organizations, the initiative did not depend on grant funds or external donor assistance. Further, the partnership structure improved the resilience of the platform, as it reduced reliance on any single organization or leader. The greatest hurdle to sustainability was the turnover of government officials, with the future of the initiative resting on each new district governor's approval in the subsequent 5-year plan. An approach to address this issue was to include the policy of Gerbangmas in the District Regulation, which was ratified by the District Legislative (Blas et al., 2011).

This case study illustrates that to establish a successful community platform that addresses the multiple social determinants of health, it is important to consider nonsectoral mechanisms that can accommodate multisectoral interests.

assurance and control protocols, insufficient funds, and a lack of standardized methods, suitably trained staff, and laboratory supplies. In many countries, the absence of networking mechanisms to link and coordinate laboratories has hampered data-sharing efforts and resulted in disconnected and unsupervised peripheral laboratories of dubious quality, which have negative implications for disease surveillance (Ndihokubwayo, Kasolo, Yahaya, & Mwenda, 2010).

A strong laboratory system is underpinned by a comprehensive national laboratory policy covering laboratory organization, structure, and coordination; staff motivation and retention (including training requirements and career development); integration of services; essential facilities, equipment, and maintenance; and biosafety and biosecurity (Ndihokubwayo et al., 2010). A national strategic plan should then be formulated to implement the national laboratory policy.

A national reference laboratory sits at the heart of an effective laboratory system, acting as a linchpin for surveillance, detection, response, and accurate diagnosis to guide treatment and prevention. This reference laboratory should set laboratory norms and standards, and participate in public health research and policy making. However, despite the threat from emerging and reemerging diseases, few laboratories in LMICs have capabilities for diagnosing highly infectious diseases. Indeed, many LMICs must ship specimens to other countries for confirmation, contributing to delays in suitable patient management and response to outbreaks (Ndihokubwayo et al., 2010). Thus, as part of the international public health infrastructure, it is critical to establish sufficiently staffed and well-equipped national public health reference laboratories that can operate as centers of excellence for laboratory services.

In addition to providing diagnostic services, the national reference laboratory should take responsibility for producing and distributing quality assurance programs to identify and address gaps in the quality of subnational laboratories. Engagement in quality assessment schemes could be linked to annual laboratory registration and renewal processes to incentivize participation. Further, to facilitate communication with and coordination of the national laboratory infrastructure, and to enable timely data sharing for purposes of surveillance efforts, laboratories at all levels of the health system should be linked through a network led by the central reference laboratory. To facilitate regional surveillance efforts, the national reference laboratories should also be connected and coordinated through a regional laboratory network.

The newly established Africa Centers for Disease Control (Africa CDC) has advocated that each African country should establish a national reference laboratory that can then be linked to form regional networks. In a first step toward realizing this goal, existing laboratories, including private laboratories, are being mapped to gauge how they might contribute to these networks. Recognizing that individual countries may lack the technological capabilities required to detect emerging threats, Africa CDC aims to establish five Regional Collaborating Centers, each equipped with laboratory and advanced diagnostic capacity to rapidly detect known and unknown pathogens (Nkengasong, 2017). These national and regional laboratory networks should also connect public health and clinical laboratories with their animal health counterparts under the One Health approach.

There is also a need to strengthen the laboratory workforce at all levels, including ensuring that technicians are trained in the operation and preventive maintenance of laboratory equipment. Unfortunately, many LMICs have the capacity to provide only basic laboratory training, greatly limiting the level of technology available on an in-country basis. To fill this gap, a number of initiatives have been set up to support the training of laboratory personnel in LMICs. For example, the Pacific Paramedical Training Centre (PPTC) based at Wellington Hospital, New Zealand, was established to support laboratories in Pacific Island countries through training and development assistance (PPTC, 2017). Since 2006, the PPTC and WHO have jointly provided distance-learning courses, including an online diploma in medical laboratory

science for Pacific laboratory workers. This course is delivered through the Pacific Open Learning Health Net (POLHN), an e-learning platform established in 2003 to deliver online learning and continuing professional development for geographically dispersed and remote health workers in the Pacific (POLHN, 2017). Currently, the platform links and provides thousands of free courses to approximately 30,000 health workers from 15 Pacific island countries. The PPTC also provides a Regional External Quality Assessment (REQA) program to help laboratories identify areas of weakness and provide solutions to improve the quality of service. Training programs such as these, and the availability of other opportunities for career development, also act to keep laboratory staff motivated, so they help LMICs retain qualified personnel in their national systems.

Another strategy to improve the delivery of quality laboratory services involves establishing or strengthening monitoring and evaluation systems with targets and measurable indicators. Monitoring and evaluation efforts should take into account activities such as adherence to standard operating procedures and safety guidelines, quality assessment activities, laboratory performance and workload, and utilization of supervisory tools.

These capacity-building activities require adequate and sustainable funding. Public health laboratory services should be funded through government budgetary provision. In LMICs, additional funding may be required through donor partners and income-generating activities. The integration of public health laboratory programs through national and regional networks can also help to alleviate financial issues through the sharing and optimal use of available resources (Ndihokubwayo et al., 2010).

Diagnostic Capability

Capacity to detect known pathogens, alongside rapid sharing of international knowledge, is key for detection of novel disease. For newly emerging pathogens, there is a need not just for reference laboratory diagnostic services with standardized testing, but also for rapidly deployable tests, ideally for point-of-care (POC) use within affected communities (POC testing is also covered in the *Innovation, Technology, and Design* chapter).

At the laboratory level, multiplex platforms that can test simultaneously for a wide array of pathogens as well as the use of whole-genome sequencing are likely to play an increasing role in future diagnosis, although their realization will require advances in computational power (Hasin, Seldin, & Lusis, 2017). For completely novel pathogens, reference microbiology services will need to offer more traditional techniques such as viral cell culture, polymerase chain reaction (PCR) amplification, serological techniques, electron microscopy, and animal models. For example, the severe acute respiratory syndrome (SARS) coronavirus was first diagnosed through cell culture in a reference microbiology laboratory, with random-amplification PCR and subsequent sequencing and comparison to other known coronaviruses (Drosten et al., 2003).

During future outbreaks, assuming they have appropriate sensitivity and specificity, POC tests could be used for rapid exclusion of pathogens and direction of patients to either mainstream health services or prompt isolation (Walker et al., 2015). POC tests are particularly useful in remote, under-resourced settings that lack the laboratory infrastructure or capacity to conduct traditional diagnostic testing. Not only could they quickly inform a prescriber whether a patient has an infection, but also identify whether it is viral or bacterial, and even indicate antimicrobial susceptibility, allowing practitioners to administer the most appropriate available therapy. This type of information would not only improve outcomes, but also facilitate infection-control efforts by shortening the time that an individual is infectious, and ensure that antimicrobial drugs are used only when indicated.

Taking diagnostic testing outside of the traditional laboratory environment requires tests to be independent of the need for a power supply and sufficiently robust that they do not need to be stored in temperature-controlled environments (Boeras, Nkengasong, & Peeling, 2017). Frontline health workers are also required to perform these tests and interpret the results with minimal training. Thus, decentralizing testing in this way can place considerable strain on fragile health systems in terms of supply chain management, training, quality assurance, and monitoring impact. As such, POC tests have been described as a double-edged sword, necessitating that technological innovations be accompanied by innovations in health service delivery.

The national reference laboratory is best positioned to take the lead on such technology introduction and implementation. A national system of training and quality assessment needs to be established in which testing data and quality assessment results are linked to the laboratory information system. This system can, in turn, be linked to the supply-chain management system to optimize timing and frequency of supply restocking to avoid stock-outs and wastage, and improve the efficiency of the health system.

Laboratories also have a key role to play in building connectivity and communication systems to facilitate the translation of raw data into intelligence to optimize disease-control efforts. Many new POC diagnostics have in-built data transmission capabilities via cellular networks or Wi-Fi. This allows results to be automatically uploaded to a central, secure health information system, enabling real-time surveillance (Boeras et al., 2017; Mazzola & Pérez-Casas, 2015). Smartphones offer fast computing, a user-friendly interface, wireless communication, and connectivity to data stored in the "cloud," all at declining prices—characteristics that have promoted their increasing uptake in LMICs. For example, Laksanasopin et al. (2015) have developed a low-cost smartphone dongle that can perform a full laboratory-quality enzyme-linked immunosorbent assay (ELISA), with all power drawn from the smartphone. Piloted in Rwanda, the tool was found to generate results in 15 minutes, and rivaled the gold-standard human immunodeficiency virus (HIV) tests with a sensitivity of 92% to 100% and a specificity of 79% to 100%. Further, 97% of patients preferred the dongle to laboratory-based tests, with most pointing to the convenience of obtaining quick results with a single finger-prick. Technologies like this are making diagnostic tests accessible to almost any population with access to smartphones.

So far, progress in diagnostic capacity has generally been slow to materialize. It is more expensive for patients and more time-consuming for doctors to use a diagnostic test than to prescribe a drug immediately, even if the test could help save costs and reduce waste at the health-system level. Further, in many parts of the world, a culture of self-medication is deeply entrenched. Where antimicrobials can be bought over the counter without prescription, it is difficult to incentivize people to pay for a test, wait for the results, and then pay for the treatment (if indicated), rather than just pay for the drug straight away. Moreover, few pharmaceutical companies have a commercial interest in the advent of rapid diagnostics, which would reduce the number of antimicrobials prescribed, or in the development of diagnostics for poverty-related diseases that offer little return on their investment.

Nevertheless, a number of recent initiatives have sought to incentivize development and catalyze innovation in this field so as to tackle antimicrobial resistance in particular. For example, the Longitude Prize will be awarded to an entrant that can develop a cheap, accurate, rapid, and easy-to-use POC test kit for bacterial infections that will allow health professionals worldwide to administer the correct antibiotics at the right time (Longitude Prize, 2017). The winning tool must be portable, easy to use anywhere in the world with minimal training, give an accurate result within 30 minutes, and be affordable to everyone who needs it. Ultimately, the full impact of investments in such technologies will be realized only if these tests are adopted, scaled up, and evaluated at all levels of the health system. It is also critical to understand the local context within which the tool will be used, with attention paid to economies, acceptability, and individual and community needs and impact (Boeras et al., 2017).

Evaluating Essential Public Health Functions

To ensure best practice in public health, performance evaluation should be developed and integrated into all levels of a country's public health infrastructure (Bellagio District Public Health Workshop Participants, 2016). However, the complexity of public health practice and the services that they provide do not easily lend themselves to a simple approach to quality improvement (Gunzenhauser et al., 2010). This perhaps explains why the field of public health quality improvement has lagged behind clinical quality assurance efforts. Yet, ongoing evaluation of public health practices is critical to ensure a well-functioning health system, and by delegating this task to a specific unit or body the likelihood of efforts being put into practice is increased. Such units should conduct assessments of the essential public health functions (EPHF) (see Exhibit 14-1) and the results shared with relevant health workers to facilitate improvement efforts.

Many toolkits are available to conduct such assessments at regional, national, and subnational levels; however, these need to be adapted to the local context and made actionable if the results are to effectively guide quality improvement strategies (Bellagio District Public Health Workshop Participants, 2016). One such assessment tool has been developed by WHO's Eastern Mediterranean Regional Office (EMRO) to evaluate the implementation of EPHF in countries of the region. The initial request for assessment is made by a member state on a voluntary basis, and the state provides funding for the assessment to ensure ownership. The involvement of a national team throughout the process secures high-level commitment, thereby contributing to strengthening national capacity in this area and ensuring sustainability. Further, the two-tier assessment process (self-assessment followed by external assessment), coupled with the involvement of all technical departments within WHO EMRO, ensures all relevant information is available and an all-inclusive approach to the inter-related functions is followed (WHO Eastern Mediterranean Regional

Office, 2017). This process is similar to that adopted by the Joint External Evaluation (JEE) process described later in this chapter. However, the former covers a wider range of public health functions, whereas the latter mainly focuses on prevention and control of outbreaks of infectious disease.

Evaluation instruments have been developed for use at subnational (district) health system level. This consideration is particularly important in LMICs with decentralized health systems (or where decentralization is in process), as it increases the complexity of operations at the district level. Bishai et al. (2016) have developed and piloted a district health system self-audit tool that can provide quantitative data on how district health authorities perceive their performance of the EPHFs. This tool was developed in partnership with local health authorities and stakeholders to ensure that it described the EPHF specific to the local health system. District health offices then used the tool to assess their performance of the EPHF, identifying gaps and weaknesses and enabling the creation of an improvement plan to address the issues flagged. The self-assessment approach facilitates local ownership of the evaluation process. It also takes advantage of local knowledge, enabling priority areas to be identified, and affords the opportunity for the broader inclusion of frontline workers on district health teams, whose views are often not heard by the central ministry. This tool has been used on a trial basis in Mozambique and Botswana, but could be adapted to meet the needs of other countries in strengthening public health practice or for the monitoring of district-level performance (Bishai et al., 2016). Methods of evaluation are also covered in the *Evaluations of Large-Scale Health Programs* chapter.

Workforce and Leadership

A sufficiently sized, qualified, and well-resourced health workforce is one of the building blocks of a strong public health infrastructure. Countries need a dedicated workforce with a defined scope of work and clear career paths to facilitate continuous learning and professional development (Bellagio District Public Health Workshop Participants, 2016). Strengthening of the public health workforce is particularly important, as its multidisciplinary nature—encompassing a diverse range of professionals from epidemiologists to public health nurses, as well as clinically qualified doctors—requires specialized training programs and targeted recruitment strategies to fill any gaps in the workforce (Bellagio District Public Health Workshop Participants, 2016).

Today, many LMICs suffer from an acute shortage of public health workers, particularly in rural areas. To address this gap, many countries have chosen to deploy community health workers, who can provide basic primary care outreach services in rural areas. The Ethiopian health system, for example, is renowned for its focus on strengthening the primary health sector, which has been achieved by deploying health extension workers (HEWs) at the community level (Reich, Javadi, & Ghaffar, 2016). HEWs provide preventive and basic curative care, and also conduct promotional work (Admasu, 2016). The work of HEWs is complemented by the Health Development Army (HDA), which engages women's groups to disseminate health information and facilitate uptake of critical health services (Admasu, 2016). The implementation of this dual-cadre community health worker program, including both HEWs and HDA volunteers, has facilitated health leadership at the local level with impressive results: Ethiopia is a leader in Africa both in delivering healthcare to communities (99% of the population has access to primary healthcare) and in advancing indicators of health (Admasu, 2016).

India is another example of a LMIC that has made considerable progress in public health recently, evidenced by accelerating declines in infant and under-5 mortality rates, and in maternal mortality ratios (Kumar, Bothra, & Mairembam, 2016). Such progress has been facilitated by the establishment of the Public Health Foundation of India (PHFI) in 2006 (PHFI, 2015). The PHFI was born out of the need to redress the limited institutional capacity in India for strengthening public health training, research, and policy development, and has as a key mission the development of the public health workforce. The value of training in public health is shown by the southernmost state of Tamil Nadu, which has among the best public health outcome indictors for states in India, performing consistently above the Indian national average (Balabanova, McKee, & Mills, 2011). This achievement has been credited, in no small part, to the presence of a distinct public health management cadre at the district level. Staff are charged with the management of primary health institutions and undergo a series of training programs in public health and allied subjects, including managerial skills. Staff members are also expected to qualify for a diploma in public health, enabling them to climb the career ladder. The main advantage of the public health cadre is that it provides managerial continuity at the district and higher levels of primary healthcare. Managers are therefore familiar with the issues faced by their area and understand how they should best be managed. This has resulted in the speedier allocation and more

effective use of resources received from the central government in Tamil Nadu in comparison with most other states, enhancing access to care. The importance of such cadres is becoming increasingly recognized, as reflected in the 2017 National Health Policy's proposal to create a public health management cadre in every state of India, both rural and urban, within 10 years (Ministry of Health and Family Welfare, 2017).

Given chronic shortages in health workers, it is also necessary to use human resources efficiently (Sales, Kieny, Krech, & Etienne, 2013). One way to do this is to apply the concept of "task-shifting," which involves the redistribution of specific tasks among health workers (WHO, 2008b). Often tasks are moved from highly qualified health workers to colleagues with less training and fewer qualifications, enabling more efficient use of the existing human resources and easing bottlenecks in service delivery. Task-shifting has been implemented as a pragmatic response to health workforce shortages in a number of LMICs. For example, in Uganda and Malawi, an HIV/AIDS basic care package is delivered to patients by nonspecialist doctors or nurses supported by community health workers and people living with HIV/AIDS (WHO, 2008b). This approach can also ensure an optimal skill mix of health workers in services delivery by blurring the boundaries of job roles (e.g., allowing community health workers to take on duties normally performed by nurses), although this practice can be met with strong opposition from professional bodies. Incentives can also be provided to health professionals to work in rural areas, or in less lucrative or prestigious health fields (Crisp, 2010).

The integration of indigenous, non-allopathic health workers into the primary health sector has been promoted as another method to boost workforce capacity and strengthen district health systems. India, in particular, has realized the potential of Ayurveda, Yoga and naturopathy, Unani, Siddha, Homeopathy (AYUSH) practitioners who can help to address the acute shortage of human resources if adequately trained, and who can play a key role in health promotion (Patel et al., 2015). Further, traditional healers are often held in high regard in the community; thus, integrating their services into the general public health system may help to instill trust in the system and encourage service uptake. Commitment to this cause in India is reflected in the creation of a national mission on AYUSH to strengthen the AYUSH network in the public sector (Ministry of Health and Family Welfare, 2014b), and the accordance of separate ministry status for indigenous medicine (previously a department of the Ministry of Health and Family Welfare) (Ministry of Health and Family Welfare, 2014a). Therefore, traditional healers can help support the work of trained health workers, and act as a valuable link between community members and the formal health system. However, this can be a risky practice unless these healers can be persuaded to forgo their harmful traditional practices (Crisp, 2010).

Of course, a strong public health workforce requires strong leadership. Effective leadership in times of crisis is particularly important, and, in light of recent public health incidents, a number of new approaches to leadership have been advocated. For example, using Liberia as a case study, Nyenswah et al. (2016) describe "distributed leadership," which views leadership as a "collective and social process" rather than focusing on an individual leader. Such a leadership model involves strategically engaging stakeholders and sharing responsibility and authority. In an analysis of the Liberian experience during the 2014–2016 West Africa Ebola virus epidemic, a shift from the traditional hierarchical leadership approach to a distributed model during the emergency phase of the epidemic is credited with bringing the crisis under control more rapidly in Liberia than occurred in Sierra Leone and Guinea. Unfortunately, in the aftermath of the outbreak, efforts were largely directed toward restoring the old hierarchies and decision-making structures that existed prior to the outbreak. A lack of institutional learning and memory is a key issue with this approach to leadership; finding ways to address obstacles to such learning is a high-priority task for public health leadership (Reich et al., 2016).

Given the multitude of actors engaged, public health systems goals can indeed be achieved only through distributed or "participatory" leadership, which is realized when diverse groups are empowered and free to contribute to the functioning of the health system (WHO, 2016d). Within each of these groups, individual leadership does matter—but such leadership has to foster an open, consultative process of decision making to bring out the collective strength of participatory leadership. Participatory leadership also allows for democratic debate and the free exchange of ideas, enabling established modes of working to be questioned or challenged—behavior that would otherwise be seen as deviant in rigid hierarchical systems. This type of leadership can also act to stabilize the health system against internal or external turbulence, as discussed in the next section.

Building Resilient Health Systems

A strong health system is one that is able to withstand shocks and maintain operations despite large perturbations. The United Nations Office for Disaster Risk

Reduction (UNISDR, 2017b) defines *resilience* as "the ability of a system, community or society exposed to hazards to resist, absorb, accommodate, adapt to, transform and recover from the effects of a hazard in a timely and efficient manner, including through the preservation and restoration of its essential basic structures and functions through risk management." This definition captures the need to build the resilience of people and communities, as well as the physical health system infrastructure. It also extends to the adaptive capacity of infrastructure, which is gained from an understanding of the risks and uncertainties in the environment (Cabinet Office, 2011).

While recent global crises have drawn attention to the necessity of resilient public health systems that can withstand acute, external shocks, it is equally important to ensure that health systems are resilient in the face of routine challenges—so-called "everyday resilience" (Gilson et al., 2017). These routine challenges may include drug stock-outs, funding constraints, patient complaints and changing expectations, unpredictable staff, compliance demands, organizational instability linked to decentralization processes, and frequently changing, and sometimes unclear, policy imperatives. Although unstable governance structures and inadequate resources certainly sow instability, everyday resilience is derived from new forms of distributed leadership (as discussed in the previous section) that reframe challenges to support problem solving; embody respect and empower others, particularly frontline health workers; facilitate learning and innovation; and develop and draw on social networks across, within, and outside the health system.

The need to take a resilience approach to health system strengthening was first brought to attention in the World Health Assembly's 2011 call to member states to "strengthen the resilience of the health system and society at large." Yet, within the public health field, the concept of resilience remains poorly conceptualized and under-researched.

The UNISDR definition of resilience refers to the need to build the resilience of both people and communities and the physical health system infrastructure. A resilient health system requires those who work within the system and those who use the system to exhibit flexibility in their roles or actions (Martineau, 2016). The capacities of health workers to reprioritize their clinical activities, of people who are unwell to alter their care-seeking practices, or of nontraditional health actors to take on new health roles, all contribute to building a resilient health system. Such systems are heavily dependent upon a strong and committed health workforce, characterized by personnel who are properly trained, equipped, and willing to deal with difficult and potentially dangerous situations. This involves training and deploying a sufficient number of health workers, and maintaining a stock of social capital in the health system before a disaster strikes. In the health system context, social capital has two aspects: (1) a sense of worth, community, and responsibility among health actors, and (2) robust community engagement with the health system. Health systems that are trusted and respected by the population they serve have a powerful resilience advantage (Kruk, Myers, Varpilah, & Dahn, 2015). Thus, strengthening the resilience of health systems is "more than simply providing assets or technology. It is about developing people's agency, it is about governance and power" (Béné, Wood, Newsham, & Davies, 2012). In other words, resilience is not just a function of what a system has, in terms of its physical infrastructure, but of what it does and how it does it (Gilson et al., 2017).

With regard to the physical infrastructure and operation of a health system, resilience can be seen to be made up of four components: resistance, reliability, redundancy, and response and recovery (**FIGURE 14-1**). In building resilience, the contribution made by each of these components needs to be considered (Cabinet Office, 2011).

The *resistance* component of resilience is concerned with providing protection. The objective is to prevent damage or disruption by providing in-built strength or protection to resist the hazard or its primary impact. However, resistance strategies have significant weaknesses, as protection is often developed against the types of threats that have been previously experienced rather than those that might occur in the future (Cabinet Office, 2011).

A *reliable* infrastructure is inherently designed to operate under a range of conditions, mitigating the service disruptions that might result after an event. Design standards can be set to ensure that the infrastructure can operate in the range of conditions likely to be experienced in a country or region. Service standards, in turn, protect users from disruptions in supply. However, infrastructure and service delivery

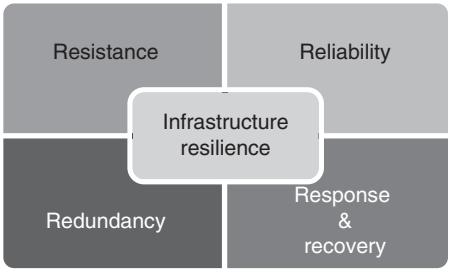


FIGURE 14-1 The four components of infrastructure resilience.

Data from Cabinet Office. (2011). *Keeping the country running: Natural hazards and infrastructure. A guide to improving the resilience of critical infrastructure and essential services.* London, UK: Author.

can adapt only so much. That is, events that exceed the specified reliability range will test the system and may stretch it beyond its capacity. Further, by focusing on a specific range of events only, the system will be ill prepared to deal with events outside of this range, leading to wider and more prolonged impacts. Thus, although reliability cannot be guaranteed, it does play a role in managing a situation at a tolerable level until full services can be restored (Cabinet Office, 2011).

The *redundancy* element of resilience is concerned with the design and capacity of the network or system. The availability of a variety of backup systems allows operations to be switched to alternative branches of the network in the event of a disruption. Having built-in redundancy ensures continuity of services. The switch to these alternative services should be made as swiftly and smoothly as possible to minimize service interruption.

Redundancy is, however, a foreign concept to the public health field, with the word "redundant" often carrying negative connotations. Operating with significant spare capacity is often deemed wasteful and inefficient; thus, in the pursuit of efficiency and as part of cost-cutting activities redundancies are often the first targets for elimination (Walker, 2010). However, by simplifying production in this way the system becomes more vulnerable to shocks and stressors, which, in the long term, works to make the system less cost-effective. Although maintaining resilience necessarily incurs costs, designers of health systems must appreciate that there is a "trade-off between foregone short-term benefits of high efficiency under narrowly constrained circumstances and the long-term persistence of the existing regime with reduced costs of crisis management" (Anderies, Walker, & Kinzig, 2006). Yet, in other areas of work, redundancies are recognized as critical elements. For example, redundancy has served as a central tenet of high reliability engineering for decades (Downer, 2009). Redundancies are also integral to natural systems. To quote Nassim Taleb (2012): "Layers of redundancy are the central risk management property of natural systems. We humans have two kidneys . . . extra spare parts, and extra capacity in many, many things (say, lungs, neural system, arterial apparatus)'" in short, "[n]ature likes to overinsure itself."

Ultimately, redundancy ensures that if a critical system fails, another is there to pick up the slack. The same principle could, therefore, be applied to the design of public health systems. To understand how such an approach might be implemented, it can be useful to look at examples of health systems that perform well under stress, such as in Palestine. This system (particularly in the Gaza Strip) consists of a distributed architecture with embedded redundancies that enable healthcare provision to withstand stressors. These redundancies include multiple service delivery points, a variety of delivery models, dispersed decision-making centers, informal power structures, multiple funding sources, and diversified supporting bodies.

The effectiveness of the *response and recovery* component depends on the thoroughness of efforts to plan and prepare in advance of events. Infrastructure owners must understand the weaknesses in their systems and have arrangements in place to respond quickly to restore services. Further, post-disaster recovery and reconstruction offers a prime opportunity to "build back better" by integrating disaster risk reduction measures into the restoration of public health infrastructure and systems so as to increase resilience (UNISDR, 2017b). This idea was first conceptualized in the aftermath of the Indian Ocean tsunami of 2004, quickly becoming the recovery effort's guiding principle, and has been advocated in many other disasters since (Fan, 2013).

All-Hazards Approach to Emergencies

Public health infrastructure must be sufficiently resilient to withstand emergencies of all types. These can threaten health directly or indirectly through the disruption of health services and critical infrastructure. Threats include natural hazards (hydrologic, geologic, climatic, and atmospheric), industrial accidents, and malicious attacks, including the use of chemical, biological, radiological, and nuclear (CBRN) weapons (Cabinet Office, 2011). While the nature of such events may vary significantly, they give rise to similar problems; thus, preparedness and response strategies for each type of disaster have much in common (Adini, Goldberg, Cohen, Laor, & Bar-Dayan, 2012). During a crisis, the need to manage resources and information, and to maintain effective lines of communications, is the same regardless of the type of incident. WHO therefore advocates a generic, "all-hazards" approach to emergency preparedness and response, while still recognizing that managing each type of crisis will require specific scientific and technical expertise (WHO, 2012).

In the coming decades, a growing number of people are expected to be affected by disasters of all types, particularly those in LMICs (Department for International Development, 2011). Disasters are expected to grow in frequency and intensity, and hazards will become more complex and unpredictable, in part due to climate change (Intergovernmental Panel on Climate Change [IPCC], 2014). However, there is

increasing recognition that disasters are not natural events, but rather the product of the interaction between the hazard and socially constructed vulnerabilities[1] and exposures[2] (Aitsi-Selmi & Murray, 2016b). The severity of a disaster depends on the level of exposure and vulnerability of the affected population, which are in turn shaped by economic, social, environmental, geographic, demographic, cultural, institutional, and governance factors (Aitsi-Selmi et al., 2016). Social vulnerability is also dynamic, in that it has been shown to vary temporally with the disaster stage and hazard context (Rufat, Tate, Burton, & Maroof, 2015). For example, in the midst of a flood, men and middle-age people tend to be most vulnerable due to risk-taking behavior and involvement in rescue operations; in the aftermath, however, women, single-parent families, and the elderly are most vulnerable due to limited resource availability and difficulties coping with disruptions to long-term care and services. Further, poor development practices, including environmental mismanagement, urbanization in hazardous areas, rapid demographic and economic changes, and ineffective governance, coupled with limited livelihood options, particularly for the poor, are all associated with increased exposure and vulnerability to disasters (IPCC, 2012). Evidence indicates that exposure of persons and assets in all countries has increased faster than vulnerability has decreased (UNISDR, 2015). This has generated new risks and a steady rise in disaster-related losses, with significant health, social, economic, and environmental short- and long-term impacts. These impacts are often felt most strongly at the local and community levels, and disproportionately affect vulnerable groups, as well as single-sector economies (UNISDR, 2015). In LMICs, where the mortality and economic losses from public health emergencies are greatest, disasters significantly impede progress toward sustainable development (UNISDR, 2015).

A single event will often give rise to a variety of cascading effects beyond the initial incident (Cabinet Office, 2011), which can have a multiplying effect on losses. For example, a prolonged period of hot weather carries the risk of thunderstorms and flash flooding. Disasters can also have adverse secondary effects that extend beyond national borders. Increasing global economic interdependence, including linked production and consumption operations, means that the impacts of a disaster can have widespread impacts on distant locations (Aitsi-Selmi, Blanchard, & Murray, 2016). To illustrate, in 2011, severe floods in Thailand stalled local manufacturing production, which in turn brought global computer supply chains and manufacture to an effective standstill (Chongvilaivan, 2012).

Consequences of Disasters on Public Health Systems

Ensuring continuity of routine care after a disaster is often a severe challenge. Population displacement and overcrowding, coupled with damage to critical infrastructure such as sewer systems, increases the risk of infectious disease outbreaks (Aitsi-Selmi & Murray, 2016b). These risks are compounded in cases where health facilities are overwhelmed or damaged, and unable to maintain normal services. Loss of medical personnel and critical infrastructure, combined with reduced service access, can result in increased morbidity and mortality. The insecure working environment for healthcare workers (HCWs) and supply-chain distribution challenges also contribute to interruptions in service delivery (Ochi, Hodgson, Landeg, Mayner, & Murray, 2014).

The ability of a health system to effectively manage a public health incident is also heavily dependent upon the degree of confidence that the affected population has in the health services, and authorities in general. This challenge attracted particular attention during the 2014–2016 West Africa Ebola epidemic (see **EXHIBIT 14-3**).

Disaster Risk Reduction

Disaster risk reduction (DRR) strategies should be "aimed at preventing new and reducing existing disaster risk and managing residual risk, all of which contribute to strengthening resilience and therefore to the achievement of sustainable development" (UNISDR, 2017b). Such efforts must consider the full spectrum of threats, including those related to rapid and unmanaged urbanization and globalization, as well as a population's social vulnerability and capacity to respond (Aitsi-Selmi & Murray, 2016b).

Countries can reduce their risk of disaster by implementing the recommendations of the Sendai Framework for Disaster Risk Reduction 2015–2030 (UNISDR, 2015). The Sendai Framework places a strong emphasis on disaster risk management (both

[1] Vulnerability is defined as the "characteristics and circumstances of a community, system, or asset that make it susceptible to the damaging effects of a hazard" (UNISDR, 2017b).

[2] Exposure is defined as the "people, property, systems, or other elements present in hazard zones that are thereby subject to potential losses" (UNISDR, 2017b).

EXHIBIT 14-3 Public Health Infrastructure Collapse During the 2014– 2016 Ebola Epidemic in West Africa

During the 2014–2016 West Africa Ebola outbreak, a serious breakdown in trust between communities and the health system was observed (Elston et al., 2015). People felt let down by the health system, which was patently unprepared for an outbreak of this scale, and also feared engaging with the health system. Rumors abounded that health facilities were deliberately infecting people, creating societal pressure to refrain from seeking help. Disengagement with the health system led to significant under-utilization of available health services, with marked reductions in patients seen at community health centers and hospitals. In the Moyamba district of Sierra Leone, the United National Children's Fund (UNICEF) reported a 10% reduction in fourth antenatal clinic attendance and a 16% reduction in births reported from May to September 2014 (UNICEF & Ministry of Health and Sanitation, 2014).

The outbreak also severely affected the healthcare workforce. Ebola-affected regions were already chronically understaffed prior to the outbreak; therefore, losses of medical personnel during the outbreak were particularly serious. In Sierra Leone, up to 28% of the total healthcare workforce was reported to have been infected with Ebola. Resources, district health management teams, and community health workers were largely reassigned to Ebola-related activities and, therefore, were unable to provide their normal services. This affected the delivery of health services, with both vertical programs and critical infrastructure being disrupted. All major health programs, including tuberculosis, HIV, malaria, and nutrition initiatives, were adversely affected, and immunization schedules were disrupted. In the Koinadugu district of Sierra Leone, the mean monthly number of children receiving all recommended childhood vaccinations by age 1 year fell from 1,152.5 during the months of January–June 2014, before the outbreak took hold, to 507.7 in the final 6 months of the year.

Disruption to communication systems also jeopardized response efforts. For example, HCWs could not call emergency ambulances or report cases, resulting in a nonfunctional surveillance system. The reductions in vaccine coverage coupled with reduced surveillance capacity raised the risk of outbreaks of vaccine-preventable infectious disease.

These factors all contributed to heightened all-cause mortality during the Ebola outbreak. In the Moyamba district of Sierra Leone, the number of deaths recorded between November 2014 and mid-April 2015 was more than 3.4 times greater than the average number of deaths registered during the same period from the previous 4 years, yet only 2% of these fatalities were confirmed Ebola cases. The cessation of usual healthcare services for malaria as a direct result of the epidemic is estimated to have led to 3.5 million additional untreated cases of malaria in Guinea, Liberia, and Sierra Leone, with 10,900 additional malaria-attributable deaths (Walker et al., 2015). The outbreak also had wider, indirect social consequences. For example, in Moyamba district, an increase in teenage pregnancies in 2014 compared to 2013 has been attributed to widespread school closures.

This case demonstrates the impact that a disaster can have on an already compromised health system, in particular through the diversion of resources and attention to disaster-related activities and the resultant reduced provision of leadership, support, and governance to the health system. It also illustrates the importance of community trust in health services, and the maintenance of this confidence during the acute phase of an incident.

preventing new risks and reducing existing risks) through pre-disaster preparedness efforts, including the need to conduct detailed risk assessments, improve early warning capabilities and impact-based forecasting, build public commitment, and establish supportive institutional frameworks (Aitsi-Selmi et al., 2016). The scope includes human-made as well as natural hazards, and related environmental, technological, and biological hazards and risks. It also recognizes the need for DRR practices to be multisectoral and transdisciplinary, engaging with relevant stakeholders at all levels, including the community. In particular, it calls for greater collaboration with the science and technology community to better understand disaster risk patterns, causes, and effects; to develop advanced methods and tools for disaster risk modeling and multiple-hazards monitoring; and to inform DRR decision making and policy.

The Sendai Framework identifies a main outcome and corresponding goal to be achieved over the next 15 years. To address the health-related aspects of the Sendai Framework, seven recommendations, known as the Bangkok Principles, were agreed in 2016 (UNISDR, Royal Thai Government, & WHO, 2016):

1. Promote systematic integration of health into national and subnational DRR policies and plans and the inclusion of emergency and disaster risk management programs in national and subnational health strategies.
2. Enhance cooperation between health authorities and other relevant stakeholders to strengthen country capacity for disaster risk management for health, the implementation of the International Health Regulations (2005), and building of resilient health systems.

3. Stimulate people-centered public and private investment in emergency and DRR, including in health facilities and infrastructure.
4. Integrate DRR into health education and training and strengthen capacity building of health workers in DRR.
5. Incorporate disaster-related mortality, morbidity, and disability data into multiple-hazards early warning systems, health core indicators, and national risk assessments.
6. Advocate for and support cross-sectoral, transboundary collaborations, including information sharing, and science and technology for all hazards, including biological hazards.
7. Promote coherence and further development of local and national policies and strategies, legal frameworks, regulations, and institutional arrangements.

The Sendai Framework and two other landmark agreements on sustainable development and climate change reached in 2015 (the SDGs and the Paris Agreement on Climate Change, respectively) provide an opportunity to build coherence across these policy streams. Indeed, the Sendai Framework and Bangkok Principles specifically highlight synergies with health, sustainable development, and climate change and the need to facilitate mutually beneficial capacity development and joint policy initiatives (Aitsi-Selmi et al., 2016).

Implementation of the Sendai Framework

The Sendai Framework is a voluntary agreement, so its implementation ultimately depends on political will, the availability of funds for DRR activities, and the willingness to collaborate across institutional and country boundaries (Aitsi-Selmi & Murray, 2016a). The framework defines precise global targets, but measuring progress may prove challenging, as the framework does not specify how these targets should be measured or which data sources to use. Only two of the targets specify a referent baseline, and none is measured in absolute terms. The minimal availability of disaster loss data is also a challenge. At present, the ability to monitor the human and economic damage from disasters is limited, and at the national level, few countries have any sort of disaster loss accounting in place (Cutter & Gall, 2015).

Progress on the Sendai Framework is due to be reviewed biannually at the Global Platform for Disaster Risk Reduction. This forum functions to enable governments, NGOs, scientists, practitioners, and UN organizations to share knowledge and experiences, and to formulate strategic guidance for the implementation of the Sendai Framework (UNISDR, 2017a).

Chemical, Biological, Radiological, and Nuclear Incidents

Acute environmental emergencies resulting from exposure to CBRN material can have hugely detrimental effects on health and life. The Bhopal gas tragedy in India, considered the world's worst industrial disaster, occurred when more than 40 tonnes (44 tons) of methyl isocyanate gas leaked from a pesticide plant, killing at least 3,800 people (Broughton, 2005). In addition, it resulted in significant morbidity and premature mortality of many more thousands of people. Incidents involving exposures to abandoned radioactive sources in LMICs have also been reported, such as the cesium incident in Brazil (International Atomic Energy Agency, 1988). The global production, trade, and use of chemicals are predicted to increase six-fold by 2050, especially in developing and transitional economies, where chemical production, extraction, processing, and use are closely linked to economic development (Marchal et al., 2011).

Radiation overexposure accidents are rare and decreasing compared to chemical incidents, but their health consequences can be severe and long term (Coeytaux et al., 2015). The Fukushima Daiichi nuclear power plant accident demonstrated the health risks of unplanned evacuation and relocation for vulnerable people, such as hospital inpatients and elderly people requiring nursing care. Furthermore, the displacement of people that followed this event created a wide range of public healthcare and social issues (Hasegawa et al., 2015).

Adequate legislation needs to be in place in LMICs to mitigate future risks of significant environmental incidents. Following the Bhopal gas tragedy, the government of India amended three key acts that deal with industrial hazards and also passed a comprehensive new environmental law (Government of India, 1986). This law has vastly improved regulatory coverage of hazardous technologies and substances.

In today's world, perhaps the most salient issue is the weaponization of CBRN materials. CBRN weapons are some of the most indiscriminate and deadly weapons in existence (Unal & Aghlani, 2016). Along with the physical damage they can inflict, these weapons generate significant economic and societal disruptions by instilling widespread fear in the populace and sowing the seeds of social, economic, and political uncertainty and instability (Cornish, 2007). Indeed, a major incentive for use of these weapons is their capacity to cause significant disruption across multiple sectors, as

well as revenue loss for governments (Unal & Aghlani, 2016). For example, decontamination after the 2001 U.S. anthrax attacks—which produced almost 3,000 tonnes (3307 tons) of contaminated waste—cost an estimated $800 million (The Royal Society, 2004).

While such incidents are rare, the likelihood of these weapons being deployed is growing, fueled by heightened terrorist activity combined with advances in technological and scientific capabilities, and the accessibility of dual-use materials (which can be used for both civilian and military purposes) (Unal & Aghlani, 2016). Inadequately secured stockpiles of decommissioned military CBRN material, and increases in the trafficking of radiologic and nuclear materials have augmented the risk in recent years (HM Government, 2010). The probability and nature of such a threat vary geographically, and are linked to the intentions, incentives and disincentives for acquisition, and capabilities of a potential perpetrator (Unal & Aghlani, 2016). The scale of the impact that can be potentially achieved through the deployment of a CBRN weapon is a clear incentive for their use; thus, despite the rarity of such incidents, there is an ongoing need to develop a robust public health infrastructure so as to reduce the risk of such an incident occurring, and to effectively respond to an event if one should occur.

Stages of an Emergency Response

The all-hazards approach to disaster preparedness and response allows for a generic framework, regardless of the cause, to guide how countries should prepare and respond to a public health emergency, be it a natural disaster or a terrorist-related CBRN incident. This framework should include the following elements:

- *Pre-disaster planning.* Examples include the European Union's Seveso Directive (European Commission, 2016) and the International Health Regulations (WHO, 2016b).
- *Exposure and risk assessment.* In the event of an incident, immediate risks to health should be addressed by advice given at an early stage (Bradley, Meara, & Murray, 2015).
- *Countermeasures and risk management techniques.* These should include measures to prevent or reduce people's exposure, and to treat and transfer casualties and vulnerable subpopulations.
- *Communicating with the public.* Prompt, honest, and informative communication with clear messages for the public during an emergency is a key factor in allaying concerns. **EXHIBIT 14-4** outlines the basic principles of good communication during public health emergencies.
- *Recovering from an incident.* The importance of the recovery phase of a disaster is highlighted in the Sendai Framework, which advocates for the need to "[enhance] disaster preparedness for effective response and to 'Build Back Better' in recovery, rehabilitation, and reconstruction" (UNISDR, 2015).
- *Follow-up of affected populations.* There are two key reasons to undertake follow-up of populations affected by public health emergencies: (1) as part of the response phase of a disaster to help and provide services for those affected, and (2) for

EXHIBIT 14-4 Basic Principles of Good Communication During Public Health Incidents

- Determine the best organizations and individuals to deliver the public health messages.
- Commence communication as soon as possible, while acknowledging inevitable uncertainties.
- Set up national telephone helplines that have sufficient capacity to deliver the public health messages and answer questions.
- Take into account people's desire to find out about the safety of family and friends.
- Comparing unfamiliar risks with familiar ones can help put risks into context. However, care must be taken to ensure that elective risks are not used as the comparison for unavoidable risks.
- Always be truthful, or give the reasons why the truth cannot be disclosed. For example, for security purposes, some facts cannot be shared.
- Develop public generic information materials in "peacetime" that can be adapted quickly during an incident to expedite delivery of public health advice. These materials should include information about the mechanisms of harm, where harm may occur, and the measures people can take to protect themselves and how effective those measures are likely to be.
- Develop skills and experience in using modern communication methods, such as social media, to deliver health messages.
- Remember that emergency responders, as well as the public, may need information and support.

Modified from Bradley, N., Meara, J., & Murray, V. (2015). Principles of public health emergency response for acute environmental, chemical, and radiation incidents. In R. Detels, M. Gulliford, Q. A. Karim, & C. C. Tan (Eds.), *Oxford textbook of global public health* (6th ed., pp. 1608–1620). Oxford, UK: Oxford University Press.

scientific purposes, by collecting information regarding the health effects of an exposure.

- *Disseminating the lessons learned.* Reporting incidents is important to share lessons learned, reduce the likelihood of the incident reoccurring, and inform future emergency responses. Reports should be published and made publicly available to enable maximum benefit.

Health Security

Health security—essentially protection from health threats—is recognized as one of the most important nontraditional security issues (Heymann et al., 2015). The notion of health security has developed and been applied to public health from a different perspective than the human rights perspective discussed earlier in this chapter. That is, health security approaches public health from the need and responsibility for authorities, whether local, national, or global, to protect the health of their citizens.

History

As cross-border trade, pilgrimage, and war intensified throughout history, international travel increased, and along with it the international spread of disease. By the 14th century, it was understood that waves of epidemic disease such as plague were seeded by international travel.

Protection of populations against the importation of these diseases soon became an important part of national security for established governments. Defense mechanisms were first developed by individual countries in an attempt to prevent the entry of disease at their borders, and later collectively by countries to provide for global health security—that is, to protect against the international spread of disease. The 14th-century city-state of Venice was the first to introduce quarantine measures, whereby arriving ships were not permitted to dock, and people at land borders were held in isolation, for 40 days.

By the mid-19th century, governments had become concerned that quarantine measures were not preventing the importation of another disease—cholera. They recognized that better cooperation between countries was required to decrease the risk of international spread of these diseases. In turn, a series of international agreements were developed, initially in Europe, to apply certain measures at international borders. These included the first International Sanitary Convention in 1892, which dealt with cholera, and the 10th International Sanitary Conference in 1897, which sought to prevent the spread of plague. Obligatory telegraphic notification of cases of cholera and plague, shortly followed by the inclusion of yellow fever as a notifiable disease, was initiated among countries in the Americas. International agreements were soon broadened to include both Europe and the Americas.

During the early 20th century, two major international sanitary bureaus were created to support the development of regional public health capacity against infectious diseases—one for the Americas, called the Pan American Sanitary Bureau (now the Pan American Health Organization), and one for European countries, called the Office International d'Hygiène Publique. Cooperation between these two sanitary bureaus continued and, after the creation of WHO, broad International Sanitary Regulations were developed as a means of fostering international cooperation in the control of cholera, plague, yellow fever, and smallpox.

The International Health Regulations

In 1969, WHO developed the International Health Regulations (IHR), which were specifically aimed at better ensuring public health security with minimal interruption of travel and trade. Under the IHR, countries were required to report to WHO whenever an outbreak of plague, cholera, smallpox, or yellow fever occurred within their borders, and other countries could then apply control measures at their border posts. If a country reported an outbreak of yellow fever, for example, information about the outbreak was published in the WHO weekly epidemiological bulletin, and all other countries could then require evidence of yellow fever vaccination before travelers from that country were allowed to cross their international border. Under the IHR, countries were also required to take other actions that would prevent the proliferation of disease vectors such as rodents and mosquitoes, and international conveyances were required to have up-to-date sanitary certificates.

The IHR provided a legal framework for global surveillance and response, with the potential to decrease the world's vulnerability to four infectious diseases that were known to cross international borders. It soon became evident, however, that countries often reported outbreaks late, or not at all, because of fear of stigmatization and economic repercussions. It was likewise understood that the IHR did not meet the challenges caused by emerging infectious diseases and their rapid global transit.

The IHR were subsequently reviewed and revised, with a change in focus from stopping disease at borders to developing the national capacity to detect and respond to disease outbreaks where and when they occur. The revised IHR (2005) introduced a

requirement that each country develop and maintain a set of core capacities for surveillance and response so that it could rapidly detect, assess, notify, report, and contain the events covered by the regulations, thereby minimizing their potential for international spread and negative economic impacts. Countries were to be monitored as to whether they had developed these core capacities by annual voluntary reports to the WHO, based on a standardized self-assessment tool.

The importance of the IHR was clearly demonstrated when the Ebola virus crossed national borders during the West African outbreak (2014–2016). There was a clear understanding that outbreaks in West Africa were a threat to health security everywhere; people with infection traveled across borders within Africa and to Europe and North America, where they unintentionally caused small chains of transmission far from the epicenter of the outbreak. This outbreak underlined the importance of strong national public health capacity to detect and respond to disease outbreaks where and when they occur, as well as the importance of a functioning safety net of collective global action as required by the IHR.

Collective health security has been the focus of attention as described previously—that is, the centuries-long, commonly understood conceptualization of national health security. But there is a second, equally important aspect to health security that is less appreciated: individual health security. This type of security comes from access to safe and effective health services, products, and technologies. Ebola-infected health workers from developed countries were repatriated from West Africa for care in their own countries, where hospitals could provide them with safe and effective access to life-saving medicines and services. Meanwhile, Ebola-infected West Africans (health workers and others) had to accept that their healthcare was not always safe, not always effective, and not always accessible. As the Ebola epidemic unfolded, the part played by substandard infection control and inadequate access to effective health products and services clearly demonstrated the wider scope of health security, in the form of the intertwining of collective and individual health security. In the future, attempts to strengthen global health security must take into account not only the need for fully implemented IHR, but also the need for strong and resilient health systems for patient management and UHC (as highlighted in the first part of this chapter).

Public Health Emergencies of International Concern

The IHR (2005) require collective action by all WHO member states in the event that an emerging or reemerging infectious disease begins to spread internationally, and the free sharing of information pertaining to this threat. In this way, the IHR provide a safety net against the international spread of emerging or reemerging infections, requiring collaboration between all countries to ensure the timely availability of surveillance information and technical resources that better support international public health security. Collective action begins with mandatory reporting of any public health emergency of international concern (PHEIC), and continues with the use of real-time evidence to recommend measures to stop their international spread.

A PHEIC is defined as an extraordinary event that could spread internationally or might require a coordinated international response. Each newly identified outbreak is evaluated for its potential to become a PHEIC by the country in which it is occurring, even though reporting might legitimately come from elsewhere, using a decision tree instrument developed for this purpose.

Once a potential PHEIC is identified and reported to WHO by the country or countries concerned, a mandatory ad hoc Emergency Committee is set up to review the evidence available to WHO and to conduct a risk assessment. A recommendation is then made to the WHO Director-General as to whether the criteria for a PHEIC are met. The Director-General then uses this recommendation, and other sources of information, to decide the course of action.

An Emergency Committee was first established when influenza A (H1N1) was reported as a PHEIC by Mexico in 2009, after which the Director-General declared a pandemic. Emergency Committees have also been set up for other potential PHEICs: MERS, the continued spread of poliomyelitis, Ebola, Zika virus and microcephaly, and yellow fever.

Under the revised IHR, countries are required to notify WHO of even a single occurrence of a disease that would always threaten global public health security—smallpox, poliomyelitis caused by a wild-type poliovirus, human influenza caused by a new virus subtype, and SARS. In addition, a second list includes diseases of documented—but not inevitable—international impact. An event involving a disease on this second list, which includes cholera, pneumonic plague, yellow fever, Ebola, and the other hemorrhagic fevers, still requires the use of the decision tree instrument to determine whether it is a PHEIC.

In summary, it is now understood that the best defense is rapid detection and response to all public health hazards when and where they occur or emerge, with a global safety net of collaboration should a disease outbreak begin to spread internationally.

Strengthening the International Health Regulations

The IHR (2005) became legally binding in 2007, with all WHO member states required to develop certain minimum core public health capacities (WHO, 2016c). To achieve this goal, member states were initially encouraged to assess their own existing national structures and resources. The results of the self-assessments were intended to form the basis of action plans to increase core capacities and improve functioning.

Unfortunately, noncompliance with these regulations is rife. In public health emergencies that have occurred since adoption of the revised regulations, decisions have been made by states that were neither scientifically sound nor consistent with WHO guidance, and the IHR principles obligating countries to respect human rights and cause minimal disruption to the international flow of people and goods have been often dismissed (Katz & Fischer, 2010) (see also the *International Trade and Health* chapter). For example, during the 2009 H1N1 influenza pandemic, several countries suspended flights to North America, and 20 countries imposed bans on the importation of pork products from Mexico, Canada, and the United States. More recently, during the West Africa Ebola epidemic, many countries imposed unjustified trade and travel restrictions, and few countries notified WHO that they were implementing additional measures or justified their reasons for doing so (as required under the IHR), significantly interfering with international traffic (WHO Executive Board, 2015). Domestic politics can also play a role in the imposition of trade and travel barriers. During the H1N1 outbreak, democracies with weak health infrastructure were more likely to erect barriers and to do so quickly to prevent a loss of domestic political support (Worsnop, 2016).

IHR noncompliance also affects how quickly outbreaks are declared by countries. States are often reluctant to declare outbreaks for fear of being subjected to trade and travel sanctions, and the subsequent economic repercussions of such restrictions, especially on tourism and food exports. This hesitation may have serious implications for the scale of the response mounted. If a government refuses to cooperate, there is little WHO can legally do to remedy the situation.

Like most international agreements, the IHR do not contain enforcement mechanisms and WHO does not have the authority to impose sanctions on countries flouting the regulations (Wilson, Brownstein, & Fidler, 2010). The World Trade Organization, while able to enforce trade regulations, does not work fast enough to apply sanctions in the context of health emergencies. Consequently, there is a need to strengthen the incentives and disincentives so that governments report disease outbreaks early and are dissuaded from implementing trade and travel restrictions without a scientific basis (Moon et al., 2015). To enhance compliance with the IHR, it has been suggested that WHO publicly request countries to justify additional measures undertaken and urge businesses to reconsider restrictions (Gostin, DeBartolo, & Friedman, 2015). To encourage good behavior and promote transparency, WHO could publicly acknowledge the nations and businesses that comply with recommendations, while publicly disclosing lists of those that do not. However, doing so would require a delicate balancing act between WHO's role as trusted interlocutor with governments on sensitive information and its role as guardian of the IHR.

To further encourage compliance with the IHR, funding bodies could create economic incentives for early reporting by committing to the rapid provision of emergency funds to assist countries when emergencies strike and compensating them for any economic losses that might occur (Moon et al., 2015). Moreover, insurance plans offered, for example, by the World Bank's proposed Pandemic Emergency Financing Facility, could mitigate the economic costs linked to outbreak reporting.

It has also been suggested that state parties consider pursuing dispute mediation from the Director-General or compulsory arbitration for economic losses incurred. Successful cases by countries affected by travel or trade restrictions or human rights violations would establish a strong precedent that would likely enhance compliance (Gostin et al., 2015). Moreover, the World Health Assembly could amend the IHR to raise temporary recommendations to a binding status.

As private actors, such as shipping companies and airlines, are not directly bound by public international law, alternative governance mechanisms are needed to prevent steps to isolate countries during outbreaks (Moon et al., 2015). The onus rests on state parties to instill confidence in private service providers by responding appropriately to emergencies, following WHO guidance, and abiding by the IHR (Garrett, 2015). It also falls on WHO to prove its competence in handling such situations so as to increase private actors' confidence in its ability to manage PHEICs.

However, as seen in both the H1N1 and Ebola epidemics, no matter how comprehensive or well embedded the IHR governance structures, the issue of national sovereignty endures. Ultimately, nations can and will make their own decisions in response to a public health emergency, regardless of the global health governance structures in place (Katz & Fischer, 2010).

For example, in 2007, the Indonesian government decided to stop sharing avian influenza virus strain information with WHO (Fidler, 2008). This decision stemmed from Indonesia's unsuccessful attempt to obtain a vaccine based on the strain that it had provided to the international surveillance network for influenza. Indonesia argued that this situation exposed inequities in the global influenza surveillance system, whereby LMICs provide vital information but do not get to share in the benefits because they are unable to afford the resultant vaccine. Indonesia asserted sovereignty over the samples originating within its territory, in accordance with the Convention on Biological Diversity. The enforcement of sovereignty is further supported by the Nagoya Protocol on Access and Benefit Sharing (Secretariat of the Convention on Biological Diversity, 2011). Sovereignty is often the only card that LMICs can play to ensure that their voices are heard on the world stage. This has many implications for data and specimen sharing, and for the expansion of research and development in LMICs.

Global Health Security Agenda

Launched in February 2014, the Global Health Security Agenda (GHSA) is a growing partnership of more than 50 nations, international organizations, and nongovernmental stakeholders to help build countries' capacity as required by the IHR (GHSA, 2017a). The GHSA takes a multisectoral, multinational approach, aiming to reduce infectious disease threats, promote global health security as an international priority, and make concrete commitments to other frameworks such as the IHR and the World Organization for Animal Health's Tool for the Evaluation of Performance of Veterinary Services (OIE PVS) (CDC, 2016; GHSA, 2017a; World Organization for Animal Health, 2017). The GHSA was a U.S. initiative conceived and established by the Obama administration, with an initial focus on ensuring the capacity of countries to take effective action against emerging threats, particularly in light of the 2009 pandemic flu response. It grew from a view that the world was not able to mount an effective response to such threats; that the core capacities integral to the IHR were not in place in many countries, especially in LMICs; and that WHO was not able to monitor effectively, or enforce compliance with, the IHR.

Whilst the GHSA is a U.S.-led initiative, it was always intended to be multilateral in nature. In its early stages, the United States approached partner countries to support the "Agenda" and commit to increasing capacities globally, including by providing funding or technical expertise to countries in need of assistance. Membership is now open to all countries, with the decision to join made by national governments and often led by the Ministry of Health. As more countries joined the initiative, the United States recognized the need for others—particularly those outside of the "global north"—to take on leadership roles within the alliance. This issue has further been addressed by increasing emphasis on the GHSA as a *support* to WHO's leadership in IHR work and health security in general.

The relationship between the GHSA and WHO has often been tense. At its inception, a widely held view was that the GHSA had a bias against WHO and that it aimed to supplant WHO as the lead organization for global health security. The counter-view was that WHO was not effectively tackling the key issues in health security (such as failing to fully implement the IHR), that WHO leadership would not (or could not) face up to this issue (because of WHO's governance position with member states), and that WHO was not fit for purpose with regard to international outbreak response (bearing in mind that this was before the Ebola outbreak in West Africa).

The GHSA's work is organized into a discrete set of agreed "Action Packages." All GHSA member countries participate in one or more Action Packages either as a "leading" or "contributing" country. Each package is accompanied by a 5-year target, indicators by which to measure progress, and lists of baseline assessment, planning, monitoring, and evaluation activities to support successful implementation (GHSA, 2017b). A Steering Group oversees the work, with a rotating country chair. The countries on the Steering Group, particularly the chairing country, fund the organizational arrangements (although in reality, activities are largely dependent on U.S. funding).

The Action Packages that constitute the GHSA's work at the country level are largely funded through a combination of country funding and donor funding, with the additional element that the United States has specific GHSA funding from Congress to support activities in line with GHSA priorities and in countries with which the United States wants to partner. The major change that the GHSA has driven is that funding for this aspect of health strengthening targets activities that are an agreed priority for the country, is part of an agreed action plan, and is monitored independently.

Response to Outbreaks: Linking Global, Regional, and National Perspectives

Though there were significant strengths in the international response to the 2014–2016 West African Ebola

outbreak, the delay between the initial Mèdecins Sans Frontières (MSF) warning in March 2014 and the WHO declaration of a PHEIC in August of the same year was lengthy and problematic (Grepin, 2015; WHO, 2014). In a retrospective reflecting on the response to the outbreak, many important lessons regarding better coordination and collaboration have been highlighted (House of Commons Science and Technology Committee, 2016).

International assistance in the event of a public health emergency can be facilitated through the Global Outbreak Alert and Response Network (GOARN), in the case of an infectious disease outbreak, or through the Global Emergency Medical Team (EMT) Registry, in the case of a humanitarian crisis. GOARN is a collaboration of existing institutions and networks that pools human and technical resources for the rapid identification, confirmation, and response to outbreaks of international importance (WHO, 2017b). WHO may coordinate an international outbreak response using resources from GOARN. Many countries and organizations have also developed Emergency Medical Teams (EMTs)—groups of health professionals who are able to provide direct clinical care to populations affected by emergencies (WHO, 2017a). EMTs often have a strong trauma and surgical focus, and have a long history of responding to sudden-onset disasters such as the Indian Ocean tsunami and the Haitian earthquake.

To coordinate the rapid deployment of highly qualified and adequately equipped medical teams in the event of a disaster, WHO has developed the Global EMT Registry. This Registry consists of EMTs that meet WHO technical standards and that have been classified according to capacity and capability. The Global EMT Registry enables a country affected by a crisis to call on quality-assured teams capable of providing the assistance required. The Registry also acts as a coordination mechanism for all partners who aim to provide clinical care in crises, and ensures the deployment of teams with the required skill set.

The United Kingdom, for example, has two mechanisms by which to deploy cadres of professionals trained in humanitarian and outbreak response. One is a new dedicated team of public health experts (including infectious disease specialists, epidemiologists, microbiologists, and logisticians) that has the capacity to deploy within 48 hours of notification. Known as the U.K. Public Health Rapid Support Team (UK PHRST), it is a partnership between Public Health England (PHE), the London School of Hygiene and Tropical Medicine, King's College London, and the University of Oxford, which brings together response, research, training, and capacity-building activities. The second mechanism involves a broader group of medical and public health reservists, deployable through UK-Med, a charity hosting the U.K. International Emergency Medical Register (Reece et al., 2017). UK-Med is certified by WHO as an EMT and is listed on the Global EMT Registry.

Internationally, WHO has developed a similar emergency response for infectious disease threats by forming the Emerging Diseases Clinical Assessment and Response Network, through which clinical scientists can be deployed to rapidly assess unknown clinical management questions or public health response challenges (WHO, 2016a). The U.S. CDC has a longstanding Epidemic Intelligence Service (EIS) field epidemiology training program, and the European Centres for Disease Control (ECDC) has founded a European Programme for Intervention Epidemiology Training (EPIET). Within the United Kingdom and some other countries, Field Epidemiology Training Programs have trainees ready to be deployed to assist with initial assessment of evolving outbreaks.

During the West African Ebola outbreak, a cornerstone of the response was the formation of a central command and control structure to host the coordination of all aspects of the outbreak response, from active case management to safe burial, and from contact tracing to quarantine activities (Brown et al., 2016). More broadly, since 2015 WHO has developed a framework for a Public Health Emergency Operations Centre (EOC) Network to ensure that novel outbreaks are appropriately resourced with a designated "public health emergency management personnel," who can rapidly come together to "coordinate operational information and resources for strategic management of public health events and emergencies" (WHO, 2015g).

Notwithstanding the array of international response options available, individual countries must develop their own capacity for meaningful surveillance and outbreak response, as highlighted by the IHR (WHO, 2016b). International response mechanisms can assist both with this process and with incident response, but sustainable control of public health threats is underpinned by the core pillars of countries' own public health systems. Consequently, investments in underlying health systems improvements, including human resources for health, postgraduate training for staff, and frontline support including surveillance officers and diagnostic facilities, are critical (Oxfam, 2015).

Bilateral partnerships are an important mechanism through which to support capacity development. For example, the U.K. Department for International Development (DFID)–funded Health Partnership

Scheme, which is managed by the Tropical Health and Education Trust (THET), links National Health Service (NHS) healthcare facilities and academic bodies with overseas partners (DFID, 2013). There are also many regional surveillance platforms and partnerships that allow for an improved and integrated outbreak response, such as the Asia Pacific Strategy for Emerging Diseases, a common framework for countries in the region to develop coherent strategies appropriate to local and national contexts (WHO Regional Office for the Western Pacific & Regional Office for South East Asia, 2010).

Many other regional organizations and alliances, such as the Africa CDC and African Union, promote strategic and operational responses to novel threats across a region. The African Field Epidemiology Network (AFENET) is a public health network established in 2005 as a nonprofit networking alliance of Field Epidemiology and Laboratory Training Programs (FELTPs) and FETPs in Africa supporting ministries of health in health system strengthening in partnership with global public health experts. During the Ebola outbreak in West Africa, African Union Support for the Ebola Outbreak in West Africa (ASEOWA) supported FETP graduates from Uganda, Zimbabwe, Ethiopia, and Tanzania as they undertook the investigation and control of the outbreak in Liberia (Lubogo et al., 2015). Further, in the wake of the outbreak, the World Bank invested $390 million in West Africa in the Regional Disease Surveillance Systems Enhancement (REDISSE) program. The program is focused on enhancing surveillance and information systems, and strengthening laboratory capacity. The project aims to support the development of national surveillance systems and their interoperability at the different tiers of the health system, and establish networks of efficient, high-quality, accessible public health, veterinary, and private laboratories for the diagnosis of infectious human and animal diseases, and to create a regional networking platform to improve collaboration for laboratory investigation (World Bank, 2017c).

Finally, many bilateral research platforms contribute to research and training for future emerging infectious disease response. These initiatives include the Southeast Asia–Europe Joint Funding Scheme for Research and Innovation (SEA-EU-NET) and the European and Developing Countries Clinical Trials Partnership (EDCTP).

Funding for Public Health Emergency Response

A major constraint to ensuring emergency response infrastructure exists in LMICs is the (lack of) availability of funding. Such funds tend to be obtained through grants from international donors, as well as borrowed from international financial institutions such as the World Bank after a disaster has occurred (Ghesquiere & Mahul, 2007). These mechanisms are often slow to mobilize and, therefore, unlikely to provide the quick liquidity needed immediately after a disaster. An alternative financing strategy would be for LMICs to commit funding for emergency response mechanisms before a disaster, which can then be rapidly released when needed. Ideally, this would be sourced from domestic funds to be sustainable.

In 2008, the WHO Global Assessment of National Health Sector Emergency Preparedness and Response found that fewer than half of all national health sectors had a specific budget allocation for emergency preparedness and response (WHO, 2008a). The Global Facility for Disaster Reduction and Recovery and the Overseas Development Institute have estimated that between 1991 and 2010, only $13.5 billion (0.45%) of the total amount of international aid ($3.03 trillion) was spent worldwide on disaster prevention and preparedness (Kellett & Caravani, 2013).

EXHIBIT 14-5 outlines some recently proposed and implemented solutions for financing disaster risk reduction strategies.

EXHIBIT 14-5 Practical Solutions to Finance Disaster Risk Reduction

- The WHO Contingency Fund for Emergencies (CEF) provides immediate funding at the beginning of an emergency until resources from other mechanisms begin to flow (WHO, 2015a).
- The World Bank Pandemic Emergency Financing Facility (PEF) provides surge funding for low-income countries to assist response efforts to help prevent rare, high-severity disease outbreaks from becoming more deadly and costly pandemics (World Bank, 2017b).
- Some countries are engaging in disaster preparedness by working collaboratively and engaging local stakeholders. For example, the Horn of Africa Risk Transfer for Adaptation (HARITA) Program is an innovative microfinance scheme being piloted in Ethiopia (Oxfam America, 2010). HARITA offers a holistic approach to risk management, integrating financial aspects such as risk transfer (i.e., insurance), prudent risk taking (e.g., credit), and risk reserves (e.g., savings) alongside risk reduction activities (e.g., improved agricultural practices and conservation activities).

Evaluation of the International Health Regulations

The 2014–2016 outbreak of Ebola virus disease in West Africa laid bare the shortcomings of the global health security strategy and the IHR. The disaster exposed challenges that nations had previously failed to handle, including sustaining and deepening global health as a political priority, reforming WHO, and improving compliance with international health laws (Heymann et al., 2015).

The last issue has since been addressed by restructuring the IHR monitoring and evaluation process through the development of the Joint External Evaluation (JEE) tool (WHO, 2015f, 2016c). The JEE is a collaborative process to assess a country's capacity under the IHR to prevent, detect, and rapidly respond to public health threats. This tool is more comprehensive than the original IHR assessment, and was developed in partnership with the GHSA. A JEE assessment measures individual countries' IHR compliance status in 19 key areas (**FIGURE 14-2**); identifies the most urgent needs, opportunities for improved preparedness, and responses and action; and should engage with current and prospective donors and partners to target resources effectively (JEE Alliance, 2017a).

Complementing the GHSA partnership is the JEE Alliance, which facilitates engagement between all relevant actors involved in WHO's assessment process for IHR capacity strengthening. The JEE Alliance promotes use of the JEE tool, and links it with other assessment processes for health security. It promotes transparency in exchanging information on the results of assessments, in particular for donors interested in funding the development and strengthening of national-level capacities and capabilities. It is envisioned that a strengthened IHR monitoring and evaluation framework, with the JEE at its center, alongside after action reviews, simulations and exercises, and annual reporting, would make compliance and implementation easier to track (**FIGURE 14-3**).

The JEE is a voluntary process initiated at the request of the recipient country. It consists of an initial national self-assessment, which is followed by an external evaluation conducted by a team of international subject-matter experts. A report is then drafted to identify gaps, opportunities, and challenges in the implementation of the IHR. The final report is made publicly available by WHO, supporting transparency and openness of data and information sharing. This evaluation assists countries to identify the most urgent needs within their health systems and prioritize efforts for enhanced preparedness, response, and action, and it facilitates engagement with current and prospective partners to target resources effectively (JEE Alliance, 2017a). The collaborative and voluntary nature of the JEE helps to make the process politically more acceptable.

The JEE works alongside the OIE PVS tool (see the section on the Global Health Security Agenda). The latter assessment framework is designed to assist veterinary services to understand their "current level of performance, to identify gaps and weaknesses in their ability to comply with OIE international standards, to

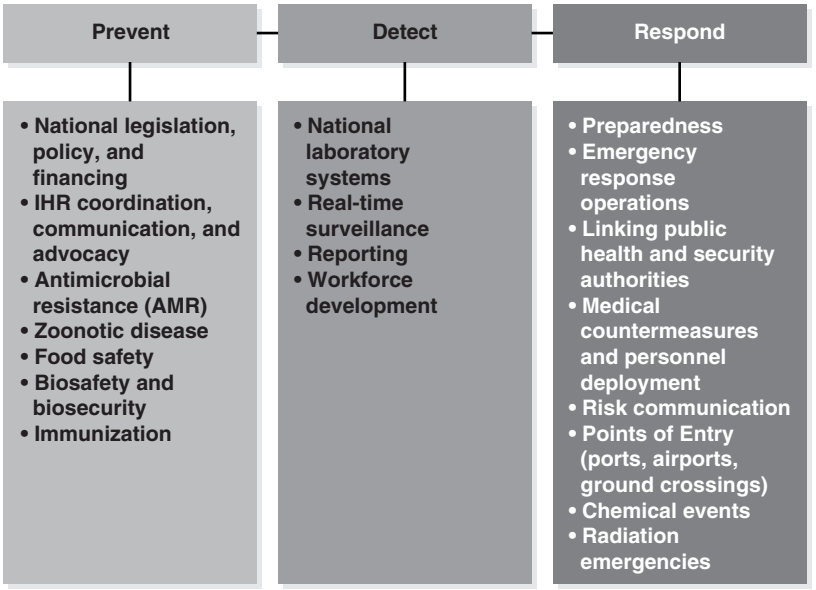


FIGURE 14-2 The 19 technical areas of the Joint External Evaluation.

Modified from World Health Organization (WHO). (2016c). *Joint external evaluation tool: International health regulations (2005)*. Geneva, Switzerland: Author. Retrieved from http://apps.who.int/iris/bitstream/10665/204368/1/9789241510172_eng.pdf. Copyright © 2016.

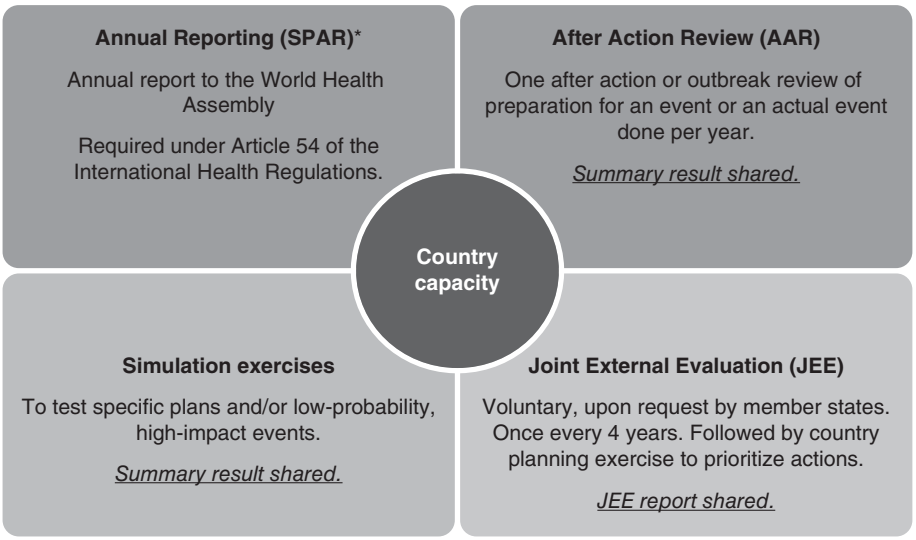


FIGURE 14-3 Complementary components of the IHR monitoring and evaluation framework.

Modified from World Health Organization (WHO). (2016). *WHO JEE SPP: From external evaluation to the development of action plans*. Jakarta, Indonesia: Author. Retrieved from https://www.ghsagenda.org/docs/default-source/default-document-library/archive-action-package-meeting/1—who-ihr-jee-spp-ludy-suryantoro-(r1)—508.pdf. Copyright © 2016.

form a shared vision with stakeholders (including the private sector), and to establish priorities and carry out strategic initiatives" (World Organization for Animal Health, 2017).

WHO covers most of the funds (via regional offices) for financing JEEs, although some countries are also able to contribute. Nevertheless, there is a need for funding to be specifically earmarked for JEEs and for the subsequent strengthening of health systems. The GHSA provides funding indirectly to the JEE through participation of its own staff in the evaluations. Further, some organizations and countries, notably Finland and the U.S. CDC, fund their own participation.

Post-JEE evaluation, countries must take action to address the gaps highlighted by the assessment. A comprehensive national plan of action should be devised and implemented. Implementation is the critical step, but the necessary activities need to be adequately funded. Inevitably, most LMICs will require additional financial resources to achieve the JEE targets and full IHR compliance, but often no concrete sources of funding or assistance are in place. This is currently a significant stumbling block to progress, as many countries are financially unable to follow through on the plans of action. Put simply, there is a need for comprehensive, clear implementation plans and sustainable funding sources to remedy this problem.

The collaborative and country-led nature of the JEE and post-JEE national action plan process is critically important. Building IHR capacity is a significant undertaking for any country, and particularly so for LMICs, where resources are scarce and competing priorities exist (Wilson, von Tigerstrom, & McDougall, 2008). An opportunity cost is incurred with any initiative, and IHR capacity building is no exception. Consequently, the IHR planning and prioritization process must be country led, with national experts deciding where to apportion their limited resources along the IHR spectrum of prevention, preparedness, detection (surveillance), and response. The collaborative process of the post-JEE national action planning, whereby national and external teams work together to set priorities and ensure synergies with existing broader national health strategies and development initiatives, supports this principle. **EXHIBIT 14-6** provides an innovative example of this collaboration in practice. To have lasting success, however, countries require sustained technical and financial inputs, which can be secured only through increased domestic financing and reduced reliance on external donors (Sands, Mundaca-Shah, & Dzau, 2016). By adopting a high-level, multisectoral approach, the JEE process intends to elevate IHR on national political agendas, paving the way for greater domestic financial commitment and sustainability.

The reformed IHR monitoring and evaluation process remains in its infancy. However, by May 2017 (just over a year after its launch), more than 40 countries had undergone a JEE assessment, with many more planned. The toolkit itself is iterative, with improvements and amendments already incorporated since it was first piloted, and plans for further refinements (JEE Alliance, 2017b).

Antimicrobial Resistance

Antimicrobial resistance (AMR) is recognized as a complex issue of increasing concern, threatening

EXHIBIT 14-6 Political and Institutional Economy Analysis of the Implementation of the IHR and Integrated Disease Surveillance and Response in Pakistan

Political and institutional economy analysis (PIEA) is a well-recognized development tool (DFID, 2009; McCulloch et al., 2017) that can be used to inform understanding and planning of health interventions. In 2016, Public Health England commissioned a PIEA of the implementation of WHO's IHR and the Integrated Disease Surveillance and Response (IDSR) in Pakistan (Public Health England & Contech International, 2017). The work was intended to answer a series of key questions at the federal level and in Punjab province, thereby facilitating planning of future activity. The PIEA addressed the following key questions:

- Why is Pakistan not compliant with the IHR?
- Why have previous recommendations for IDSR not been implemented?
- Why is progress slow?
- Who is responsible for IHR compliance and delivery of IDSR, and what influences them?
- What has or has not worked well to date, and why?

This work complemented a JEE in Pakistan undertaken in April–May 2016 (WHO Alliance for Country Assessment, 2016) and detailed situational analyses of existing disease surveillance and response systems in Pakistan (Public Health England & Contech International with the Punjab Government, 2016). It was the first time that formal political economy analysis had been used to study implementation of the IHR.

The findings of the PIEA from stakeholder interviews and desk review identified the following key governance issues:

- Expressed political commitment that had not translated into implementation
- Lack of a system-wide approach to the IHR and IDSR
- The need to improve the governance of delivery (e.g., defining roles, responsibilities, and standards, as well as systems for accountability and management)
- Deficient specialized workforce, HR planning, support, and career structures
- The role/approach of donors and technical agencies, with their dominance of the health landscape giving donors undue influence over domestic agendas and priority setting, undermining efforts for long-term sustainable system change
- Limited fiscal space for IHR compliance and IDSR implementation

Based on the PIEA diagnosis, the following recommendations and key actions were proposed to improve delivery of IHR capacities and an IDSR system:

- Ensure political prioritization of the IHR and IDSR so as to push the agenda at all levels, and improve communications around the IHR and IDSR to political leaders.
- Develop and deliver a system-wide approach to the IHR and IDSR, and coordinate institutions, partners, and individuals around this approach. IDSR strategic and operational plans should be developed in conjunction with provinces and wider stakeholders.
- Develop and put in place clear governance and accountability structures for individuals, organizations, and tiers of the health system for IDSR and delivery of IHR capacities with well-defined roles and responsibilities for delivery.
- Strengthen workforce planning, support, and career structures; develop specific eligibility and qualification standards for various cadres of the public health workforce; assess if there are immediate training needs.
- Encourage donors and technical agencies to work together and align their support to system-wide national plans for delivery of IHR capacities and IDSR. Governance mechanisms should be developed to assess whether donor assistance will support or hinder national plans.

These PIEA recommendations informed roadmap development and strategic and operational planning that occurred subsequent to the technical JEE assessment and IDSR situational analyses.

health and biosecurity across multiple sectors (OECD, 2017). It is an explicit technical area covered in the JEE, and increasing political awareness of this threat has placed AMR firmly on the global agenda, with numerous initiatives and policies established to tackle this growing crisis. Indeed, the Global Action Plan on Antimicrobial Resistance has called on member states to put into place national plans to combat AMR within 2 years (WHO, 2015d). The goal is to "ensure, for as long as possible, continuity of successful treatment and prevention of infectious diseases with effective and safe medicines that are quality-assured, used in a responsible way, and accessible to all who need them" (WHO, 2015d).

Since AMR is a shared, global problem, preserving antibiotic susceptibility will depend largely on the strength of the weakest national efforts. The development of comprehensive national action plans by LMICs hinges upon the availability of information pertaining to the etiology and prevalence of AMR in

these contexts. However, large information gaps in the global governance of AMR remain, concerning both the magnitude of the issue and national responses to it (Wernli et al., 2017). Surveillance data from LMICs are often scarce or incomplete (Tadesse et al., 2017), and even where epidemiologic data are collected, the full complement of information—particularly from private hospitals and laboratories—is often unavailable. Methodological inconsistencies in data collection and reporting further hamper comparability of data between laboratories, countries, and regions. Consequently, our understanding of the overall clinical and public health burden of AMR, as well as the state of national response efforts, is limited.

As will be discussed in more depth in the "Surveillance and Health Information Systems" section, it is important to strengthen the capacity of LMICs to monitor AMR. Efforts should also be made to ensure harmonization of data collection and reporting across countries, as this improves the comparability of information and enables countries to benchmark their performance against others. For example, laboratories in Europe are encouraged to use the European Committee on Antimicrobial Susceptibility Testing (EUCAST) breakpoints and methods to ensure consistent clinical reporting of antimicrobial susceptibility results and comparability for surveillance purposes (European Centre for Disease Prevention and Control, 2015). Improved situational monitoring may also result in greater engagement of health professionals and consumers, whose participation is critical to tackling AMR (Wernli et al., 2017). However, equally as important as improving the capacity of countries to monitor AMR, is deciding which factors should be measured in the first place. Given the diversity among and within countries (from the ecological determinants of AMR to social and political systems), it has been proposed that national monitoring efforts focus on three main components (Wernli et al., 2017):

- The epidemiologic situation and its impact on human health (outcomes)
- Drivers and practices (processes)
- Regulations and control policies (structures)

A variety of measurements, drawn from human, animal, and environmental health, can be attached to each of these components. For example, measurements of the drivers of AMR encompass access to sanitation and consumption of meat products, whereas the adoption of a national action plan and the enactment and implementation of regulations are measurements of structures.

AMR monitoring in LMICs can be bolstered through engagement with the GHSA, which supports country capacity to collect data. Another strategy is to exploit synergies with successful disease-specific surveillance programs (such as tuberculosis surveillance) to expand AMR surveillance. Data from each country should be collated and synthesized on a common platform, with methodologic differences addressed transparently to ensure comparability of results. As the information collected is aimed at informing policy, attention should be paid to communicating the findings clearly and widely through various media outlets (Wernli et al., 2017).

Improvements in AMR monitoring ultimately requires the engagement and participation of a number of global actors. For example, the recently launched Conscience of Antimicrobial Resistance Accountability (CARA) initiative aims to track national responses to AMR (Gelband, 2016). CARA comprises an alliance of multisectoral organizations (public, private, and nonprofit) from the human, animal, and environment sectors. Each member organization will commit to reporting on mutually agreed indicators in countries around the world, so as to monitor the state of national AMR responses.

Antimicrobial Stewardship and Infection Prevention and Control

Hand hygiene (as part of infection prevention and control [IPC]) and antimicrobial stewardship are thought to be two of the most effective evidence-based interventions in tackling AMR. Whilst antimicrobial stewardship programs (ASPs) are increasingly common in higher-income settings, they are more difficult to implement in LMICs due to multiple factors—for example, a lack of investment, human workforce, logistical infrastructure, and strong surveillance mechanisms (Bebell & Muiru, 2014), as well as lack of access to life-saving antimicrobials in some areas (Laxminarayan et al., 2016). The last issue is particularly important: Simultaneously expanding appropriate access to antimicrobials and restricting inappropriate access to these agents is a key challenge, requiring new approaches to financing and delivering healthcare services (addressed in detail in the *Pharmaceuticals* chapter).

Although antibiotic consumption is increasing worldwide, driven by economic growth and prosperity, access to these medications is not uniform. Indeed, more deaths are currently caused by limited access to antibiotics than by antibiotic resistance (Laxminarayan et al., 2016). For example, pneumonia-related deaths in children younger than age 5 years are strongly correlated with availability of antibiotics, and it is estimated that universal provision of antibiotics could reduce deaths due to community-acquired

pneumonia in this age cohort by 75%. Although ASPs that act to restrict the use of quality-controlled antimicrobials, while also enabling effective monitoring and offering alternative therapies (e.g., vaccinations), are essential in ensuring good antimicrobial stewardship (Buckland Merrett, 2013), such policies should be implemented with caution: Achieving child survival goals will depend on continued and even expanded access to effective antibiotics for pediatric patients (Laxminarayan et al., 2016).

Further, creative solutions are required that recognize the limited regulatory capacities in many LMICs. For example, banning the sale of nonprescription antibiotics is likely to be counterproductive—as well as unenforceable—because it would restrict access for poorer populations that rely on private drug sellers for healthcare (Laxminarayan & Heymann, 2012). In Nigeria, for example, drug vendors are the first source of care for as many as 55% of cases of under-5 illnesses, and in some parts of sub-Saharan Africa drug shops provide as much as 83% of all child health services (Webster, 2017).

New tools to provide highly specific diagnoses of pathogens can decrease diagnostic uncertainty, improving clinical management and stewardship efforts. Rapid, low-cost, and readily available diagnostic tools should be an essential component of ASPs. Currently, however, the vast majority of antimicrobial prescriptions are written without the use of a diagnostic tool. The decision as to whether an antimicrobial is prescribed is often based on an "empirical diagnosis," whereby expertise, intuition, and professional judgment are used to "guess" the causative agent of an infection. This behavior stems from the length of time traditional diagnostic tests take to confirm the offending pathogen (e.g., many types of bacteria must be cultured for a number of days before an infection can be confirmed). Acutely ill patients cannot afford to wait this long, and even in non-urgent cases time, patient, and financial pressures often force prescribers to address needs much faster (Review on Antimicrobial Resistance, 2016). As a consequence, huge quantities of antimicrobials are wasted globally on patients who do not need them (driving AMR) and on microbes against which they are ineffective. One solution to this problem is the development of rapid point-of-care diagnostic tests, as discussed in depth in the "Diagnostic Capability" section.

Surveillance and Health Information Systems

Disease surveillance is a cornerstone of the IHR (WHO, 2016b). Countries must be able to detect threats within their own borders if they are to implement effective control measures and forewarn other countries of potential risks. Data generated from both public health and clinical laboratories are essential for public health surveillance activities, and should feed into a national health information system. Countries need to establish a robust interoperable and interconnected electronic reporting system that can link and analyze multisectoral surveillance data including epidemiologic, clinical, laboratory, environmental testing, and bioinformatics data (GHSA, 2017b). An electronic reporting system capable of generating automatic reports or alerts will also facilitate real-time surveillance efforts. A strong health information system is necessary to track progress toward health-related goals, to monitor the performance of the health system, and to inform policy makers and health system managers. Unfortunately, health information systems in many LMICs are weak and fragmented, so they cannot supply sound data in a timely manner (Mbondji et al., 2014).

To facilitate information sharing and the subsequent analyses, laboratory protocols and data collection, coding, and storage need to be standardized. Within countries, data are often collected or produced by multiple actors and stored on various information management platforms. To streamline data sharing, identify gaps in data collection, and avoid duplication of efforts, there is a need to ensure greater complementarity and coordination between agencies.

For example, in India, health information data are collated by many different agencies and surveillance systems (Patel et al., 2015). Little coordination exists between these actors, and there is little integration and reconciliation of diverse data sources. Health workers often collect several types of health metrics to meet the divergent requirements of central, state, and local governments. Data collection is also incomplete and restricted to public health facilities. Failure to collect data from the private sector excludes the major provider of healthcare in India from the surveillance system. Recognizing these limitations, India's 2017 National Health Policy outlined the need for an integrated health information system, which necessitates private-sector participation (Ministry of Health and Family Welfare, 2017). The National Rural Health Mission has established an integrated nationwide health information portal that facilitates the flow of information including more than 300 data elements, and maintains a repository of published reports from national data sets (Patel et al., 2015). This system has been expanded to capture monthly information from 195,000 facilities across almost all states of India. The government aims to establish a federated national

health information architecture to link systems across public and private health providers at the state and national levels.

Surveillance is a particular issue for monitoring AMR, because AMR surveillance in LMICs is limited. The lack of data constrains national governments and international organizations in their efforts to devise action plans and detect evolving trends and emerging threats. However, a number of initiatives have been established to build surveillance capacity. For example, the Global AMR Surveillance System (GLASS) collects and reports data on AMR rates aggregated at the national level (WHO, 2015e). This system aims to collect comparable, validated data on AMR, which are then analyzed and shared to inform evidence-based action and decision making. Similarly, a roadmap has been created for implementation of this system in LMICs, as these countries may not have the resources or capacity to implement all components in the GLASS manual (London School of Hygiene and Tropical Medicine, 2016). Further, regional AMR surveillance networks have been established in Latin America (Red Latinoamericana de Vigilancia de la Resistencia a los Antimicrobianos, ReLAVRA) and Central Asia and Eastern Europe (CAESAR), and successes have been achieved in India through the National Programme on Antimicrobial Surveillance and in Kenya, South Africa, and Mozambique. Nevertheless, significant hurdles to AMR surveillance remain in many LMICs—namely, weak health systems, difficulty in enforcing regulations, healthcare worker recruitment and retention, poor public health infrastructure with a higher burden of infectious disease, and limited resources (Abdula, Macharia, Motsoaledi, Swaminathan, & VijayRaghavan, 2016). An integrated, multisectoral approach combining leadership, training, supervision, and laboratory quality assurance (including site visits and panel testing) would facilitate strengthened laboratory capacity in LMICs (Dacombe, Bates, Bhardwaj, Wallis, & Pulford, 2016).

Given the close ties between AMR and animal husbandry, and recognizing that the majority of infectious diseases are zoonoses (Jones et al., 2008), it is valuable to take an integrated or "One Health" approach to surveillance activities. WHO (2017c) defines One Health as "an approach to designing and implementing programmes, policies, legislation, and research in which multiple sectors communicate and work together to achieve better public health outcomes." This approach recognizes that the health of humans, animals, and ecosystems is interconnected (**FIGURE 14-4**), and has been gaining momentum ever since the Asian avian influenza outbreak in 2005. An added value of this holistic approach is the resultant shift in the paradigm of emergency response upstream, toward prevention and preparedness, through management of the risks that lead to disease emergence (Dixon, Dar, & Heymann, 2014; Heymann & Dar, 2014).

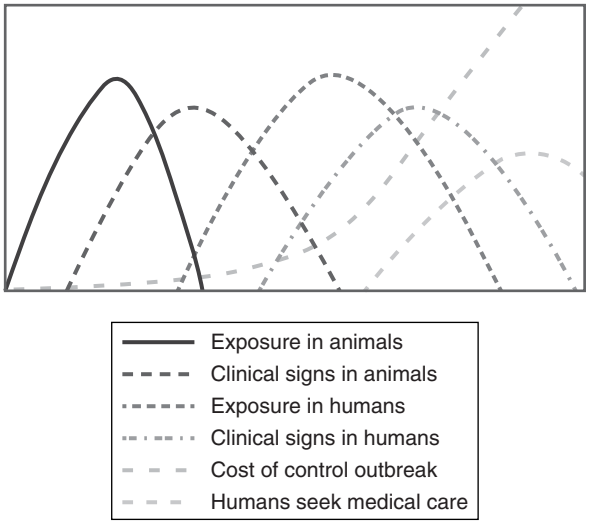


FIGURE 14-4 Early control of zoonotic infections is cost-effective and reduces human disease.

Reproduced from Heymann, D.L., & Dar, O.A. (2014). Prevention is better than cure for emerging infectious diseases. *BMJ, 348,* g1499. Copyright © 2014 with permission from BMJ Publishing Group Ltd.

One Health surveillance comprises the systematic collection, validation, analysis, interpretation, and dissemination of data collected on humans, animals, and the environment so as to inform decisions for more effective evidence- and system-based public health interventions (Stark et al., 2015). Integrating surveillance in this way can be achieved by aggregating databases at the human–animal interface. For example, the Global Early Warning System (GLEWS) database includes animal or zoonotic events for which information has been jointly gathered by the Food and Agriculture Organization (FAO), World Organization for Animal Health (OIE), and WHO, and confirmed by national authorities (GLEWS, 2017). Another example is the "four-way linking" platform implemented in Egypt, Vietnam, and Indonesia (Forcella et al., 2015). This project promotes the development of a framework for data sharing, risk assessment, and risk communication among governmental public health and animal health influenza laboratories as well as epidemiology units. The need to integrate human and animal interventions to better prevent, detect, and control human diseases is also reflected in the adoption and promotion of One Health approaches by the GHSA, IHR, and JEE (GHSA, 2017b; WHO, 2016c).

Experience from recent major outbreaks has shown the need for a wide array of surveillance systems and diagnostic capability, alongside the ability to implement emergency response mechanisms (Crawford, Rutz, & Evans, 2016). Novel surveillance

tools, such as mobile phone technology to track and monitor outbreaks (Jia & Mohamed, 2015), multimedia platforms including image and pattern recognition to monitor mass gatherings (Al-Salhie, Al-Zuhair, & Al-Wabil, 2014), and use of Internet search terms to predict outbreaks (Cook, Conrad, Fowlkes, & Mohebbi, 2011), are also likely to be applied more frequently in the future.

Networks such as the International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC, 2017b) are developing standardized data collection tools to provide for obtaining early information about novel disease progression and management. Countries increasingly will need to integrate with regional partners, such as ECDC, and through global exchanges of information and novel pathogens, such as with WHO collaborating centers. To support this type of cooperation, agreements are required at the global and regional levels on data and specimen sharing and, where relevant, shared intellectual property. The latter issue is particularly important when specimens originating in LMICs are used to develop diagnostics, vaccines, and therapeutics that are then distributed and sold globally through the private sector (Fidler, 2008).

As many countries move toward more privatized provision of diagnostic services (Corti, Passini, Lanzavecchia, & Zambon, 2016), steps are needed to ensure that reference services receive adequate and representative samples of pathogens. One way to ensure this access is to develop a mechanism whereby provision of appropriate samples leads to diagnostic and therapeutic development, leading to local deployment of novel tests or drugs. Integrated disease surveillance, including a One Health approach including animal health and laboratory infrastructure investment, is key to "rapid identification and containment of . . . emerging and re-emerging infectious diseases" (Houlihan, Youkee, & Brown, 2017).

Data Sharing

There is a growing consensus among stakeholders about the importance of data sharing in advancing health (Wellcome Trust, 2011). Sharing public health surveillance data enables regional collaborations, capacity strengthening, and insight into public health system performance. This, in turn, leads to improvements in risk management, and enhances public health responsiveness (Chatham House, 2017). Seven principles have been proposed to help create an environment conducive to data sharing (**FIGURE 14-5**). They are intended to facilitate good practice and encourage ethical sharing to the highest achievable standards, while also helping to identify opportunities for capacity building (Chatham House, 2017).

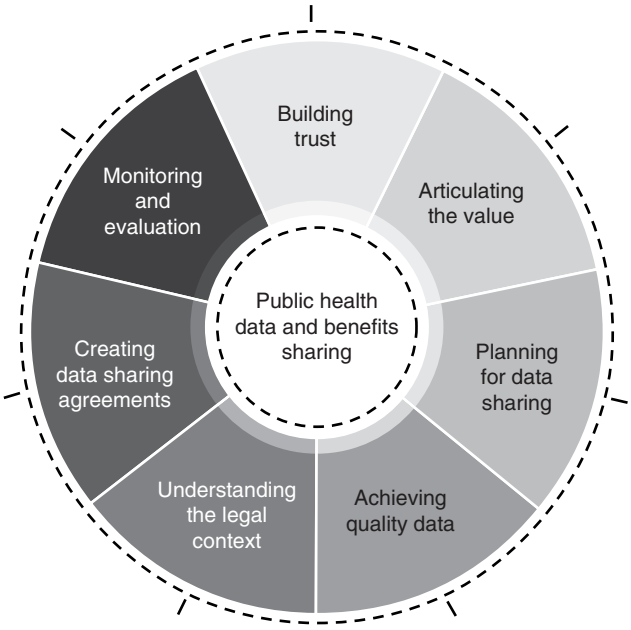


FIGURE 14-5 Data sharing principles for public health surveillance.

Reproduced from Chatham House, Royal Institute of International Affairs. (2017). *A guide to sharing the data and benefits of public health surveillance.* London, UK: Author. Copyright © The Royal Institute of International Affairs, 2017. With permission of RIIA through PLSclear.

Data sharing is particularly pertinent during a public health emergency, where it is critical to ensure the timely dissemination of information that might have value in combating the crisis. The ability to mobilize data-sharing networks enables researchers to more quickly develop and test hypotheses, provide evidence to inform the public health response, and thereby facilitate a prompt and effective global response (GloPID-R Data Sharing Working Group, 2017). Further, sharing data between stakeholders helps prevent the duplication of work, and pooling efforts both extends the value of limited resources and reduces the lag time between the identification of a public health emergency and the mounting of a response.

Many different types of data, ranging from clinical, epidemiologic, and pathogen genome data to information on experimental diagnostics, therapeutics, and vaccines, may contribute to the research response. The increasing depth and breadth of international collaborations, coupled with technological advancements including informatics, rapid whole-genome sequencing of pathogens, and novel methods to collect and analyze clinical and epidemiological data, present a huge potential to advance public health and transform responses to outbreaks (Wellcome Trust, 2011; Yozwiak, Schaffner, & Sabeti, 2015). However, the power of these "big data" to aid such responses can be realized only if the data are shared widely and promptly (Yozwiak et al., 2015).

Despite this necessity, in practice, the timely and transparent sharing of data remains a challenge. The shortcomings of existing data-sharing mechanisms were highlighted during the Ebola outbreak in West Africa (Whitty et al., 2015). In the early stages of the epidemic, individuals and organizations, including WHO, were reluctant to share data in real time. In these type of situations, the main reason given for withholding information is the perceived disincentives to share data, including concerns that data sharing may jeopardize subsequent publication, would allow other researchers to use the data for their own publications, and might violate confidentiality agreements. Additional reasons given for lack of sharing include the lack of incentives and the absence of a mechanism to enable information dissemination. This unwillingness to share data had implications for the effectiveness of the early response and put the lives of those HCWs on the frontline at risk, many of whom had provided the data in the first place. Further, the particularities of outbreaks can present additional barriers to data sharing. Outbreaks often bring together people and communities from all over the world, and who often have different norms and values. This "culture clash" can create uncertainties regarding ownership of data (Yozwiak et al., 2015). Further, in emergencies it can be difficult to obtain patient consent for data sharing, especially in situations, such as in the Ebola outbreak, where patients and their families are vulnerable to stigmatization and exploitation.

Thus, there is an obvious need to develop clear guidelines and initiatives to encourage and enable data sharing. In 2015, an agreement was reached on the need for data sharing, particularly during public health emergencies (WHO, 2015b). In this meeting, a consensus was reached that the risks and potential harms of nondisclosure of important information provide a strong ethical rationale for the rapid sharing of data. Among other things, it was recognized that while epidemiologic data belong to the country of origin, the default option is that data should be shared (i.e., opt-out policy) to ensure that knowledge generated becomes a global public good. It was also agreed that the disclosure of information should not be delayed by publication timelines, and that prepublication disclosure must not prejudice journal publication. Prepublication information sharing should become the global norm in the context of public health emergencies, with the exception of some conflict settings (Karunakara, 2013).

In implementing data-sharing initiatives, a number of issues need to be considered. First, incentives and safeguards should be created to encourage researchers to disclose their data (Yozwiak et al., 2015). Ethical, legal, and governance issues also need to be addressed to safeguard participants and communities, overcome regulatory hurdles, and mitigate challenges associated with benefit sharing and reciprocity. A data-sharing infrastructure needs to be developed to facilitate prompt dissemination of information, including the creation of networks and norms for repositories and best practices.

The Global Research Collaboration for Infectious Disease Preparedness (GloPID-R) is a platform that brings together research funding organizations on a global scale to facilitate an effective research response to a significant infectious disease outbreak (GloPID-R, 2015). Data sharing is a key priority of GloPID-R, and a working group has been established to develop systems for data sharing during public health emergencies (Littler, 2016). The data-sharing working group has also drafted a set of seven core principles to underpin and support the development of systems for data sharing in public health emergencies (GloPID-R Data Sharing Working Group, 2017); these principles are described in **TABLE 14-1**.

Another initiative is the Infectious Diseases Data Observatory (IDDO, 2017), which acts as a "central repository for evidence of optimal management and treatment efficacy for selected infectious diseases." This platform for collaboration facilitates the investigation of key scientific questions and rapid response to public health emergencies. It offers a variety of disease-specific data-sharing platforms where data can be collated, standardized, and stored securely.

ISARIC provides a collaborative platform for clinical researchers to share data and access protocols needed to facilitate a rapid response to outbreaks of emerging diseases (ISARIC, 2017a). ISARIC can be described as a network of networks having gathered in one place more than 70 networks and individuals. This organization also promotes the standardization and harmonization of definitions, endpoints, and data collection methods to streamline the aggregation, tabulation, and analysis of data across many different settings.

As in other responses to public health emergencies, preparedness is key to ensure successful data sharing during a public health crisis. The systems, standards, and principles of data sharing should already be in place as part of standard data-sharing practices. These mechanisms can then be scaled up or extended to address the specific challenges associated with data dissemination during public health emergencies (Littler, 2016). Not only does this expedite the process of data sharing in an emergency,

TABLE 14-1 GloPID-R Data Sharing Principles

Principles	Description
Timely	Data must be shared promptly with few access limitations. This may include the use of harmonized study protocols and the development of clear guidelines.
Ethical	Data sharing must abide by applicable ethical and legal standards. Informed consent models that allow for secondary use of data (broad consent) should be utilized.
Accessible	Data sharing should be subject to as few restrictions as possible. Data providers should clearly indicate which conditions are in place.
Transparent	Information outlining how data can be requested and how requests are considered should be provided, including timelines and conditions governing use and access.
Equitable	To ensure equal access, data should be made available to all parties at no cost, or at a cost-recovery level only.
Fairness	The provision and use of data must be done in such a way that ensures fair treatment of all parties involved and recognition of their contributions.
Quality	While speed is of the essence in public health emergencies, a minimum quality standard of data must be ensured by the provider.

Modified from GloPID-R Data Sharing Working Group. (2017). Principles for data sharing in public health emergencies. Version 1. GloPID-R. Retrieved from https://s3-eu-west-1.amazonaws.com/pstorage-wellcome-4792389547823904/7897831/20170330PrinciplesforDataSharinginPublicHealthEmergenciesAdobe.pdf

but it also builds advance confidence in the system. It is difficult to rapidly cultivate trust in the context of an emergency; therefore, it is crucial that such relationships are already in place to facilitate data sharing (GloPID-R Data Sharing Working Group, 2017). Further, implicit in the principle of fairness mentioned earlier is a need for a circular model of data collection, in which the benefits derived from the sharing of data feed back to the community from which the data are derived. This also fosters trust and confidence in the system, in turn facilitating the future collection and sharing of data as people and communities are more likely to be willing to engage with researchers.

Conclusion

While population health has improved greatly in the past few decades, major health disparities persist around the world. In the majority of countries, the underlying cause of these disparities is an inadequate health system. Public health infrastructure remains the poor relation of healthcare services within the health systems of LMICs. It tends to be less visible than the curative services, and so often does not garner much political support, despite some clear examples of the popularity among the electorate of an improved health system and strong economic arguments for investing in public health initiatives (Jamison et al., 2013). While there have been promising advances in strengthening essential public health, most of the recent attention has focused on enhancing health security. While this approach to public health has a long history, it has been stimulated recently by a series of serious epidemics that have amply demonstrated how unprepared is the world to respond to major outbreaks and crises. While these health security initiatives are welcome, they are by their nature more limited in scope than a human rights–derived public health agenda. There is an urgent need to work to align and consolidate these two public health approaches to provide a comprehensive strengthening of public health infrastructure for the future.

Acknowledgments

We wish to thank Ala Alwan, Colin Brown, Emmeline Buckley, Katie Carmichael, Supamit Chunsuttiwat, Mishal Khan, Annette Luker, Rishma Maini, Brian McCloskey, and Monica Sethi for their contributions and insights into the writing of this chapter.

Discussion Questions

1. How valid are concerns that public health has received inadequate attention and resources relative to healthcare, and that this bias in favor of healthcare has limited the potential for population health gains? Give examples to support your position.
2. Which evidence-based arguments might you make to show that strengthening public health infrastructure is an essential component for achieving the Sustainable Development Goals?
3. Discuss the challenges to strengthening public health systems at the district and community levels. Which approaches can be used to address these challenges?
4. Which policies and actions could support the creation of a resilient public health system in a LMIC?
5. Discuss ways in which the reformed International Health Regulations' monitoring and evaluation process can improve national public health systems. Outline gaps in the process that could be addressed for comprehensive strengthening of essential public health functions.

References

Abdula, N., Macharia, J., Motsoaledi, A., Swaminathan, S., & VijayRaghavan, K. (2016). National action for global gains in antimicrobial resistance. *The Lancet, 387*(10014), e3–e5.

Acheson, E. D. (1988). *Public health in England: The report of the Committee of Inquiry into the Future Development of the Public Health Function. Cmnd 289*. London, UK: Her Majesty's Stationery Office.

Adini, B., Goldberg, A., Cohen, R., Laor, D., & Bar-Dayan, Y. (2012). Evidence-based support for the all-hazards approach to emergency preparedness. *Israel Journal of Health Policy Research, 1*(40).

Admasu, K. (2016). Designing a resilient national health system in Ethiopia: The role of leadership. *Health Systems and Reform, 2*(3), 182–186.

Aitsi-Selmi, A., Blanchard, K., & Murray, V. (2016). Ensuring science is useful, usable and used in global disaster risk reduction and sustainable development: A view through the Sendai framework lens. *Palgrave Communications, 2*(16016).

Aitsi-Selmi, A., & Murray, V. (2016a). The Chernobyl disaster and beyond: Implications of the Sendai Framework for Disaster Risk Reduction 2015–2030. *PLoS Medicine, 13*(4), e1002017.

Aitsi-Selmi, A., & Murray, V. (2016b). Protecting the health and well-being of populations from disasters: Health and health care in the Sendai Framework for Disaster Risk Reduction 2015–2030. *Prehospital and Disaster Medicine, 31*(1), 74–78.

Aitsi-Selmi, A., Murray, V., Wannous, C., Dickinson, C., Johnston, D., Kawasaki, A., . . . Yeung, T. (2016). Reflections on a science and technology agenda for 21st century disaster risk reduction. Based on the scientific content of the 2016 UNISDR Science and Technology Conference on the Implementation of the Sendai Framework for Disaster Risk Reduction 2015–2030. *International Journal of Disaster Risk Science, 7*(1), 1–29.

Alford, J. (2009). *Engaging public sector clients: From service-delivery to co-production*. London, UK: Palgrave Macmillan UK.

Al-Salhie, L., Al-Zuhair, M., & Al-Wabil, A. (2014). Multimedia Surveillance in Event Detection: Crowd Analytics in Hajj. In A. Marcus (Ed.), *Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments: Third International Conference, DUXU 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22–27, 2014, Proceedings, Part II* (pp. 383–392). Cham: Springer International Publishing.

Alsan, M., Schoemaker, L., Eggleston, K., Kammili, N., Kolli, P., & Bhattacharya, J. (2015). Out-of-pocket health expenditures and antimicrobial resistance in low-income and middle-income countries: An economic analysis. *The Lancet Infectious Diseases, 15*(10), 1203–1210. http://dx.doi.org/10.1016/S1473-3099(15)00149-8

Anderies, J. M., Walker, B. H., & Kinzig, A. P. (2006). Fifteen weddings and a funeral: Case studies and resilience-based management. *Ecology and Society, 11*(1), 21.

Balabanova, D., McKee, M., & Mills, A. (Eds.). (2011). *"Good health at low cost" 25 years on: What makes a successful health system?* London, UK: London School of Hygiene and Tropical Medicine.

Beaglehole, R., & Bonita, R. (2004). *Public health at the crossroads: Achievements and prospects*. Cambridge, UK: Cambridge University Press.

Bebell, L., & Muiru, A. (2014). Antibiotic use and emerging resistance. *Global Heart, 9*(3), 347–358.

Bellagio District Public Health Workshop Participants. (2016). *Public health performance strengthening at districts: Rationale and blueprint for action.* Bellagio, Italy.

Béné, C., Wood, R. G., Newsham, A., & Davies, M. (2012). Resilience: New utopia or new tyranny? Reflection about the potentials and limits of the concept of resilience in relation to vulnerability reduction programmes. *IDS Working Papers, 2012*(405), 1–61. doi:10.1111/j.2040-0209.2012.00405.x

Berman, P., Ahuja, R., & Bhandari, L. (2010). The impoverishing effect of healthcare payments in India: New methodology and findings. *Economic and Poltical Weekly, XLV*(16), 65–71.

Bettcher, D. W., Sapirie, S., & Goon, E. H. T. (1998). Essential public health functions: Results of the international Delphi study. *World Health Statistics Quarterly, 51*(1), 44–54.

Bishai, D. M., Sherry, M., Pereira, C., Chicumbe, S., Mbofana, F., Boore, A., . . . Borse, N. N. (2016). Development and usefulness of a district health systems tool for performance improvement in essential public health functions in Botswana and Mozambique. *Journal of Public Health Management & Practice, 22*(6), 586–596.

Blas, E., Sommerfeld, J., & Kurup, A. S. (2011). *Social determinants approaches to public health: From concept to practice.* Geneva, Switzerland: World Health Organization.

Boeras, D. I., Nkengasong, J. N., & Peeling, R. W. (2017). Implementation science: The laboratory as a command centre. *Current Opinion in HIV & AIDS, 12*(2), 171–174.

Bradley, N., Meara, J., & Murray, V. (2015). Principles of public health emergency response for acute environmental, chemical, and radiation incidents. In R. Detels, M. Gulliford, Q. A. Karim, & C. C. Tan (Eds.), *Oxford textbook of global public health* (6th ed., pp. 1608–1620). Oxford, UK: Oxford University Press.

Broughton, E. (2005). The Bhopal disaster and its aftermath: A review. *Environmental Health, 4*(1), 6.

Brown, C. S., Elsherbiny, M., Johnson, O., Kamboz, A., Lado, M., Leather, A., . . . Ngauja, R. (2016). Ebola and health partnerships: Action in a time of crisis. In C. Podlipnik (Ed.), *Ebola* (Chapter 02). Rijeka, Croatia: InTech.

Buckland Merrett, G. (2013). *Tackling antibiotic resistance for greater global health security.* London, UK: Chatham House, Royal Institute of International Affairs.

Burns, A., van der Mensbrugghe, D., & Timmer, H. (2008). *Evaluating the economic consequences of avian influenza.* Washington, DC: World Bank.

Cabinet Office. (2011). *Keeping the country running: Natural hazards and infrastructure. A guide to improving the resilience of critical infrastructure and essential services.* London, UK: Author.

Centers for Disease Control and Prevention (CDC). (2014). The public health system and the 10 essential public health services. Retrieved from https://www.cdc.gov/stltpublichealth/publichealthservices/essentialhealthservices.html

Centers for Disease Control and Prevention (CDC). (2016). One Health: Global activities. Retrieved from https://www.cdc.gov/onehealth/global-activities/index.html

Chatham House, Royal Institute of International Affairs. (2017). *A guide to sharing the data and benefits of public health surveillance.* London, UK: Author.

Cheng, T. M. (2008). China's latest health reforms: A conversation with Chinese health minister Chen Zhu. *Health Affairs, 27*(4), 1103–1110.

Chongvilaivan, A. (2012). *Thailand's 2011 flooding: Its impact on direct exports and global supply chains.* ARTNeT Working Paper Series, No. 113. Bangkok, Thailand: ESCAP.

Coeytaux, K., Bey, E., Christensen, D., Glassman, E. S., Murdock, B., & Doucet, C. (2015). Reported radiation overexposure accidents worldwide, 1980–2013: A systematic review. *PLoS One, 10*(3), e0118709.

Commission on a Global Health Risk Framework for the Future. (2016) *The neglected dimension of global security: A framework to counter infectious disease crises.* Washington, DC: National Academies Press.

Commission on Social Determinants of Health. (2008). *Closing the gap in a generation: Health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health.* Geneva, Switzerland: World Health Organization.

Cook, S., Conrad, C., Fowlkes, A. L., & Mohebbi, M. H. (2011). Assessing Google flu trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic. *PLoS One, 6*(8), e23610. doi:10.1371/journal.pone.0023610

Cornish, P. (2007). *The CBRN system: Assessing the threat of terrorist use of chemical, biological, radiological and nuclear weapons in the United Kingdom. International Security Programme Report.* London, UK: Chatham House, Royal Institute of International Affairs.

Corti, D., Passini, N., Lanzavecchia, A., & Zambon, M. (2016). Rapid generation of a human monoclonal antibody to combat Middle East respiratory syndrome. *Journal of Infection and Public Health, 9*(3), 231–235. doi:10.1016/j.jiph.2016.04.003

Crawford, R., Rutz, D. C., & Evans, D. P. (2016). "Between combat boots and Birkenstocks": Lessons from HIV/AIDS, SARS, H1N1 and Ebola. *Public Health, 141*, 186–191.

Crisp, N. (2010). *Turning the world upside down: The search for global health in the 21st century*. London, UK: Royal Society of Medicine Press.

Cutter, S. L., & Gall, M. (2015). Sendai targets at risk. *Nature Climate Change, 5*, 707–709.

Cyril, S., Smith, B. J., Possamai-Inesedy, A., & Renzaho, A. M. N. (2015). Exploring the role of community engagement in improving the health of disadvantaged populations: A systematic review. *Global Health Action, 8*(1), 29842.

Dacombe, R., Bates, I., Bhardwaj, M., Wallis, S., & Pulford, J. (2016). *Fleming Fund: Supporting surveillance capacity for antimicrobial resistance. An analysis of approaches to laboratory capacity strengthening for drug resistant infections in low and middle income countries.* Liverpool, UK: Liverpool School of Tropical Medicine.

Dar, O., Buckley, E. J., Rokadiva, S., Huda, Q., & Abrahams, J. (2014). Integrating health into disaster risk reduction strategies: Key considerations for success. *American Journal of Public Health, 104*(10), 1811–1816. doi:10.2105/AJPH.2014.302134

Department for International Development (DFID). (2009). *Political economy analysis: How to note.* London, UK: Author.

Department for International Development (DFID). (2011). *Humanitarian emergency response review.* London, UK: Author.

Department for International Development (DFID). (2013). Health links scheme. Retrieved from https://www.gov.uk/guidance/health-partnership-scheme

Dixon, M., Dar, O., & Heymann, D. (2014). Emerging infectious diseases: Opportunities at the human–animal–environment interface. *Veterinary Record, 174*(22), 546–551.

Downer, J. (2009). *When failure is an option: Redundancy, reliability and regulation in complex technical systems.* London, UK: Centre for Analysis of Risk and Regulation, London School of Economics and Political Science.

Drosten, C., Gunther, S., Preiser, W., van der Werf, S., Brodt, H. R., Becker, S., . . . Doerr, H. W. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *New England Journal of Medicine, 348*(20), 1967–1976. doi:10.1056/NEJMoa030747

Elston, J. W. T., Moosa, A. J., Moses, F., Walker, G., Dotta, N., Waldman, R. J., & Wright, J. (2015). Impact of the Ebola outbreak on health systems and population health in Sierra Leone. *Journal of Public Health, 38*(4), 673–678.

European Centre for Disease Prevention and Control (ECDC). (2015). *Antimicrobial resistance surveillance in Europe 2014. Annual report of the European Antimicrobial Resistance Surveillance Network (EARS-Net).* Stockholm, Sweden: Author.

European Commission. (2016). Major accident hazards: The Seveso Directive: Summary of requirements. Retrieved from http://ec.europa.eu/environment/seveso/legislation.htm

Fan, L. (2013). *Disaster as opportunity? Building back better in Aceh, Myanmar and Haiti.* London, UK: Overseas Development Institute.

Fidler, D. P. (2008). Influenza virus samples, international law, and global health diplomacy. *Emerging Infectious Diseases, 14*(1), 88–94.

Field Epidemiology Training Program (FETP). (2017). History of FETP-Thailand. Retrieved from http://www.interfetpthailand.net/about_index.php

Forcella, S., El-din El Tantawy, N., Yilma, J., AbdelNabi, A., Claes, F., Dauphin, G., & Mumford, E. (2015). The development of a four-way linking framework in Egypt: An example of the FAO, OIE and WHO joint activities to facilitate national risk assessment. *Veterinaria Italiana, 51*(1), 45–50. doi:10.12834/VetIt.220.680.1

Fried, L. P. (2015). Public health and the UN's Sustainable Development Goals. Retrieved from http://www.huffingtonpost.com/linda-p-fried/public-health-and-the-uns_b_8739212.html

Garrett, L. (2015). Ebola's lessons: How the WHO mishandled the crisis. Retrieved from https://www.foreignaffairs.com/articles/west-africa/2015-08-18/ebolas-lessons?campaign=Garrett

Gelband, H. (2016). CARA: The Conscience of Antimicrobial Resistance Accountability, and the next big thing for CDDEP. Retrieved from https://www.cddep.org/blog/posts/cara_conscience_antimicrobial_resistance_accountability_and_next_big_thing_cddep/

Ghesquiere, F., & Mahul, O. (2007). *Sovereign natural disaster insurance for developing countries: A paradigm shift in catastrophe risk financing.* World Bank Policy Research Working Paper. Washington, DC: World Bank.

Gilson, L., Barasa, E., Nxumalo, N., Cleary, S., Goudge, J., Molyneux, S., . . . Lehmann, U. (2017). Everyday resilience in district health systems: Emerging insights from the front lines in Kenya and South Africa. *BMJ Global Health, 2*, e000224. doi:10.1136/bmjgh-2016-000224

GLEWS. (2017). About GLEWS and GLEWS+. Retrieved from http://www.glews.net/

Global Health Security Agenda (GHSA). (2017a). About Global Health Security Agenda. Retrieved from https://www.ghsagenda.org/about

Global Health Security Agenda (GHSA). (2017b). Action Packages. Retrieved from https://www.ghsagenda.org/packages

GloPID-R. (2015). Global research collaboration for infectious disease preparedness. Retrieved from https://www.glopid-r.org/

GloPID-R Data Sharing Working Group. (2017). Principles for data sharing in public health emergencies. Version 1. GloPID-R. Retrieved from https://s3-eu-west-1.amazonaws.com/pstorage-wellcome-4792389547823904/7897831/20170330PrinciplesforDataSharinginPublicHealthEmergenciesAdobe.pdf

Gostin, L. O., DeBartolo, M. C., & Friedman, E. A. (2015). The International Health Regulations 10 years on: The governing framework for global health security. *The Lancet, 386*(10009), 2222–2226.

Government of India. (1986). *Environmental (Protection) Act.* New Delhi: Ministry of Environment and Forests, Government of India.

Grepin, K. A. (2015). International donations to the Ebola virus outbreak: Too little, too late? *BMJ, 350*, h376. doi:10.1136/bmj.h376

Gunzenhauser, J. D., Eggena, Z. P., Fielding, J. E., Smith, K. N., Jacobson, D. M., & Bazini-Barakat, N. (2010). The quality improvement experience in a high-performing local health department: Los Angeles County. *Journal of Public Health Management & Practice, 16*(1), 39–48. doi:10.1097/PHH.0b013e3181bedcf1

Hasegawa, A., Tanigawa, K., Ohtsuru, A., Yabe, H., Maeda, M., Shigemura, J., . . . Ishikawa, T. (2015). Health effects of radiation and other health problems in the aftermath of nuclear accidents, with an emphasis on Fukushima. *The Lancet, 386*(9992), 479–488.

Hasin, Y., Seldin, M., & Lusis, A. (2017). Multi-omics approaches to disease. *Genome Biology, 18*(1), 83.

Heymann, D. L., Chen, L., Takemi, K., Fidler, D. P., Tappero, J. W., Thomas, M. J., . . . Rannan-Eliya, R. P. (2015). Global health security: The wider lessons from the West African Ebola virus disease epidemic. *The Lancet, 385*, 1884–1901.

Heymann, D. L., & Dar, O. A. (2014). Prevention is better than cure for emerging infectious diseases. *BMJ, 348*, g1499.

Hill, P. S., Pavignani, E., Michael, M., Murru, M., & Beesley, M. E. (2014). The "empty void" is a crowded space: Health service provision at the margins of fragile and conflict affected states. *Conflict and Health, 8*, 20. doi:10.1186/1752-1505-8-20

HM Government. (2010). *The United Kingdom's strategy for countering chemical, biological, radiological and nuclear (CBRN) terrorism.* London, UK: Author.

Houlihan, C. F., Youkee, D., & Brown, C. S. (2017). Novel surveillance methods for the control of Ebola virus disease. *Int Health, 9*(3), 139–141. doi:10.1093/inthealth/ihx010

House of Commons Science and Technology Committee. (2016). *Science in emergencies: UK lessons from Ebola.* London, UK: Her Majesty's Stationery Office.

Infectious Disease Data Observatory. (2017). Our work. Retrieved from https://www.iddo.org/about-us/our-work

Intergovernmental Panel on Climate Change (IPCC). (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change.* Cambridge, UK: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC). (2014). *Climate change 2014: Synthesis report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Geneva, Switzerland: Author.

International Atomic Energy Agency. (1988). *The radiological incident in Goiânia.* Vienna, Austria: Author.

International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC). (2017a). Home. Retrieved from https://isaric.tghn.org/

International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC). (2017b). Protocols & data tools. Retrieved from https://isaric.tghn.org/protocols/

Jamison, D. T., Summers, L. H., Alleyne, G., Arrow, K. J., Berkley, S., Binagwaho, A., . . . Yamey, G. (2013). Global health 2035: A world converging within a generation. *The Lancet, 382*(9908), 1898–1955.

JEE Alliance. (2017a). Joint External Evaluation (JEE). Retrieved from https://www.jeealliance.org/global-health-security-and-ihr-implementation/joint-external-evaluation-jee/

JEE Alliance. (2017b, May 20). *Summary of Alliance for Country Assessments for Global Health Security and IHR implementation meeting*, Geneva, Switzerland: Author.

Jia, K., & Mohamed, K. (2015). Evaluating the use of cell phone messaging for community Ebola syndromic surveillance in high risked settings in Southern Sierra Leone. *African Health Sciences, 15*(3), 797–802. doi:10.4314/ahs.v15i3.13

Jonas, O. B. (2014). *Pandemic risk.* Washington, DC: World Bank.

Jones, K., Patel, N., Levy, M., Storeygard, A., Balk, D., Gittleman, J., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature, 451*(7181), 990–993.

Karunakara, U. (2013). Data sharing in a humanitarian organization: The experience of Medecins Sans Frontieres. *PLoS Medicine, 10*(12), e1001562. doi:10.1371/journal.pmed.1001562

Katz, R., & Fischer, J. (2010). The revised International Health Regulations: A framework for global pandemic response. *Global Health Governance, 2*(3).

Kellett, J., & Caravani, A. (2013). *Financing disaster risk reduction: A 20 year story of international aid.* London, UK: Global Facility for Disaster Reduction and Recovery & Overseas Development Institute.

Kennedy, S. B., & Nisbett, R. A. (2015). The Ebola epidemic: A transformative moment for global health. *Bulletin of the World Health Organization, 93*(1), 2. doi:10.2471/blt.14.151068

Kenny, A., Hyett, N., Sawtell, J., Dickson-Swift, V., Farmer, J., & O'Meara, P. (2013). Community participation in rural health: A scoping review. *BMC Health Services Research, 13*(64).

Kruk, M. E., Myers, M., Varpilah, S. T., & Dahn, B. T. (2015). What is a resilient health system? Lessons from Ebola. *The Lancet, 385*, 1910–1912.

Kulig, J. C. (2011). *Health in rural Canada*. Vancouver, Canada: University of British Columbia Press.

Kumar, S., Bothra, V., & Mairembam, D. S. (2016). A dedicated public health cadre: Urgent and critical to improve health in India. *Indian Journal of Community Medicine, 41*(4), 253–255.

Laksanasopin, T., Guo, T. W., Nayak, S., Sridhara, A. A., Xie, S., Olowookere, O. O., . . . Sia, S. K. (2015). A smartphone dongle for diagnosis of infectious diseases at the point of care. *Science Translational Medicine, 7*(273), 273re271. doi:10.1126/scitranslmed.aaa0056

Laxminarayan, R., & Heymann, D. L. (2012). Challenges of drug resistance in the developing world. *BMJ, 344*, e1567. doi:10.1136/bmj.e1567

Laxminarayan, R., Matsoso, P., Pant, S., Brower, C., Rottingen, J. A., Klugman, K., & Davies, S. (2016). Access to effective antimicrobials: A worldwide challenge. *The Lancet, 387*(10014), 168–175. doi:10.1016/s0140-6736(15)00474-2

Littler, K. (2016). GloPID-R Data Sharing WG: Addressing the challenges of data sharing in a public health emergency. Retrieved from https://www.glopid-r.org/wp-content/uploads/2016/12/zika-workshop-2016-day-1-katherine-littler.pdf

London School of Hygiene and Tropical Medicine. (2016). *AMR surveillance in low- and middle-income settings: A roadmap for participation in the Global Antimicrobial Surveillance System (GLASS)*. London, UK: Author.

Longitude Prize. (2017). The challenge: Reduce the use of antibiotics. Retrieved from https://longitudeprize.org/challenge

Lubogo, M., Donewell, B., Godbless, L., Shabani, S., Maeda, J., Temba, H., . . . Berhanu, N. (2015). Ebola virus disease outbreak: The role of field epidemiology training programme in the fight against the epidemic, Liberia, 2014. *Pan African Medical Journal, 22*(suppl. 1), 5. doi:10.11694/pamj.supp.2015.22.1.6053

Marchal, V., Dellink, R., Van Vuuren, D., Clapp, C., Chateau, J., Magné, B., & van Vliet, J. (2011). *OECD environmental outlook to 2050*. Paris, France: Organisation for Economic Co-operation and Development.

Martineau, F. P. (2016). People-centred health systems: Building more resilient health systems in the wake of the Ebola crisis. *International Health, 8*(5), 307–309.

Mazzola, L. T., & Pérez-Casas, C. (2015). *HIV/AIDS diagnostics technology landscape* (5th ed.). Geneva, Switzerland: UNITAID Secretariat, World Health Organization.

Mbondji, P. E., Kebede, D., Soumbey-Alley, E. W., Zielinski, C., Kouvividila, W., & Lusamba-Dikassa, P. S. (2014). Health information systems in Africa: Descriptive analysis of data sources, information products and health statistics. *Journal of the Royal Society of Medicine, 107*(1 suppl.), 34–45. doi:10.1177/0141076814531750

McCulloch, N., Barnett, A., Duncan, A., Kingsmill, W., Kydd, J., Owen, O., & Williams, G. (2017). *The evolving role of political economy analysis in development practice. Policy Practice Brief 11*. Brighton: The Policy Practice.

Médecins Sans Frontières. (2014). Guinea: Ebola epidemic declared, MSF launches emergency response. Retrieved from http://www.msf.org/article/guinea-ebola-epidemic-declared-msf-launches-emergency-response

Mills, A. (2014). Health care systems in low- and middle-income countries. *New England Journal of Medicine, 370*(6), 552–557.

Ministry of Health and Family Welfare. (2014a). De-linking of AYUSH Department from the Health Ministry [Press release]. Retrieved from http://pib.nic.in/newsite/PrintRelease.aspx?relid=113440

Ministry of Health and Family Welfare. (2014b). National Ayush Mission (NAM) [Press release]. Retrieved from http://pib.nic.in/newsite/PrintRelease.aspx?relid=113941

Ministry of Health and Family Welfare. (2017). National health policy 2017: Government of India. Retrieved from https://www.mohfw.gov.in/sites/default/files/9147562941489753121.pdf

Moon, S., Sridhar, D., Pate, M. A., Jha, A. K., Clinton, C., Delaunay, S., . . . Piot, P. (2015). Will Ebola change the game? Ten essential reforms before the next pandemic. The report of the Harvard–LSHTM Independent Panel on the Global Response to Ebola. *The Lancet, 386*(10009), 2204–2221.

Ndihokubwayo, J. B., Kasolo, F., Yahaya, A. A., & Mwenda, J. (2010). Strengthening public health laboratories in the WHO African Region: A critical need for disease control. *African Health Monitor, 12,* 47–52.

Nicholson, D., Yates, R., Warburton, W., & Fontana, G. (2015). *Delivering universal health coverage: A guide for policymakers.* Doha, Qatar: The World Innovation Summit for Health (WISH) Universal Health Coverage Forum.

Nkengasong, J. N. (2017). Establishing the Africa Centres for Disease Control and Prevention: Responding to Africa's health threats. *The Lancet, 5*, e246–e247.

Nyenswah, T., Engineer, C. Y., & Peters, D. H. (2016). Leadership in times of crisis: The example of Ebola virus disease in Liberia. *Health Systems and Reform, 2*(3), 194–207.

Ochi, S., Hodgson, S., Landeg, O., Mayner, L., & Murray, V. (2014). Disaster-driven evacuation and medication loss: A systematic literature review. *PLoS Currents Disasters, 18*(6), ecurrents.dis.fa417630b417566a417630c417637dfdbf945910edd417696.

Organisation for Economic Co-operation and Development (OECD). (2017). Antimicrobial resistance. Retrieved from http://www.oecd.org/els/health-systems/antimicrobial-resistance.htm

Oxfam. (2015). *Never again: Building resilient health systems and learning from the Ebola crisis.* Oxford, UK: Author.

Oxfam America. (2010). *Horn of Africa Risk Transfer for Adaptation. HARITA project report: November 2007—December 2009. Rural Resilience Series*. Boston, MA: Author.

Pacific Open Learning Health Net (POLHN). (2017). About POLHN. Retrieved from https://www.polhn.org/about-polhn/

Pacific Paramedical Training Centre (PPTC). (2017). Home. Retrieved from http://pptc.org.nz/

Patel, V., Parikh, R., Nandraj, S., Balasubramaniam, P., Narayan, K., Paul, V. K., . . . Reddy, K. S. (2015). Assuring health coverage for all in India. *The Lancet, 386*, 2422–2435.

Powles, J., & Comim, F. (2003). Public health infrastructure and knowledge. In R. Smith, R. Beaglehole, D. Woodward, & N. Drager (Eds.), *Global public goods for health: Health economic*

and public health perspectives (pp. 159–175). Oxford, UK: Oxford University Press.

Public Health England & Contech International. (2017). *Political & institutional economy analysis. WHO International Health Regulations (2005) and Integrated Disease Surveillance & Response in Pakistan.* London, UK: Public Health England.

Public Health England & Contech International with the Punjab Government. (2016). *Integrated Disease Surveillance and Response System—Punjab: Situational analysis.* London, UK: Public Health England.

Public Health Foundation of India (PHFI). (2015). About PHFI. Retrieved from https://www.phfi.org/about-us

Reece, S., Brown, C. S., Dunning, J., Chand, M. A., Zambon, M. C., & Jacobs, M. (2017). The UK's multidisciplinary response to an Ebola epidemic. *Clinical Medicine (London), 17*(4), 332–337. doi:10.7861/clinmedicine.17-4-332

Reich, M. R., Javadi, D., & Ghaffar, A. (2016). Introduction to the special issue on "Effective Leadership for Health Systems." *Health Systems and Reform, 2*(3), 171–175.

Review on Antimicrobial Resistance. (2016). Tackling drug-resistant infections globally: Final report and recommendations. Retrieved from https://amr-review.org/sites/default/files/160518_Final%20paper_with%20cover.pdf

Rufat, S., Tate, E., Burton, C. G., & Maroof, A. S. (2015). Social vulnerability to floods: Review of case studies and implications for measurement. *International Journal of Disaster Risk Reduction, 14*(4), 470–486.

Rutherford, D. (1992). *Routledge dictionary of economics*. London, UK: Routledge.

Sales, M., Kieny, M.-P., Krech, R., & Etienne, C. (2013). Human resources for universal health coverage: From evidence to policy and action. *Bulletin of the World Health Organization, 91*, 798.

Sands, P., Mundaca-Shah, C., & Dzau, V. J. (2016). The neglected dimension of global security: A framework for countering infectious-disease crises. *New England Journal of Medicine, 374*(13), 1281–1287. doi:10.1056/NEJMsr1600236

Schäferhoff, M., Suzuki, E., Angelides, P., & Hoffman, S. (2015). *Rethinking the global health system.* London, UK: Chatham House, Royal Institute of International Affairs.

Secretariat of the Convention on Biological Diversity. (2011). *Nagoya protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilization to the convention on biological diversity.* Quebec, Canada: United Nations Environmental Programme.

Sherry, M., Ghaffar, A., & Bishai, D. (2017). Community platforms for public health interventions. Retrieved from http://dcp-3.org/chapter/2589/community-platforms-public-health-interventions

Stark, K. D., Arroyo Kuribrena, M., Dauphin, G., Vokaty, S., Ward, M. P., Wieland, B., & Lindberg, A. (2015). One Health surveillance: More than a buzz word? *Preventive Veterinary Medicine, 120*(1), 124–130. doi:10.1016/j.prevetmed.2015.01.019

Tadesse, B. T., Ashley, E. A., Ongarello, S., Havumaki, J., Wijegoonewardena, M., Gonzalez, I. J., & Dittrich, S. (2017). Antimicrobial resistance in Africa: A systematic review. *BMC Infectious Diseases, 17*(1), 616. doi:10.1186/s12879-017-2713-1

Taleb, N. N. (2012). *Antifragile: Things that gain from disorder*. London, UK: Penguin Books.

The Royal Society. (2004). *Making the UK safer: Detecting and decontaminating chemical and biological agents.* London, UK: The Royal Society.

Unal, B., & Aghlani, S. (2016). *Use of chemical, biological, radiological and nuclear weapons by non-state actors: Emerging trends and risk factors. LLoyd's Emerging Risk Report—2016*. London, UK: Lloyds & Chatham House, Royal Institute of International Affairs.

UNICEF & Ministry of Health and Sanitation. (2014). *Sierra Leone health facilities survey 2014: Assessing the impact of the EVD outbreak on Sierra Leone's health system.* Freetown, Sierra Leone: Government of Sierra Leone Ministry of Health and Sanitation.

United Nations. (2015). Transforming our world: The 2030 agenda for sustainable development. Retrieved from https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf

United Nations Development Group. (2015). *Socio-economic impact of Ebola virus disease in West African countries: A call for national and regional containment, recovery and prevention.* Addis Ababa, Ethiopia: United Nations Development Group—Western and Central Africa.

United Nations Office for Disaster Risk Reduction (UNISDR). (2015). *Sendai Framework for Disaster Risk Reduction 2015–2030.* Geneva, Switzerland: Author.

United Nations Office for Disaster Risk Reduction (UNISDR). (2017a). The Global Platform for Disaster Risk Reduction. Retrieved from https://www.unisdr.org/we/coordinate/global-platform

United Nations Office for Disaster Risk Reduction (UNISDR). (2017b). Terminology on disaster risk reduction. Retrieved from https://www.unisdr.org/we/inform/terminology

United Nations Office for Disaster Risk Reduction (UNISDR), Royal Thai Government, & World Health Organization (WHO). (2016). *Bangkok Principles for the implementation of the health aspects of the Sendai Framework for Disaster Risk Reduction 2015–2030.* Paper presented at the International Conference on the Implementation of the Health Aspects of Sendai Framework for Disaster Risk Reduction 2015–2030, Bangkok, Thailand.

Walker, B. (2010). A social–ecological perspective of resilience. In A. Almedom (Ed.), *Resilience: Interdisciplinary perspectives on science and humanitarianism* (Vol. 1, pp. i–iii). Medford, MA: The Fletcher School, Tufts University.

Walker, N. F., Brown, C. S., Youkee, D., Baker, P., Williams, N., Kalawa, A., . . . Brooks, T. (2015). Evaluation of a point-of-care blood test for identification of Ebola virus disease at Ebola holding units, Western Area, Sierra Leone, January to February 2015. *Euro Surveillance, 20*(12), 1–6.

Walker, P. G. T., White, M. T., Griffin, J. T., Reynolds, A., Ferguson, N. M., & Ghani, A. C. (2015). Malaria morbidity and mortality in Ebola-affected countries caused by decreased health-care capacity, and the potential effect of mitigation strategies: A modelling analysis. *The Lancet Infectious Diseases, 15*(7), 825–832.

Webster, P. (2017). Drug shops as primary point of care: The case of Nigeria. *The Lancet, 390*(10089), 15–17. doi:10.1016/S0140-6736(17)31722-1

Wellcome Trust. (2011). *Sharing research data to improve public health: Full joint statement by funders of health research.* London, UK: Author.

Wernli, D., Jorgensen, P. S., Harbarth, S., Carroll, S. P., Laxminarayan, R., Levrat, N., . . . Pittet, D. (2017). Antimicrobial resistance: The complex challenge of measurement to inform policy and the public. *PLoS Medicine, 14*(8), e1002378. doi:10.1371/journal.pmed.1002378

Whitty, C. J. M., Mundel, T., Farrar, J., Heymann, D. L., Davies, S. D., & Walport, M. J. (2015). Providing incentives to share data early in health emergencies: The role of journal editors. *The Lancet, 386*(10006), 1797–1798.

Wilson, K., Brownstein, J. S., & Fidler, D. P. (2010). Strengthening the International Health Regulations: Lessons from the H1N1 pandemic. *Health Policy and Planning, 25*(6), 505–509. doi:10.1093/heapol/czq026

Wilson, K., von Tigerstrom, B., & McDougall, C. (2008). Protecting global health security through the International Health Regulations: Requirements and challenges. *Canadian Medical Association Journal, 179*(1), 44–48. doi:10.1503/cmaj.080516

World Bank. (2014). *The economic impact of the 2014 Ebola epidemic: Short and medium term estimates for West Africa.* Washington, DC: Author.

World Bank. (2017a). *From panic and neglect to investing in health security: Financing pandemic preparedness at a national level.* Washington DC: Author.

World Bank. (2017b). Pandemic Emergency Financing Facility. Retrieved from http://www.worldbank.org/en/topic/pandemics/brief/pandemic-emergency-facility-frequently-asked-questions

World Bank. (2017c). Regional Disease Surveillance Systems Enhancement (REDISSE). Retrieved from http://projects.worldbank.org/P154807/?lang=en&tab=overview

World Bank. (2017d). The World Bank in Haiti: Overview. Retrieved from http://www.worldbank.org/en/country/haiti/overview

World Health Assembly. (2011). *Resolution WHA64.10. Strengthening national health emergency and disaster management capacities and resilience of health systems.* Geneva, Switzerland: World Health Organization.

World Health Organization (WHO). (2000). *The world health report 2000. Health systems: Improving performance.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2008a). *Global assessment of national health sector emergency preparedness and response.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2008b). *Task shifting: Rational redistribution of tasks among health workforce teams: Global recommendations and guidelines.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2010). *Monitoring the building blocks of health systems: A handbook of indicators and their measurement strategies.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2012). *Assessment of health-system crisis preparedness: England (2011).* Copenhagen, Denmark: World Health Organization Regional Office for Europe.

World Health Organization (WHO). (2013). *Universal health coverage: Supporting country needs.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2014). *Ebola response roadmap situation report update—November 7th.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2015a, June 29). Contingency Fund for Emergencies. Retrieved from http://www.who.int/emergencies/funding/contingency-fund/en/

World Health Organization (WHO). (2015b). Developing global norms for sharing data and results during public health emergencies. Retrieved from http://www.who.int/medicines/ebola-treatment/blueprint_phe_data-share-results/en/

World Health Organization (WHO). (2015c). *Factors that contributed to the undetected spread of the Ebola virus and impeded rapid containment one year into the Ebola epidemic*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2015d). *Global action plan on antimicrobial resistance.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2015e). Global Antimicrobial Resistance Surveillance System (GLASS). Retrieved from http://www.who.int/glass/en/

World Health Organization (WHO). (2015f). *Implementation of the International Health Regulations (2005). Report of the Review Committee on Second Extensions for Establishing National Public Health Capacities and on IHR Implementation.* A68/22 Add.1. Geneva, Switzerland: Author.

World Health Organization (WHO). (2015g). *WHO framework for a public health emergency operations centre.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2016a). Emerging Diseases Clinical Assessment and Response Network (EDCARN). Retrieved from http://www.who.int/csr/edcarn/en/

World Health Organization (WHO). (2016b). *International Health Regulations (2005)* (3rd ed.). Geneva, Switzerland: Author.

World Health Organization (WHO). (2016c). *Joint external evaluation tool: International Health Regulations (2005).* Geneva, Switzerland: Author. Retrieved from http://apps.who.int/iris/bitstream/10665/204368/1/9789241510172_eng.pdf

World Health Organization (WHO). (2016d). *Open mindsets: Participatory leadership for health.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2017a). Emergency Medical Teams. Retrieved from http://www.who.int/hac/techguidance/preparedness/emergency_medical_teams/en/

World Health Organization (WHO). (2017b). Global Outbreak Alert and Response Network (GOARN). Retrieved from http://www.who.int/ihr/alert_and_response/outbreak-network/en/

World Health Organization (WHO). (2017c). One Health. Retrieved from http://www.who.int/features/qa/one-health/en/

World Health Organization (WHO). (2017d). WHO estimates cost of reaching global health targets by 2030. Retrieved from http://www.who.int/mediacentre/news/releases/2017/cost-health-targets/en/

World Health Organization (WHO) & World Bank. (2015). New report shows that 400 million do not have access to essential health services. Retrieved from http://www.who.int/mediacentre/news/releases/2015/uhc-report/en/

World Health Organization (WHO) Alliance for Country Assessment. (2016). Joint External Evaluation of IHR core capacities of the Islamic Republic of Pakistan. Mission Report: April 27–May 6, 2016. Retrieved from https://www.ghsagenda.org/docs/default-source/jee-reports/pakistan-jee-report.pdf

World Health Organization (WHO) Eastern Mediterranean Regional Office. (2017). *Assessment of essential public health functions in countries of the Eastern Mediterranean Region: Assessment tool.* Geneva, Switzerland: Author.

World Health Organization (WHO) Executive Board. (2015). *IHR and Ebola.* EBSS/3/INF./4, EB136/INF./7. Geneva, Switzerland: Author.

World Health Organization (WHO) Regional Office for the Western Pacific & Regional Office for South-East Asia. (2010). *Asia Pacific strategy for emerging diseases.* Geneva, Switzerland: Author.

World Organization for Animal Health. (2017). The OIE Tool for the Evaluation of Performance of Veterinary Services

(OIE PVS Tool). Retrieved from http://www.oie.int/solidarity/pvs-evaluations/oie-pvs-tool/

Worsnop, C. Z. (2016). Domestic politics and the WHO's International Health Regulations: Explaining the use of trade and travel barriers during disease outbreaks. *Review of International Organizations*. doi:10.1007/s11558-016-9260-1

Yozwiak, N. L., Schaffner, S. F., & Sabeti, P. C. (2015). Make outbreak research open access. *Nature, 518*(7540), 477–479.

Zanotti, L. (2010). Cacophonies of aid, failed state building and NGOs in Haiti: Setting the stage for disaster, envisioning the future. *Third World Quarterly, 31*(5), 755–771. doi:10.1080/01436597.2010.503567

CHAPTER 15

Management and Planning for Global Health

Tolib Mirzoev, Andrew Green, and Marsha Orgill[1]

What Is Management and Planning?

This chapter focuses on an area essential within global health for all parts of public health—management and planning. We focus particularly on application of these concepts at the national level and below, and management primarily within and by the public sector. Although we occasionally refer solely to either management or planning, the key issues discussed in this chapter—such as the wider contextual influences and key themes—apply to both management and planning.

The importance of health management and planning for strengthening national health systems has been increasingly recognized. However, management and planning in the health field still have an unfortunate reputation. They are often regarded as unnecessary and cumbersome activities that at best divert resources from the real frontline activities of providing health care or preventing ill health. At worst, health management and planning are seen as interfering in these activities in an unhelpful and bureaucratic manner. When they are working well, their presence is not always noticed; when they are malfunctioning, they are likely scapegoats for the health system's problems. In this chapter, a more balanced perspective on management and planning is sought, beginning by asking what management really is, and subsequently discussing health planning.

The key to understanding management and planning is the relationship between resources and objectives. Management is a process of making decisions about how resources will be generated, developed, allocated, combined, and used in pursuit of particular organizational objectives. It is difficult to deny the need for such decisions, given that resources—including, for example, financial, staff, medical supplies, physical space, and transport—are limited. Health planning is a "method of trying to attain explicit objectives for the future through the efficient and appropriate use of resources, available now and in the future" (Green, 2007, p. 3).

Planning and management are related, yet different activities. Both management and planning are concerned with ensuring the best possible use of resources to achieve the intended objectives. An important distinction between the two relates to their time frame and the balance between strategic and

[1] Acknowledgment is given to Charles Collins, who co-authored the previous editions of this chapter.

operational perspectives in their foci. Management is primarily concerned with shorter-term, operational decisions regarding resource use, whereas planning focuses on linking resource requirements to achievement of projected objectives in the future. Planning can be both strategic (i.e., longer term) and operational (i.e., shorter term). Health planning has also been described as a bridge between longer-term strategic policies and shorter-term operational management (Collins & Green, 2014). Drawing on this conceptualization, we interpret policy, planning, and management as a continuum. The policy end of this continuum is primarily about defining the strategic vision and longer-term direction; the management end is mostly about the day-to-day decision making; and health planning serves as the connecting process.

How limited resources within the health system are deployed is critical. Decisions are needed at a macro level within a health system regarding how resources are to be allocated among different areas of activity, and at a more operational level among different approaches to the delivery of health services. Management and planning are concerned with improving the allocation of such resources. For example, there may be an imbalance between the resources targeted at curative activities versus public health, between levels of care (primary versus secondary versus tertiary), between disease control programs, between geographic areas, between social or ethnic groups, between spending on different items such as personnel versus medicines, and between allocations of different personnel such as physicians versus nurses.

How such decisions are made is critical, and this process in part relates to the question of *who* manages. Management and planning functions are best carried out as a shared responsibility between staff whose only or primary activity is management or planning (i.e., managers or planners) and those whose primary function is working as health workers. One important role of specialist managers is to provide health workers with the space and resources they require to carry out their roles (see the *Public Health Infrastructure* chapter for more on the public health workforce and leadership).

Although the discussion presented here so far has focused on the limited resources available within health systems and the key role of managers and planners in responding to this major constraint, it is important to recognize other constraints that they face, including political (in the widest sense) limitations. These healthcare personnel may respond in different ways to such constraints—that is, by accepting, challenging, or looking for ways of maneuvering around them. Indeed, the success of management is often a function of the ability of a manager to work creatively with, rather than passively accept, constraints. At an operational level, this key attribute of management is one way in which this role differs from administration, which is more related to routine implementation of existing rules and procedures.

The definition of management presented earlier in this chapter may suggest it comprises a technocratic mechanical process. In reality, effective management is as much an art as a science, and the same description also applies to planning. One particular aspect of management that is often forgotten is its role in dealing with contradictions (discussed later in this chapter). To be effective, managers must recognize the context within which they operate and adapt their approach accordingly. Such contextual factors include the health situation, the degree of political stability, the general attitude toward public-sector reforms, the level of the country's economic growth, and international influences. Solutions that are appropriate in one situation may not work in another. One key message of this chapter, therefore, is to exercise caution when called upon to adopt blueprint approaches to management.

Management and planning occur at different levels of a country's health system, including at the organizational (e.g., within a ministry of health or district), sectoral (health sector), and health system (across the social sectors) levels. Health managers and planners are therefore staff within organizations at each level of a country's health system. Health managers can work within a specific health program or area (e.g., tuberculosis control program manager), work within a specific component of a health system (e.g., human resources manager), or focus on the whole health system (e.g., district medical officer). Similarly, health planners, and the resultant health plans, can focus on a particular program, a health systems component, or a health system as a whole. Given that both management and planning are concerned with the best possible use of available resources, there are inevitable overlaps between the work of planners and managers. For example, planners are expected to work closely with finance managers in projecting the budget for implementing health-related plans.

Management and planning at sectoral and system levels can also be interpreted as a reflection of approaches to, and principles of, wider governance within the health system such as participation, accountability, and use of information to support decisions (Mikkelsen-Lopez, Wyss, & de Savigny, 2011; Siddiqi et al., 2009). Earlier we outlined the relationships among policy, planning, and management. An important distinction between the concepts

of governance and management and planning is the following: Health systems governance is defined by the World Health Organization (WHO) as "ensuring strategic policy frameworks exist and are combined with effective oversight, coalition-building, the provision of appropriate regulations and incentives, attention to system-design, and accountability" (WHO, 2007a). This is a much broader concept that frames the earlier-mentioned continuum of policy–planning–management. In the last decade, there has been a substantial proliferation of studies exploring governance or its individual elements, such as regulation, and proposing frameworks and principles of good governance and detailed variables for measuring governance within health systems (Lehmann & Gilson, 2015; Mikkelsen-Lopez et al., 2011; Ramesh, Wu, & He, 2014; Siddiqi et al., 2009; Vian & Bicknell, 2014; Vian, Brinkerhoff, Feeley, Salomon, & Vien, 2012; WHO, 2010a). A recent systematic review has identified as many as 16 frameworks for understanding governance that stem from various disciplines, including institutional economics, political science, public management, and development studies (Pyone, Smith, & van den Broek, 2017). Siddiqi et al.'s (2009) 10 principles of good governance (strategic vision, participation and consensus orientation, rule of law, transparency, responsiveness, equity and inclusiveness, effectiveness and efficiency, accountability, intelligence and information, and ethics) underpin one of the most comprehensive frameworks for understanding and assessing health systems governance. A recent framework for good governance, proposed by Collins and Green (2014), emphasized eight similar criteria: a proactive state, effective accountability, decentralization of decision making, intersectoral action, inclusive involvement, ethical conduct, effective regulation, and transparency and democratizing information.

The political character of management and planning also needs to be emphasized. Although management and planning require certain technical skills, they inevitably deal with change, which can be threatening to affected groups and, therefore, stimulates opposition. Thus, for management and planning to succeed, their practitioners must have capabilities in political analysis as well as technical skills. Various techniques have been developed to map attitudes to particular interventions, of which the best known is stakeholder analysis (Brugha & Varvasovszky, 2000). How different groups are involved in management is likely to affect the quality of the process, the speed of decision making, and the ownership of the resultant decisions and action. For example, consultation may occur at different points in the management and planning process—with different aims and implications. It can be seen as the following types of engagement:

- Seeking views as to priorities and strategies at the beginning of decision making
- Seeking views on alternative options once these choices have been formulated
- Seeking views on a formulated plan of action

If consultation takes place later in the process, the more it may be viewed as a formal and even tokenistic process. The earlier consulation takes place, the more likely it is to influence the thinking behind the development of management actions and plans. Consultation can take place in a variety of ways, including directly with the stakeholders (e.g., through surveys or focus groups) or indirectly through representative organizations. It can take place through special one-off mechanisms or through ongoing management processes. Each approach has both advantages and disadvantages in terms of the resources and time required and the robustness of the information gathered.

A fundamental point regarding both management and planning is this: These processes are not and cannot be value-free. One determinant affecting the particular choice of management and planning approach is the values underpinning the health system (Collins & Green, 2014), which could include some of the following:

- Equity, including its different interpretations
- Efficiency
- Choice
- Gender sensitivity
- Transparency and accountability of decision processes
- Market values, including the pursuit of profits
- Participative and accountable decision making
- Solidarity
- Long-term perspective

Some values may be inconsistent with and perhaps even contradict others, most notably market values and equity. The choice, pursuit, and realization of these values through management processes is political in that it affects the interests and views of different groups in society. Furthermore, managers themselves are not value-neutral, but rather bring to the job their own political perspectives and positions.

A health sector committed to the pursuit of accountability and transparency is bound to approach its decision-making processes and consultation differently from one that assigns less importance to this value. Closely related to this are differences between approaches to management and planning within the

public and private sectors.[2] Although the main difference between these two sectors is often suggested to be the freedom of maneuvering enjoyed by a manager, a more salient difference stems from the values and principles that underpin the public and private sectors, and in particular their organizational motives. The private (for-profit) sector is aimed at profit generation for its owners. This stance contrasts with the social goals of the public sector, whose formal objectives may be stated as health promotion and protection, equity and accountability, responsiveness to people's medical and nonmedical expectations, and, increasingly, efficiency and cost containment. This difference in motivation suggests that different styles of management will be used in the public and private sectors, with the former usually emphasizing longer-term sustainability, often collaborative as opposed to competitive strategies, and political debate and negotiation regarding social values. This chapter focuses primarily on public-sector management.

FIGURE 15-1 depicts a framework for management and planning, and the approach taken in this chapter. The chapter follows this schema. It first draws out the important aspects of context. This is followed by a discussion of the structures within which health care may be organized. We then look at health planning as a key activity in a health system that determines the pattern of activities and services to be developed to meet an organization's objectives. Next, the chapter examines issues in the management of resources, focusing on finance, staff, transport, and information. It concludes by discussing some cross-cutting themes in management and planning. The reader is cautioned that, because the areas of health management and health planning are broad, the analysis in this chapter cannot be exhaustive and is, in parts, necessarily selective. It is hoped that this introductory chapter will lead the reader on to further reading on specific topics in health management and planning.

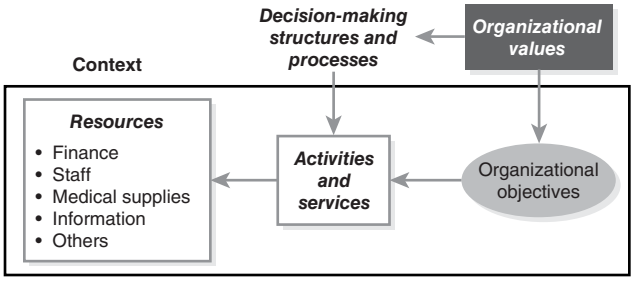


FIGURE 15-1 Key elements in management and planning.

The Context of Management and Planning

Management and planning are social processes of relations that differ according to the social context in which they are found. One would not expect the processes of a public health service in Mali to be the same as those of a multinational corporation based in the United States, for example.

This potential for variation poses dilemmas for the manager and the planner, who need to recognize that the management and planning process does not occur in a vacuum. The way in which managers (and planners) relate resources to objectives needs to be adapted to the environment in which that management takes place. However, the manager faces two difficult issues in ensuring this flexibility. First, he or she must confront universalism—that is, the idea that there is one best way for management to work, irrespective of time, place, or condition. Second, the relationship between the management process and the context is not easy to understand and respond to. On the one hand, the context strongly influences the management process. Management and planning cannot be simply transplanted from one situation to another; rather, the manager must read the current and ever-changing context and determine the most appropriate response given that background. The economic, social, and political context determines the social well-being of communities and the health needs they express, together with the pressures on healthcare organizations. The context affects the feasibility and effectiveness of the options available to managers and planners. On the other hand, the influences do not just flow in a single direction: Managers can (and indeed are expected to) have an impact on the environment.

This section analyzes the complex relationship between the management and planning processes and the context in which they operate. Although much of the discussion refers specifically to the context of management, the same context is also applicable to health planning.

Different definitions of context exist. Dobrow et al. (2004) defined context as including "all factors within an environment where a decision is made" (p. 209). A more detailed approach, introducing a three-tier distinction between macro (systems), meso (organizational), and micro (individual and interpersonal) levels of context, has been used in health

[2] Particular aspects of the analysis in this chapter referring to the public sector and its characteristics and contradictions draw on and develop points raised by Stewart and Ranson (1994).

policy analyses (see, for example, Hudson & Lowe, 2009; Mirzoev et al., 2017; Ricketts, 2010). For organizational purposes, the context in this chapter is broken down into different thematic areas, economic, social, political, and international. These are analytical constructs, however, and do not represent real-life boundaries. In reality, such boundaries are overlapping, permeable, and flexible.

Economic Context

The economic circumstances of a society will influence employment and income and the extent to which individuals and groups can meet their basic needs. These macroeconomic circumstances provide the fundamental backdrop to social well-being—or the lack of it—to which management needs to respond.

The management process is also influenced by the amount and type of resources available to managers. The flow of resources to, and within, the health sector reflects the country's economic circumstances (e.g., economic growth and amount of government taxation) and resource allocation decisions. The flow of resources has important effects on the freedom with which managers are able to work through various options. Furthermore, availability of resources can affect the degree and nature of involvement of different stakeholders in management processes.

Social Context

The broader processes of social change set the scene for the health context in which managers work. For example, unemployment, aging of the population, and migration all raise social and health issues, placing new demands on the health sector. At the same time, social factors influence the feasibility and effectiveness of options. For example, although managers may seek to develop community participation, the feasibility of this endeavor will depend on factors such as past experiences with community participation, geographic settlement patterns, and the degree of social homogeneity. **EXHIBIT 15-1** provides an example from Brazil that illustrates the contextual impact on health management.

EXHIBIT 15-1 Social Context and Health Management: Local Context and Decentralized Health Management in Brazil

Research analyzed the factors that influence the effectiveness of decentralized health management in three Brazilian municipalities. The study identified a variety of contextual factors, including political patronage, different ethical notions of acceptable practices, and differing commitments of staff to the localities. While showing the effects of context on the management process, the study also highlighted the role of managers in influencing that context by referring to the "space for the formal health system to influence local social organisation and political culture and offer a potential for change" (Atkinson, Medeiros, Oliveira, & de Almeida, 2000, p. 632).

Political Context and Public Sector

Four issues in the political environment in which the public sector operates are also contextually important.

Public Service Orientation, Corruption, and Patronage

The management process should be concerned with health and healthcare objectives for the public. Defining these objectives is difficult enough given the contradictory interests of social and political groups in society. Public-sector managers, however, constantly face a further challenge, related to the contradiction between this public service role and the opportunities for private gain through the public sector that arise in four interrelated ways (Green & Collins, 2003):

- *Corporate gain*: Private companies often profit from the public sector. Obvious examples include pharmaceutical, construction, and information technology companies. Tax concessions or public-sector training for the private sector may subsidize the private medical industry. The private sector may also be able to capture those public agencies designed to regulate it.
- *Corruption*: Illegal and unethical use of public resources for private gain can take many forms, ranging from bribes to theft, and can have an adverse effect on government activities (Gaitonde, Oxman, Okebukola, & Rada, 2016). It is important to note that corruption is not solely the preserve of low- and middle-income countries (LMICs).
- *Patronage*: Public resources may be used to strengthen the political position of a patron or political leader. Complex networks of patronage can emerge and sustain themselves, particularly through the manipulation of employment (Collins, Omar, & Hurst, 2000).
- *Professionalism*: Although its contribution to standards and the quality of care can be positive, professionalism can also lead to factionalism and the manipulation of public resources to favor the interests of particular professional groups to the detriment of the public interest.

Integrity and Cohesion of the Public Sector

Public-sector management takes place within a system that is not always cohesive. Such discontinuities are particularly problematic in some LMICs. The impact of neocolonial domination, economic crises, political conflict, famine, and national disasters can lead to disintegration and fragmentation of state authority.

Structural Change of the Public Sector

The public sector in many countries is undergoing significant change. Health-sector management and planning has been part of that change, as witnessed by the three major waves of international health reform affecting LMICs—the Alma Ata primary health care (PHC) movement, subsequent market-based reforms (discussed in the *Design of Health Systems* chapter), and more recent efforts to apply more integrated approaches to health systems development in achieving universal health coverage, signaling the revival of the PHC agenda. **TABLE 15-1** identifies the potential implications of these broader changes for the role and operation of health-sector management in LMICs.

The management changes outlined in Table 15-1 have been mediated by other factors and have had different impacts both between and within country health systems. First, alternative health-sector approaches, drawing their inspiration from the PHC approach, have continued to emphasize, for example, the importance of citizenship, equity, and social justice, and have focused on community participation in health systems decision making. This multiplicity of aims can generate contradictions for management

TABLE 15-1 Health-Sector Change and Management

Management Change	Examples
New management responsibilities	The introduction of competitive relationships, contracting, voucher systems for service users, and market research to assess customer responses generates new responsibilities for managers.
New management skills	The introduction of the above responsibilities can require new skills in areas such as contracting, customer relations, and quality assurance techniques.
New management boundaries	The introduction of public–private joint ventures, contracting out to the private sector, and the use of competitive markets have blurred the distinction between the public and private sectors. This evolution suggests new boundaries for public-sector managers, with implications for management values and relations.
Changing management actors	Reforms can lead to greater diversity within the health system, as public-sector facilities gain semiautonomous status and private-sector facilities (both nonprofit and for profit) become involved in health care. Managers need to take note of this diversity in developing a range of relations, from collaboration to competition.
Changing management objectives	Efficiency and health facility financial survival and growth can eclipse objectives and values based more on social justice and equity. For some, the generation of profit or financial surplus becomes an accepted objective of management action.
Changing management structures, systems, and processes	Management structures, processes, and systems associated with the private sector have been introduced into the public sector—for example, performance-based incentives in pay structures and competition between health facilities.
New management options	Reforms in many countries have opened up the option of contracting out service provision to the private sector, as opposed to relying on internal service provision.
Dealing with new challenges and contradictions	Decentralization has been associated in some countries with the requirement that health facilities be financially self-sufficient. Managers face the challenge of simultaneously ensuring the continuity of healthcare services provision, meeting equity objectives, and generating resources for institutional survival.

(Flynn, 1997). For example, market-based strategies of service delivery emphasizing individual customer choice may clash with more citizen-based strategies based on community participation and collaborative relations within the public sector.

Second, the capacity for making the type of management changes outlined in Table 15-1 is constrained in many LMICs by the prevalence of informal arrangements in the economy and public sector, and the scarcity of general administrative and managerial skills and systems (Schick, 1998). Corruption and patronage frequently maintain a stranglehold on health management and planning, thereby limiting systems' capacity for change. The result is often a changing blend of different approaches, leading to a muddled, confused, and contradictory character for public-sector management. At times, a thin veneer of reform fails to disguise a cumbersome bureaucracy bent on corruption and patronage. At other times, purposive action by dedicated health staff overcomes strong constraints to provide health care to communities.

Increasing Role of the Private Sector

The private sector consists of both for-profit and not-for-profit institutions, each with a different set of objectives. The primary objective of the for-profit private sector is to benefit financially, whereas the objectives of the not-for-profit private health sector are closer to the social objectives of the public sector. The private sector can include in-country institutions (e.g., private hospitals or private insurance companies), international agencies (e.g., international pharmaceutical companies) and global institutions (e.g., Global Fund to Fight AIDS, Tuberculosis and Malaria). The distinction between the private and public sectors is often blurred—for example, many public-sector health staff may have private practices in addition to fulfilling their public-sector jobs.

The role of the private sector, which is discussed in more detail in the *Design of Health Systems* chapter, differs across countries. Nevertheless, it typically includes engagement in the areas of policy making, financing, provision of services (often contracted out from the public sector), and regulation.

Outsourcing of public services through contracts with the private sector (e.g., cleaning or transport services) is seen as a way of benefiting from economies of scale and the technical expertise of specialist contractors in areas generally removed from the central responsibilities of healthcare organizations. Some also argue that the private sector's perceived efficiency and the control and accountability conferred by contracts are other advantages associated with this model. Concerns raised by outsourcing of these responsibilities include the transaction costs of contracting out; questions about whether lower costs are merely the result of lower salary costs; the lack of sustainability and management learning in contracts; potential problems with poor quality, collusion, and corruption; and the lack of contracting skills in purchasing organizations (Mills, 1998).

The public and private sectors can engage in different types of interrelationships, referred to by many as the public–private mix (PPM) or public–private partnerships (PPPs). The underlying assumptions for the PPM model are its perceived greater efficiency in service delivery; a recognition that the public sector alone may be unable to ensure comprehensive coverage of, and access to, health services; and the complementary nature of the public and private sectors in achieving this aim.

The role of the private health sector (both the for-profit and nonprofit segments) was widely promoted in the mid-1990s through, for example, incentives to the private sector (e.g., tax concessions on capital invested or on private health insurance), public-sector contracting with the private sector (e.g., in Brazil), outright privatization, and outsourcing in the private sector (e.g., support services in a public-sector hospital). There are examples of successful PPPs, such as in tuberculosis control (Karki, Mirzoev, Green, Newell, & Baral, 2007). Nevertheless, the uncontrolled growth of the private sector in many countries has raised concerns about the quality and coverage of many health services delivered through private-sector companies (Mills, Brugha, Hanson, & McPake, 2002), leading to further health inequalities and calls for better regulation (Lê, Ha, Mirzoev, & Nguyen, 2015; Sheikh, Saligram, & Hort, 2015), as discussed in the *Design of Health Systems* chapter. Examples of challenges for public-sector managers include understanding the new range of providers and negotiating new relationships through, for example, contracts, leasing, and concessions; the need for public-sector managers to devise strategies to retain staff in the public sector; and pressures to incorporate the style and techniques of private-sector management into the workings of public agencies.

International Context

The international economic and political context has an important effect on the national and local economic, social, and political processes of a country through investment, interest rates on debt repayments, trade, and technical advice and support, among other mechanisms. For example, significant funding is provided

by the Bill and Melinda Gates Foundation and other philanthropic institutions for health systems development in LMICs. These initiatives, which are known as global public–private partnerships (GPPPs), inevitably affect management and planning practices at country and local levels. For example, while providing increased funding, the GPPPs may be

> depriving specific stakeholders a voice in decision-making; . . . [contributing to] misguided assumptions of the efficiency of the public and private sectors; . . . wasting resources through inadequate use of recipient country systems and poor harmonisation; and inappropriate incentives for staff engaging in partnerships. (Buse & Harmer, 2007, p. 259)

The growth of the international policy presence of the World Bank during the 1980s and 1990s had an important impact on health management reform in LMICs. The powerful financial presence of such international donors can leave health ministries in a dependent relationship. Moreover, the development of budget support, Poverty Reduction Strategy Papers (PRSPs), and health sector-wide approaches (SWAps) throughout the 1990s influenced changes in the relationship between national governments and international agencies. More recently, the global financial crisis in 2008 affected the availability of resources for health sectors at both the national and international levels (Bednarz, 2010). The general concern that underfunded and weak health systems had hindered the achievement of health-related Millennium Development Goals (MDGs) also spurred an exploration of alternative health financing mechanisms, as shown by the establishment of a High-Level Taskforce on Innovative International Financing for Health Systems (Fryatt, Mills, & Nordstrom, 2010).

More recent changes in international context include large-scale population displacements, increasingly frequent and severe conflicts and emergencies, disease outbreaks (e.g., Ebola and Zika viruses), and political commitments and targets related to the Sustainable Development Goals (SDGs) that replaced the MDGs. Many of these changes highlight the growing number of fragile states on a worldwide basis, and call for integrated, cross-sectoral management approaches to help ensure health system resilience and responsiveness as well as more sustainable health systems development (O'Hare, 2015; Roome, Raven, & Martineau, 2014; Schmidt, Gostin, & Emanuel, 2015).

The general point raised by the discussion in this section is the importance of understanding the interrelationship between context and health management and planning. Managers need to interpret the wide range of changing and complex contextual factors that influence the way in which management is actually conducted in the health system. At the same time, managers can be proactive; they are not powerless (Atkinson et al., 2000; Grindle, 1997). There is a margin of maneuverability in which they can operate, which varies in time but allows purposive action to be developed. One example of this flexibility can be seen in the different forms of sector-wide approaches that have emerged in response to different country contexts (WHO, 2006a).

Organizing

An important management function is developing the organizational structure—that is, the way that work is assigned, both vertically and horizontally, together with the formal framework of links with other organizations and groups. Managers usually inherit an established structure. Recognizing this fact of life, this section takes the form of a review of organizational structure, indicating key issues and options available to managers to address those issues.

In reviewing this organizing function, several points need to be emphasized. First, health managers need to consider the full range of structural changes open to the organization. For this purpose, this section presents a framework setting out the dimensions of a review of the organizational structure. For purposes of illustration, it refers to ministries of health, although the framework could be adapted to decentralized health authorities or health facilities. The role of managers in structuring an organization will depend on their authority within the organization and the significance of the structural issue to the organization.

Second, in deciding on organizational structure, managers need to take into account contextual factors, such as the staff capacity, the health policy, and the overall government structure.

Third, there is a need to secure a balance between two key factors. On the one hand, managers have to keep the organizational structure under review and implement necessary changes. Structural changes are not the only component of change, however; they are just one factor to be considered and balanced against changes in resources, systems, values, and skills. On the other hand, managers need to recognize that structural changes can be both expensive in terms of resource use and disruptive to staff motivation and service delivery. Change should also not be used to mask more difficult and controversial issues of resource availability or health policy content.

Three overlapping dimensions of the organizational structure are reviewed here:

- Center–periphery relations
- Relationships with other organizations and groups
- Internal structure

These dimensions relate system-level issues to more micro-level issues. Reference may be made to the *Design of Health Systems* chapter to understand some these issues in greater depth.

Center–Periphery Relations

A ministry of health (MOH) is organized based on relations that link the center to the periphery. This raises important dimensions of geographic decentralization, purchaser–provider relations, and delegated semi-autonomy.

Decentralization

Decentralization is "a transfer of authority to make decisions, to carry out management functions and use resources. Focusing on the public sector, it means the passing of these from central government authorities to such bodies in the periphery as local government, field administration, subordinate units of government, specialised authorities and semiautonomous public corporations" (Collins, 1994). An assumption, or perhaps a value-driven judgment, here is that greater local decision space within the decentralized context can help ensure greater accountability at local levels. Conversely, an argument can be made that in a decentralized health system where each regional government is sufficiently autonomous, there is a clear danger of losing an inter-regional equity perspective.

Purchaser–Provider Relations

As explained in the *Design of Health Systems* chapter, health-sector reforms have encouraged a shift from organizationally integrated and hierarchically structured systems to a separation between the purchaser organization, which has the financial resources and interprets the needs, and the provider, which is responsible for services provision. Purchasing involves identifying interventions or services to be purchased (based on need, priority, and cost-effectiveness), choosing service providers (based on quality, equity, and efficiency), and determining how services will be purchased, including contractual arrangements and provider payment mechanisms (RESYST, 2014). "Passive purchasing implies following a predetermined budget or simply paying bills when presented. Strategic purchasing involves a continuous search for the best ways to maximize health system performance by deciding which interventions should be purchased, how, and from whom" (WHO, 2000).

Delegated Semi-Autonomy

Delegated semi-autonomy, as a form of decentralization, involves the transfer of semi-autonomous authority to manage an organization. In health care, a common form involves hospitals organizationally attached to a health ministry. These facilities may be managed by a board that is only partly appointed by the ministry; staff may be hospital (not ministry) employees, while the hospital has powers of staff appointment, revenue generation, determination of salaries, and purchasing authority. This delegation can take the form of a purchaser–provider separation and be executed through a contract that formalizes the hospital's responsibilities and the ministry budget allocations to the hospital.

Another form in which purchaser–provider separation is combined with delegated semi-autonomy occurs when the whole operational side of health care is taken out of the MOH and located in a separate health services structure. This arrangement is found in Ghana, for example. Within the Ghanaian health sector, the MOH wields power over public policy decision-making processes, while the Ghana Health Service (GHS) is a semi-autonomous agency responsible for service delivery at regional, subdistrict, and community levels. Other service delivery agencies also include teaching hospitals and national ambulance services (Couttolenc, 2012; Koduah, 2016). Metropolitan, municipal, and district assemblies do not run or manage facilities, but rather provide certain health-related services, mostly involving support to MOH and GHS activities (Couttolenc, 2012).

Managers should have a responsibility for developing the center–periphery relations, or at least be consulted as important stakeholders in the process of organizational change.

Relationships with Other Organizations and Groups

The public sector—in our example, the MOH—needs to have clear and defined relations with a variety of external organizations and groups. Private-sector growth requires organizational links with government such as joint ventures, contracting out, and policy consultation. Community participation raises issues of how the community will be organizationally integrated into the health system. Many LMICs depend on international organizations, both bilateral and

multilateral, for significant financing of the health sector and develop important relations around health policy formulation and implementation. This organizational relation can take the form of aid agencies developing links with parts of the healthcare system when they take responsibility for financing particular disease-control programs or healthcare services for particular geographic zones. Notably, the development of health SWAps in the 1990s led to organizational linkages between international donors and MOHs to develop health policy frameworks.

The MOH also relates to other government agencies—an interaction typically required as part of the cross-governmental process of management via ministries such as finance, social affairs, and labor. An important aspect of this interagency cooperation is the development of an intersectoral approach to health development. Structuring relationships with special interests, such as the private, for-profit healthcare sector, forms an important part of consultation. At the same time, the MOH is required to exercise regulatory powers over the same special interests.

Internal Structure[3]

Related to the two dimensions of center–periphery and external relations is the organization's own internal relations and assignment of roles. These aspects of internal structure encompass four interrelated factors:

- Hierarchy and span of control
- Relations of authority
- Horizontal divisions
- Internal linking

Hierarchy and Span of Control

Organizations usually adopt some form of hierarchical shape consisting of different levels of management. These recognized levels of authority have a specific depth of authority and span of vision over the organization; the higher up the hierarchy, the greater the authority and the broader the vision. An organization can adopt different hierarchical shapes, although it is often thought that a flatter organizational shape is more appropriate for effective communication and inclusive decision making.

Health ministries often suffer from overextended and overly tall hierarchies. This structure can result from bloated bureaucracies and the tendency to confuse public-service grading systems with management levels (the number of management levels increases to accommodate new grades). Hierarchies can be made flatter by allowing several grades to occupy the same management levels, or by reducing the number of management levels. However desirable it might seem to widen the span of control for entities within the hierarchy, this endeavor can be difficult. Attention has to be paid, for example, to the capacity of both managers and subordinates to deal with a widened span of control, the changed nature of the work, and the wider organizational and systems context in which the managers and planners operate.

Relations of Authority

Relations of authority are important in binding the organization together. Healthcare organizations, depending on the type of decentralization, typically exhibit different forms of overlapping authority, such as strategic, main line managerial, technical, professional, and supervisory. Strategic authority is fundamental to a MOH and involves decision making with implications for the direction in which the whole or a significant part of the organization is working.

One notable feature of healthcare organizations is the existence of dual authority relations in which staff have more than one relation of authority In **FIGURE 15-2**, for example, the district tuberculosis (TB) officer is under the managerial authority of the district health officer but under the technical authority of the regional TB officer.

Dual authority relations can be used to widen spans of control. **FIGURE 15-3** shows how a health manager can increase his or her span of control by appointing a supervisor or support staff to help in the management of staff. The supervisor falls under the main line management control of the health manager, as do the subordinates. Each subordinate, however, is

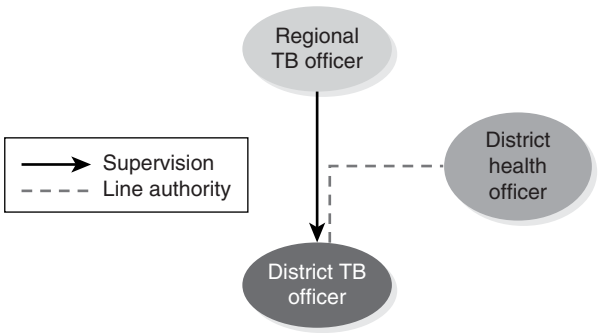


FIGURE 15-2 Dual authority relations.

Modified with permission from Rowbottom, R., & Billis, D. (1987). *Organisational design: The work level approach*. Aldershot, UK: Gower.

[3] This section draws on the work of Collins (1994).

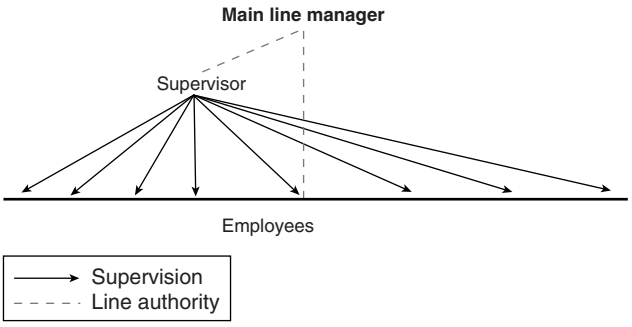


FIGURE 15-3 Reducing the burden of a wide span of control on a main line manager through the creation of supervisory staff.

Modified with permission from Rowbottom, R., & Billis, D. (1987). *Organisational design: The work level approach.* Aldershot, UK: Gower.

under the joint authority of both the supervisor and the main line manager. Because the pressure of direct supervision has been taken off the main line manager, the span of control can be widened.

Although dual authority relations may appear strange and confusing, they are commonplace in health systems. Increasing health system decentralization through devolution to local government presents challenges for authority relations, however. Technical and hierarchical relations with the center need to be balanced with the strengthened horizontal relations within the local government authority.

Managers need to design authority relations in such a way as to avoid potential confusion. There are two complementary ways of clarifying dual authority situations. First, the different forms of authority in operation should be clarified and explained to the staff. For example, the relationships in Figure 15-2 among the regional TB officer, the district TB officer, and the district health officer might be clarified by identifying one superior, the district health officer, as possessing main line managerial authority and the other as holding another form of authority, such as technical supervisory, monitoring, coordinating, or prescribing authority. Second, the dual authority structure could be clarified by specifying the areas of management for the exercise of authority. Key areas of authority to be allocated among the authority holders can include combinations of issues related to staff appointments, induction and training, performance review, supervision and development, development and implementation of policies and programs, introduction of standards and methods, and control over budgets. Both methods of dual authority relations require a clear specification of authority and responsibilities and a culture of understanding between the persons involved.

Horizontal Grouping

At each managerial level, work has to be assigned to staff within different groupings, such as divisions, departments, and units. This division of work may be based on different combinations of geography (e.g., in a decentralized system), functions (e.g., in planning or national disease control programs), or staff cadre (e.g., in a nursing department). The MOH could be based on a simple logic of planning, doing, and supporting, with directorates of planning (including aid coordination and information systems), health services (including primary healthcare services, national health programs, and hospital services), and support services (including human resources and financial administration) (Collins, 1994). This sort of arrangement has some drawbacks, however: Notably, it can lead to a lack of horizontal links between the directorates and to imbalances in power, such that one directorate becomes the most powerful.

Internal Linking

The previous discussion focused on the application of internal structures to divide work, which runs the risk of fragmenting the organization. The health organization needs to be brought together and act as an integrated whole. One way to ensure such cohesiveness is through the operation of strategic authority. Strategic policy authority and resource allocation flow downward, maintaining vertical integration. At the same time, horizontal collaborative links are developed between the various divisions, departments, and units in the organization. Several structural arrangements may be employed to accomplish this goal:

- Job descriptions to specify lines of authority in the particular job in addition to job liaison responsibilities and membership of teams
- Organizational devices, such as interdepartmental committees and task groups, in addition to specific linking responsibilities assigned to staff
- Matrix systems (e.g., staff may belong to the basic departments in the organization in addition to specific task units and projects)

Planning[4]

Like the management function, planning does not have a good reputation in many parts of the health sector. Its record is not good, with plans often not being implemented or implemented ineffectively. The

[4] This section draws on the work of Green (2007).

reasons for this failure vary, but frequently include top-down, rigid, and often bureaucratic centralist processes; a failure to integrate planning processes with other decision-making processes such as budgeting; a failure to fully understand and take account of changing context; and a failure to involve key groups, including health services managers, professional groups, and users in planning. These criticisms should not, however, be interpreted as pointing to the inevitable failure of planning, but rather as recognizing the need to develop systems appropriate to the particular healthcare needs and context of a country. This section looks at various background planning issues and provides an overview of the planning cycle.

Why Plan?

Planning is an essential element of health systems governance that is concerned with making decisions today to influence the future. It is a response to the dilemma that faces these organizations, which must deal with an inevitable shortfall in resources compared with increasing health needs and hence the need to make choices between competing uses for the resources—that is, to set priorities.

Four other critical issues underpin the importance of planning. First is the changing nature of health and health care. Decisions on resource usage need to take account of likely future changes to health needs, resources, and potential health service strategies and technologies. Health need changes include both new diseases—such as human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS) or, more recently, Ebola and Zika viruses—and changes in the relative prevalence of particular problems as a result of epidemiologic or demographic transition. Variations on this theme include, for example, the growth of multidrug-resistant strains of TB. Resource changes that require consideration include both financial forecasts and the availability of key resources—in particular, health workers. In many countries, the large population displacements and international migration of key staff such as doctors and nurses may lead to critical future shortages unless appropriate policies are developed now. The last area of forecasting relates to future technology developments and their impact on the health sector.

Second, as noted earlier in this chapter, decisions about priorities take place within economic, social, and political contexts that will vary between countries and within countries over time. Health planning, therefore, can be seen as means of identifying and responding to the context-specificity of different health systems. Closely linked to these priority decisions are questions as to both what the underpinning values of the health system are and who determines these priorities.

The third issue is the recognition that the current allocation of resources within the health sector is not optimal, as discussed earlier. As a consequence, planning is concerned not just with dealing with changes in the future, but also with addressing current problems in a way that will have effects in the future (e.g., effects of distribution of human resources on availability and quality of services and, ultimately, on health status). Shifts in resources could lead to a more effective and efficient use of resources.

Last, health planning is a way of avoiding decisions driven by status quo, taken by a small elite group(s) of decision-makers such as policy makers, external agencies, or health workers without consulting other relevant stakeholders.

What Is Planning?

Health planning has been defined by Green (2007, p. 3) as follows:

> A systematic method of trying to attain explicit objectives for the future through the efficient and appropriate use of resources, available now and in the future.

The important components of this and other similar definitions are as follows:

- Where one is going (objectives)
- With what (resources)
- How (efficient and appropriate implementation)
- When (future)
- With what degree of formalization (explicitness, systematic, and method) about the process

There are, however, different approaches to, and types of, planning. In particular, we can contrast strategic (or allocative) and operational planning. *Strategic planning*, which closely resembles policy making, aims to provide an open and formalized process for making these difficult decisions as to which health needs will be met by the limited resources and how. Thus it attempts to provide a broad direction of travel for the health sector. In contrast, *operational planning* (also known as activity planning) focuses on the detail of implementation by setting out time frames for activities in the short term. In practice, the two types of planning should be linked, and often there will be elements of both within any particular plan document. Highlighting the conceptual difference between the two is helpful, however.

Health planning also involves a chain of interrelated processes at different levels of a country's health

TABLE 15-2 Misperceptions About Planning

Misperception: Planning Is . . .	But . . .
About the production of plans	Planning is concerned with change, not documents; these may be necessary but are not a sufficient part of the planning process
About capital budgets	Planning should also focus on recurrent budgets
Only concerned with projects	Projects are simply one way of achieving change
A highly technical and specialist activity	Much planning is common sense
Carried out by specialist planners	Planning needs to be shared by a wide group of actors
An objective and neutral activity	Planning involves value judgments

Reproduced with permission from Green, A. (2007). *An introduction to health planning in developing health systems* (3rd ed.). Oxford, UK: Oxford University Press.

system. The following levels of planning can be distinguished: country (e.g., PRSPs), sector (e.g., SWAps, policy, and strategy), and program and project (e.g., maternal and child health [MCH]) (Mirzoev & Green, 2017).

One criticism of planning is that it is seen as being unfeasible during periods of uncertainty or instability. However, the reverse can also be argued—that planning is itself a means of dealing with uncertainty, while retaining a strategic direction. However, it needs to be sufficiently flexible to achieve this aim.

Several common misperceptions about planning need to be addressed. **TABLE 15-2** sets out these sources of misunderstanding, along with the counter-views.

Planning for Health and the Health Sector

The arguments for a government lead in planning hinge on whether health care is viewed as a special good for which the normal market mechanisms are not appropriate (e.g., because of equity implications or because of their public-good nature). The following discussion takes the widely held position that (1) at a minimum the state has a responsibility to set and regulate policies, and (2) the state will continue to provide certain key healthcare services for the foreseeable future.

Historically, planning in the public sector has tended to focus on the state's own healthcare services. More recently, it has become more widely recognized that planning by government needs to recognize the actual and potential inputs (both positive and negative) of other healthcare agencies, such as those in the private for-profit sector and nongovernmental organizations (NGOs). This expanded reach calls for the development of new policy and planning tools to implement such strategies, and will be needed for developing strategic long-term plans and policies to achieve universal health coverage (WHO, 2017).

In addition, although we tend to label plans in the health sector as "health plans," in reality they tend to focus on health care, with little recognition of the positive and negative effects of other (non-healthcare) agencies on health. Genuine health plans need to broaden their scope and incorporate appropriate actions related to other sectors.

The changes in the structure and roles of government that have been taking place in many health sectors in recent years often require corresponding changes in the government planning approach. In particular, the increasing number of healthcare providers that are not directly managed or controlled by one government agency and the decentralization of authority mean that governments need to develop new ways of achieving change in providers other than those whom they directly manage. New forms of incentives and regulatory powers are needed, in contrast to the traditional managerial command-and-control approaches. This trend also implies that a greater onus is placed on the lead government agency to provide overall policy frameworks for the health sector that specify appropriate roles for other agencies.

Approaches to Planning and Their Relationship to the Context in Which Planning Occurs

Various approaches to planning are possible. Following are descriptions of two continua of approaches,

although it is important to note that they are not mutually exclusive:

- *Problem-solving reactive approaches versus longer-term needs assessment*. Planning can focus on, and try to identify solutions to, existing problems. A variety of techniques have been developed to assist planners and managers in problem identification and solving, including problem tree and fish-bone analysis. One danger with such an approach is that its focus on current problems may detract from longer-term needs assessment and timely responses to emerging needs (Green, 2007; Mirzoev & Green, 2017).
- *Structured logical frameworks versus looser strategic directional approaches*. During the 1990s and early 2000s, a number of organizations, and in particular donor agencies, adopted an approach to planning (generally projects) using logical frameworks (logframes) that set out in a structured manner a hierarchy of objectives and activities, together with means of identifying whether they have been achieved, and the potential future risks or assumptions (Nancholas, 1998). Such an approach can be contrasted with a looser narrative strategic plan.

Plans can also have a different focus. Plans may focus on any of the following:

- An organization such as a hospital
- An administrative geographical level such as a district and its population
- Particular health programs (such as reproductive health or TB) or time-bound projects
- Health systems components (such as human resources or information systems)

The appropriate approach to and focus of planning will depend on the context in which the planning occurs. As an extreme example, the presence of a conflict or an emergency situation will affect planning. For example, in Sierra Leone during the civil war, quarterly plans were set, in recognition of the inability to take a longer-term strategic view during this period.

Planning and Organizational Levels

Planning occurs—or should occur—at all levels in the health system, as illustrated in **FIGURE 15-4**. As decentralization policies are increasingly implemented within health sectors, it is important that they incorporate a clear expression of the relative planning responsibilities of the different levels of the health system. **TABLE 15-3** gives an example of the division of responsibilities in a two-level system.

Types of Planning Time Scales

Planning, as we have seen, focuses on actions related to the future. However, decisions have to be made as to the time scale. For many planning systems, a period of five years has been taken as the standard time frame.

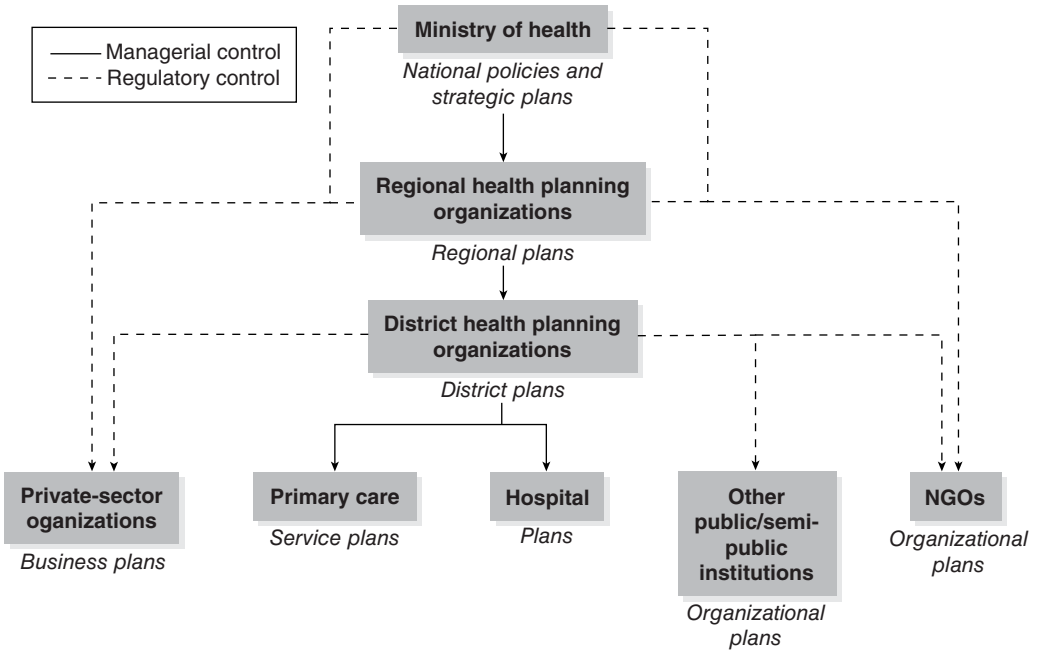


FIGURE 15-4 Examples of planning at different organizational levels within a deconcentrated system.

Modified with permission from Green, A. (2007). *An introduction to health planning in developing health systems* (3rd ed.). Oxford, UK: Oxford University Press.

TABLE 15-3 Example of the Division of Responsibilities Between Center and Periphery

Central Functions	Joint Activities	Local Functions
Broad policy leadership	Monitoring and evaluation	Local needs assessment
Resource generation and allocation		Development of local plans, and local generation of resources in decentralized contexts
Donor coordination		Implementation of local plans
Liaison with central ministries		Liaison with local government authorities
Coordination of local plans		Provision of primary and secondary health services
Planning of central specialist services		Human resources performance appraisal and support
Human resources planning		
Technical planning support		
Legislation		Local operational policies

Modified from Green, A. (2007). *An introduction to health planning in developing health systems* (3rd ed.). Oxford, UK: Oxford University Press.

Other health systems have found this practice too rigid and have adopted a rolling process of planning, as exemplified in **FIGURE 15-5**, where each year the plan period (in this example, three years) is rolled on by a year. To maintain an overall set of strategic direction, this type of plan is often combined with a long-term perspective plan that sets out a broad set of policies or direction of strategic travel. Such an approach underlines the importance of ensuring consistency and full integration between the wider policies, strategic plans and operational plans.

Planning timetables need to take account of both the need for wide consultation at different stages of planning and the fact that planning at each level (e.g., district) must link to planning both at higher and lower levels and other parts of the overall system (e.g., local government plans).

Elements of Planning

FIGURE 15-6 sets out diagrammatically the health planning spiral, which shows the various stages involved in planning. Two general points need to be made. First, conceptually, this process is a spiral rather than a cycle—suggesting movement in a direction rather than a repetitive cycle. Second, planning refers to a process rather than a chronological sequence. Several activities may be occurring at the same time, together with various iterations within, and across, the six stages of the planning process. Each of these stages in planning is briefly introduced here.

Situational Analysis

In the situational analysis stage, planners assess the current situation (in other words, the current context) and project future changes to it. This stage also represents a useful mechanism for getting a planning team working well together early in the planning process and for opening up the process to a wider group of organizations and individuals. Examples of key information needs for situational analysis, covering both the current situation and likely future trends, include population characteristics (e.g., demographic data, morbidity and mortality), area characteristics (e.g., geography, infrastructure, resources), the policy environment, the functioning of the health system (e.g., existing policies, and availability and quality of

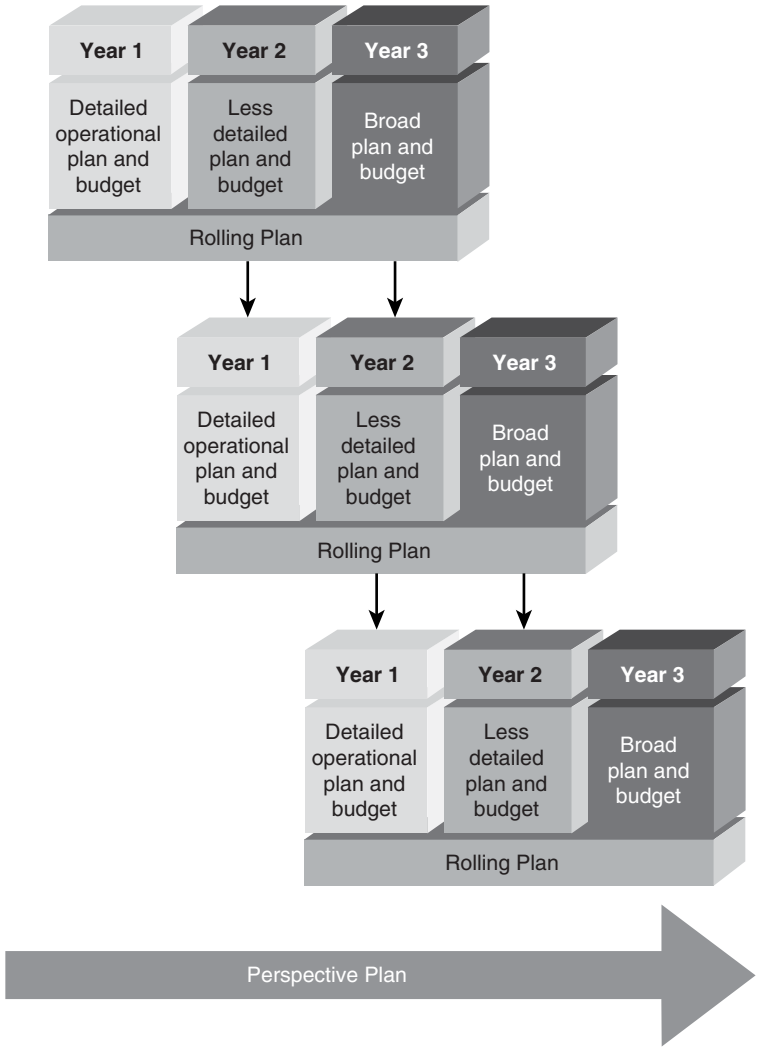


FIGURE 15-5 Three-year rolling plan and long-term perspective plan.

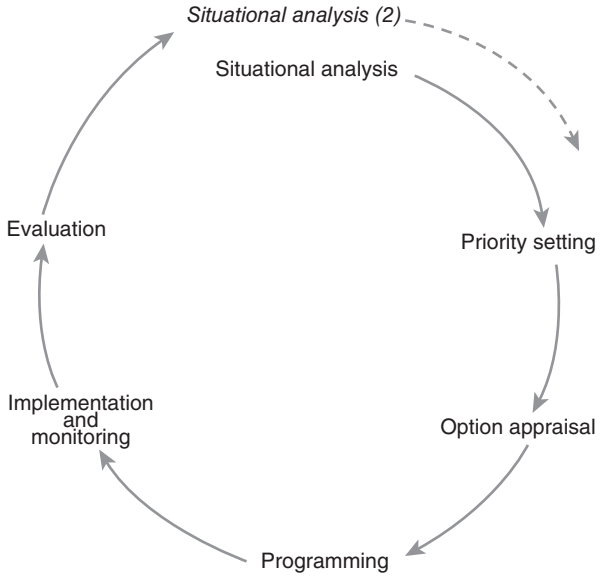


FIGURE 15-6 The planning spiral.

Reprinted with permission from Green, A. (2007). *An introduction to health planning in developing health systems* (3rd ed.). Oxford, UK: Oxford University Press.

services provided), and health needs (as perceived by health staff and communities) (Green, 2007).

A SWOT (strengths, weaknesses, opportunities, and threats) analysis may also be used during situational analysis (**FIGURE 15-7**). This technique directs attention toward internal and external aspects of institutions and is generally better suited to assessing an institution such as a commercial firm, but there may be occasions when it provides a useful format for health-sector organizations.

The end product of the situational analysis stage is usually a focused overview of the current situation with identification of specific areas requiring actions.

Priority Setting

In the second stage, planners set priorities for the organization, in the light of competing needs and limited resources. Setting priorities is perhaps the most critical and most difficult planning stage—but a part of

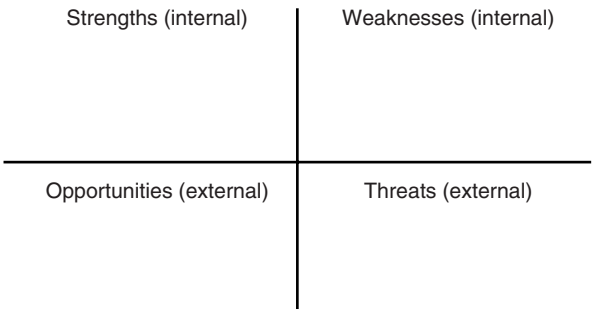


FIGURE 15-7 SWOT analysis framework.

the planning process that cannot be avoided. Priority setting is also a key contributor to health systems performance (Barasa, Molyneux, English, & Cleary, 2015).

For the state providing an overall strategic plan, its priorities need to be sufficiently broad to allow for local variations as a result of differing needs. Underpinning all of the issues regarding how priorities are set is a tension between attempts to make decisions evidence driven and recognition of the essentially political or value-laden nature of such decisions.

What Should Priorities Be Based On? At one level, the most obvious answer to the question of the basis for prioritization is that priorities should be set on the basis of greatest health need. However, this response in itself raises various further questions, such as how health need is perceived. In particular, is a broad or a narrow view of health taken, how is health need measured, and should priorities focus on health needs or healthcare needs? It also (and related to the preceding discussion) is affected by whose perception of health is taken.

Clear criteria are needed so that the process of priority setting can be as open as possible. These criteria should be derived from overall policy and could include the following:

- The maximum health gain given the available resources (efficiency)
- The effects on equity
- Public demands for change

These criteria may sometimes work against one another. For example, there may be a tradeoff between equity and efficiency, unless efficiency is seen as a means to achieve objectives that include the distributional aspects of the health gain.

Who Should Set Priorities? A critical issue within the planning process relates to who has the right or responsibility to set priorities at which level in the national health system. For example, good arguments may be made for any of the following groups to be involved in priority setting: health workers, administrators and managers, users of services, and, more widely, communities or politicians. The priorities set will depend significantly on who makes them. Therefore, adequate engagement, empowerment, and satisfaction of key stakeholders in relation to the process of priority setting are important (Barasa et al., 2015).

Stakeholders' power is an important consideration in understanding the broader political environment of priority setting. For example, a study in Ghana found that interventions that targeted breast cancer, spearheaded by the country's First Lady, received greater attention despite the epidemiologic and economic evidence suggesting that cervical cancer should be prioritized (Reichenbach, 2002). Stakeholder analysis (Brugha & Varvasovszky, 2000) can be a useful technique in assessing the strengths of different groups in society.

Establishing Priorities Within a Planning Framework. It is important that the planners make explicit how priorities are to be set. This process needs to allow a broad view of health, rather than health care alone; find an appropriate balance between decision making at the national and local levels; be transparent; and end up with feasible objectives. A common flaw in the priority-setting process is that everything is viewed as a high priority, meaning effectively that no real priorities have been set. A good test of a robust priority-setting system is whether it clarifies those areas that are not viewed as high priorities.

Various processes and techniques can be used to set priorities, including economic appraisal (e.g., using cost per disability-adjusted life year [DALY] as the basis for measurement), multivariable decision matrices, and Delphi techniques. The resource allocation processes from the center to lower health services levels are also an important vehicle for ensuring that broad priorities (and particularly those of equity) are reflected in the budgets allocated for service delivery. However, there is a danger in assuming that priority-setting is value-free, when it actually incorporates a number of implicit values. **FIGURE 15-8** provides examples of value-driven judgments exercised in conducting economic appraisal of health programs.

Multivariable decision matrices offer another approach in which any number of criteria can be incorporated and the information used either quantitatively or qualitatively. **FIGURE 15-9** provides an example. Caution is needed when such tools are used, however, because they can easily mask implicit value judgments such as relative weightings between the criteria.

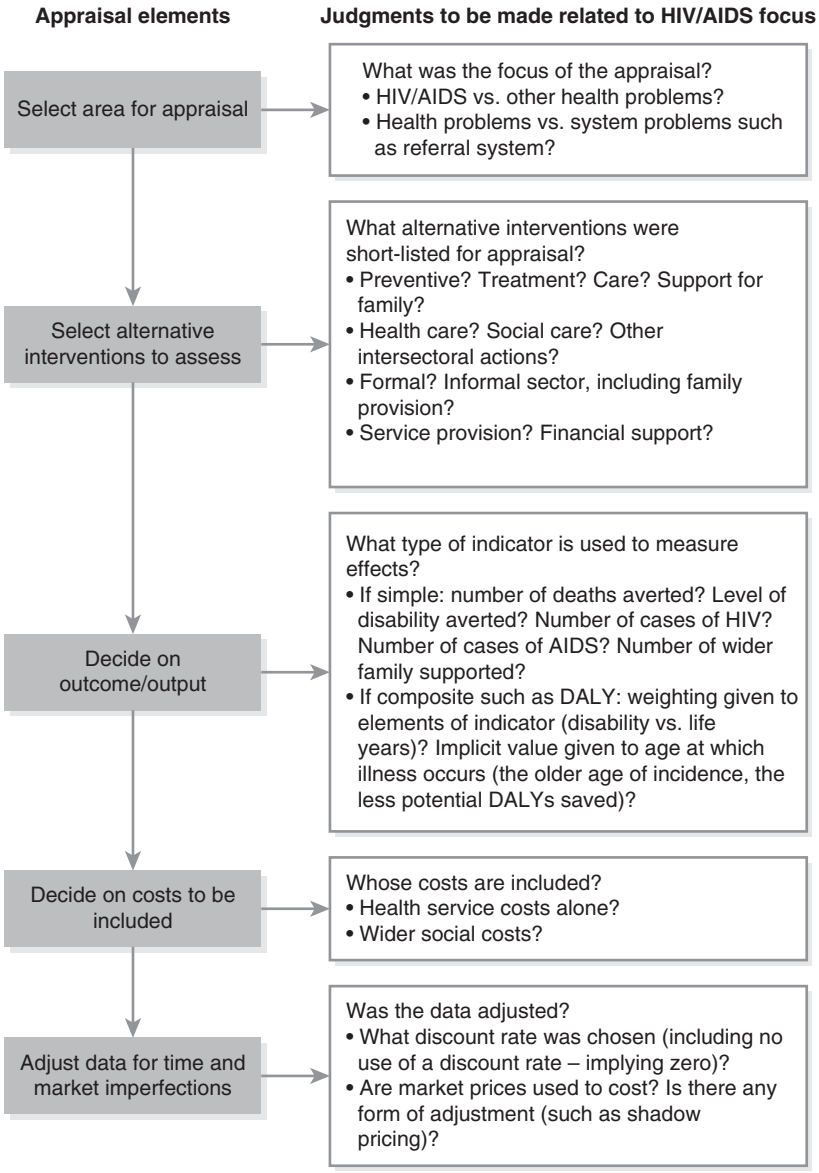


FIGURE 15-8 Examples of value judgments required in economic appraisal using an HIV/AIDS program example.

Reproduced from Collins, C., & Green, A. (2014). *Valuing health systems: A framework for low and middle income countries.* New Delhi, India/Los Angeles, CA: Sage.

	Criteria			
Allocated Score	**Cost per DALY**	**Public Demands**	**Mortality Rates**	**Disability Rates**
4	Measles	AIDS	AIDS	Polio
3	TB	Alcoholism	TB	Alcoholism
2	Malaria		Malaria	
1			Gastroenteritis	

Scoring		
	AIDS	8
	Alcoholism	6
	TB	6
	Measles	4
	Malaria	4
	Polio	4
	Gastroenteritis	1

FIGURE 15-9 Hypothetical example of a multivariable decision matrix.

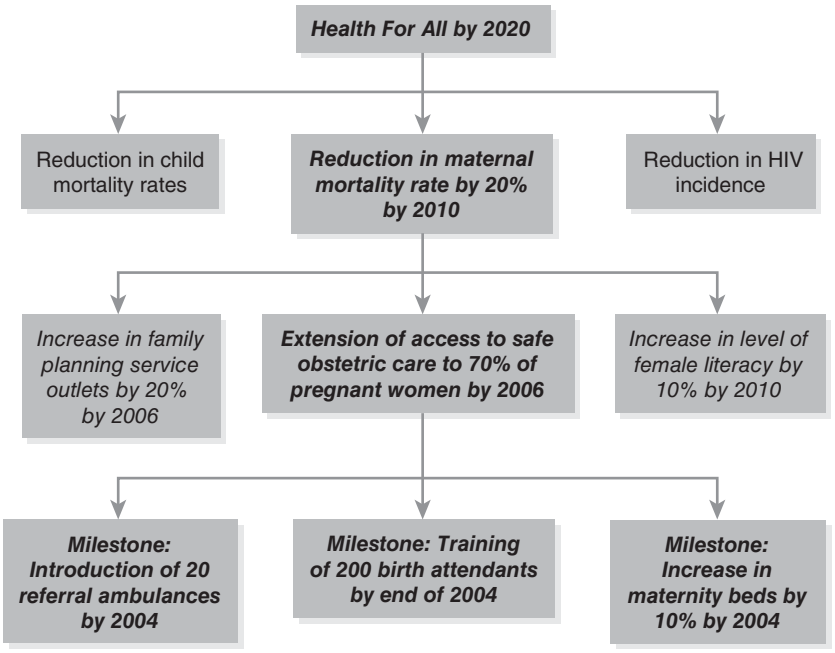


FIGURE 15-10 Example of the hierarchy of objectives.

Courtesy of Dr. Nancy Gerein.

The end result of a priority-setting stage should be a set of clear objectives for an organization. Although different terms are often used by different organizations for objectives (e.g., goals, purpose, aims, objectives, and targets), the important feature is that they are structured in a hierarchy. For example:

- Broad overall health goals achieved through *x*
- Specific health aims related to particular health problems to be achieved by *x*
- Health-sector activity objectives to be monitored by *x*
- Targets that are milestones along the way to achieving aims and objectives

FIGURE 15-10 sets out an example of a hierarchy of objectives. The mnemonic SMART is often applied to objectives to suggest that they should be Specific, Measurable, Attainable, Relevant, and Time bound.

Option Appraisal

For each priority area, there may be various strategies that can be followed to achieve those aims. In the third stage of planning, planners appraise each alternative to determine which is most appropriate according to various criteria. These criteria should include those that underpin economic appraisal techniques (cost and effectiveness) as well as others such as equity, feasibility, and acceptability.

A variety of options are available to any organization, but are particularly feasible for use by the state, which has multiple "carrots and sticks" that it can deploy to encourage or require action by other organizations. In traditional command-and-control planning systems, plans were expected to be implemented through managerial commands and compliance, rather than through incentives and creation of an enabling environment for better performance. With the recognition of the importance of other sectors and providers, a wider range of tools has become necessary. Thus the state may see, as one of its options to be appraised, the contracting out of services to other, nonstate organizations. The end product of the option appraisal stage should usually be a detailed record of alternative options considered, along with a clear framework for assessment and justifications for the chosen planning option.

Programming

The next stage in planning encompasses programming for services and projects. The aims of and agreed-upon approaches for meeting the organization's objectives are brought together in a document that sets out mechanisms to achieve these objectives, along with a time frame and financial plan. The level of detail, particularly regarding the budget and time frame, will depend on whether the plan is primarily strategic or more operational. **EXHIBIT 15-2** outlines the components of a typical plan document.

Logframes (logical frameworks) are often advocated as a means of both ensuring that a logical

approach to project design is followed and providing a means of monitoring progress. **FIGURE 15-11** gives an example of a logframe layout.

The end product of the programming stage should be a detailed plan of action identifying clearly the resource implications, time frame, and distribution of responsibilities.

Implementation and Monitoring

Planning is useful only if it culminates in implementable action. Unfortunately, planning has a poor implementation record, with failure being possible for a variety of reasons—for example, lack of funds, lack of relevant resources, poor timing of inputs, resistance to change, neglect of institutional or legal requirements, unexpected results, poor coordination, and unforeseen circumstances. Frequently, the root cause is poor planning design, often as a result of a failure to recognize the political nature of planning at the design stage and the development of overly optimistic and infeasible objectives.

A key activity that can improve implementation is monitoring. It requires an explicit time frame for well-specified activities and a clear understanding of who is responsible both for implementation and for monitoring of the activities. Monitoring techniques such as Gantt charts set out, in tabular form, activities by their expected dates of completion. In addition, planners must ensure that monitoring does not end up being viewed as an end in itself, but rather as a means of facilitating the achievement of the set objectives. As such, it is important that only the minimum number of monitoring indicators are chosen and *used*, and that monitoring is seen as a supportive (rather than punitive) activity. This understanding is closely linked to the concept of performance management of an organization, which seeks to identify the progress toward the organization's objectives and any barriers to this advancement.

The end product of this stage should be a record of implementation (e.g., services provided) including the challenges faced and any deviations from, or changes to, the original plan.

EXHIBIT 15-2 Possible Outline for a Plan

- Situational analysis, including the health needs or problem being tackled
- Objectives of the plan
- Strategies to meet these objectives, and specific activities as appropriate
- Resources required, including funding to provide the services and sources of these resources
- Timetable with identification of specific roles and responsibilities
- Foreseeable constraints or risks

Objectives	Objectively verifiable indicators	Means of verification	Assumptions/risks
Goal			
Purpose/objectives			
Outputs			
Activities	Inputs		

FIGURE 15-11 Example of a logframe layout.

Note: See Nancholas (1998) for a description of the process.

Evaluation

The final stage of planning involves the evaluation of the plan. Similar to monitoring, evaluation seeks to establish whether and to what degree the objectives have been achieved. Evaluations may be formative (process evaluation) or summative (end evaluation). While monitoring is primarily seen as an in-house activity, evaluation can often be outsourced to an "external" (i.e., not involved in the implementation) person or a group. Evaluation is also regarded by many as a separate activity in itself, involving a series of stages that are similar to those occurring in research (i.e., identification of questions to answer, methodology design, data collection, analysis of findings, and report writing). The key areas that an evaluation is likely to focus on are the following:

- Inputs: Did the resources planned arrive? Were they sufficient for the services provided?
- Processes: How were resources allocated, and by whom? How were services provided, and how was quality assurance performed?
- Outputs: Were the services provided? Were the services appropriate, relevant, and adequate?
- Outcomes: What were the objectives of the activity being evaluated? Were they appropriate? Were the objectives set achieved? If not, why not? Were there any other effects of the activity?
- Impact: Did any health improvements occur as the direct result of the activity? What are the other benefits or unintended consequences of the implemented activity?

The results of the evaluations should be fed into the next round of the situational analysis, thereby completing one round of the planning spiral.

Political Aspects of Planning

The preceding discussion outlined different stages and techniques used in developing plans. It is important to recognize that during all of these stages, there is a need to consider the political dimension of planning, which is inevitable in a process that is designed to determine who gets what in the health sector (**EXHIBIT 15-3**). There are two levels at which planners need to be aware of the political nature of planning as a process. First, they need to recognize that techniques and decisions are rarely value free. Second, they need to understand that techniques such as stakeholder analysis and political mapping may use priorities set at different points in the planning cycle to analyze levels of support for different strategies.

EXHIBIT 15-3 Examples of Political Decisions in Planning

Situational Analysis

- Who chooses the information (e.g., in determining health needs)?
- Whose view is taken? What emphasis is placed on the different information?

Priority Setting

- Whose views are taken?
- Which view of health is chosen, and by whom?
- Which criteria for priorities are chosen (e.g., equity versus efficiency)?

Option Appraisal

- Who chooses, and how, the original options that are appraised?
- Which criteria are chosen, and by whom, for the option appraisal?

Programming

- How are resources distributed, and by whom?
- Who is responsible for the different tasks?

Implementation and Monitoring

- Who is coordinating the implementation, and how it is done?
- Who is using the information from monitoring and how?

Evaluation

- Who is involved in the evaluation?
- What are the criteria for evaluation?

Management of Resources

Management, as we argued earlier, is about resources—how they are generated and developed; allocated; and combined and used with a view toward achieving objectives. Management of resources can be either direct by public-sector managers or indirect (e.g., through contracting out this function to the private sector or other agencies within the public sector), and may be informed, and driven, by the managers' underlying values such as equity, efficiency, and participative and accountable decision making (Collins & Green, 2014). Resources of particular importance to the health sector include those in the financial, human, information, supplies (including pharmaceuticals), and transport categories.

This section first presents an overview of the main issues related to resource generation, allocation, and use; it is followed by a more detailed discussion of how the principal resources—money and staff—are managed. This is followed by brief comments on the management of supplies, transport, and information. In most countries, management of specific resources is undertaken by separate subsystems (e.g., the health management information system) that have their own staff, organizational rules, and objectives. Although much of the discussion in this section refers specifically to management or managers, in reality many issues (e.g., related to resource generation and allocation) apply to planning as well, so they should be of relevance to health planners.

Resource generation and development relates to the different processes for production and enhancement of various types of health resources. Examples of these processes include preservice training of health staff, manufacturing of pharmaceuticals and medical supplies, different types of health financing (e.g., taxation, insurance, user fees), and procurement of transport and technology. Given the complementarity of such resources in their eventual use, it is important that the generation of these resources is planned in a coordinated way. Unfortunately, this is not always the case: There may, for example, be an imbalance between the number of health workers trained and the health facilities in which they are to work.

A broad interpretation of the concept of resource allocation defines it as "the process by which available resources are distributed between competing uses" (Pearson, 2002, p. 7), highlighting the links between resource allocation and priority setting or rationing (Martin & Singer, 2003). A well-known example of the result of prioritization of resources within the health sector is a minimum or essential package of healthcare services. Although some systems of resource allocation focus on nonfinancial resources, such as the distribution of staff, the most common interpretation of the term "resource allocation" relates to financial resources.

Different classifications of *approaches to resource allocation* exist, including a distinction between normative and empirical approaches (Martin & Singer, 2003), as well as negotiation and political compromise, incremental budgeting, and allocation according to health needs (Pearson, 2002). Approaches to resource allocation rarely exist in isolation. For example, the allocation of financial resources for health in Mexico, Nicaragua, and Peru is done through a combination of legal, political, and technical approaches (Arredondo, Orozco, & De Icaza, 2005). Any approach to resource allocation is linked with (i.e., driven by, and likely to affect) the underlying health system values. For example, a true needs-based system of resource allocation will address equity, whereas an incremental approach is likely to reinforce existing inequities.

The *types of resource allocation* can be described using two continua: (1) the allocation of resources across different levels (e.g., between the center and decentralized levels) and (2) the allocation of resources within a single level (e.g., within a district health system). The types of resource allocation also reflect the underlying *methods for allocating health resources*. Resource allocation may include different combinations of (1) *input-based models* (e.g., based on numbers of facilities or staff) and (2) *needs-based models* (e.g., based on identified health needs requiring specific types of health services). Other principles in allocating resources include continuation of historical trends and patterns in utilization of health services.

Different economic tools and techniques are available to support allocation of resources, with the most commonly used tools including cost-effectiveness, cost-utility, and cost–benefit analyses (Arnesen & Kapiriri, 2004; Hutubessy Chisholm, Edejer, & WHO, 2003; Kapiriri, Norheim, & Heggenhougen, 2003). An important caveat here is that, irrespective of which tools and techniques are used, there is a need to ensure consistency between the overall strategic direction identified within the health policy or plan, the underlying values, and the allocation of resources.

From a management perspective, the use of different resources refers to the most efficient and effective combination of available resources to achieve intended objectives such as providing high-quality services. Where different management subsystems are set up to manage different resources, a particular challenge is ensuring the coordinated and complementary deployment of these resources.

Financial Management

A key resource for any organization is, self-evidently, finance. This section looks at the main elements of good financial management.[5] Finance is, of course, important in that it allows human and physical resources to be obtained, such as by hiring health worker and administrative staff, purchasing equipment, and investing in infrastructure. Good financial management supports the health organization in

[5] Acknowledgment is given to Jane Shaw for her input into this section.

the achievement of its objectives. As such, it must be closely linked to other key decision-making processes and, in particular, to planning. Financial management, like general management, is not value-free, and must be viewed as a means rather than an end in itself. The elements of financial management are described next.

Resource Generation

Various sources of financial support are available to a health services organization, and it is important to recognize the interrelationships between these different sources and consider them as a whole in planning. These sources are explored in the *Design of Health Systems* chapter.

Financial Resource Allocation and Budgeting

Financial resource allocation refers to the distribution of block financial resources from a higher level to a lower level in an organization (a vertical process). Budgeting refers to the allocation of financial resources within an organization.

Within the public health sector, financial resource allocation from the center is a common feature. Its importance will depend on the health financing system and the degree of decentralization of resource generation within the health system (Green, Ali, Naeem, & Ross, 2000).

Frequently, financial resource allocation is based (explicitly or implicitly) on one or more of the following factors:

- Previous allocations
- Service or facility patterns and norms
- Capital developments and associated recurrent implications
- Political profile and compromise

These approaches, although potentially nonthreatening to existing budget holders, fail to address issues of efficiency and equity. In contrast, an equity-focused strategy bases its resource allocation on an assessment of the needs of particular areas or population groups. **FIGURE 15-12** sets out a conceptual model for such a process. The potential components for allocation are as follows:

- The population's size, composition, and health needs
- The costs of providing services
- Variations in costs between different areas
- The costs of activities other than health care, such as research or teaching
- Flows of patients across administrative boundaries

FAIR: Formula for Allocation of Internal Resources

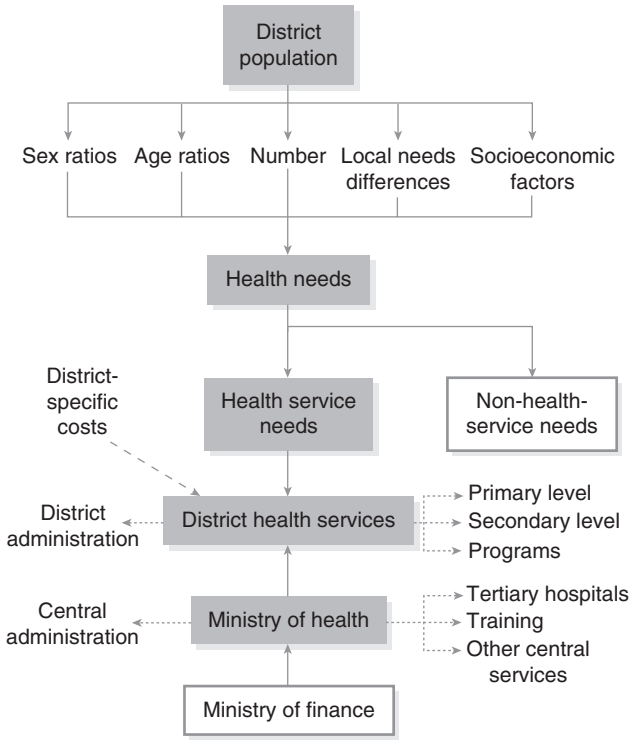


FIGURE 15-12 Conceptual model for needs-based resource allocation.

Data from Green, A. (2007). *An introduction to health planning in developing health systems* (3rd ed.). Oxford, UK: Oxford University Press.

"Health need" is the most significant and most difficult element to measure, as its ingredients include demographic factors; morbidity, mortality, and disability; and indicators of deprivation.

Formulae to allocate resources on the basis of needs have been developed in a number of health systems. The historical U.K. RAWP (Resources Allocation Working Party) model is a well-known example, with similar approaches that target specific groups such as the poor being the subject of resource allocation experiments in other countries such as Cambodia, South Africa, and Uganda (Pearson, 2002). A number of issues arise regarding the implementation of such a formula, however—most importantly, the choice of indicators, and the availability of information. Other issues affecting the introduction and implementation of resource allocation a formula include the absorptive capacity of areas receiving significant additional funding and the political resistance of areas that are "losing" funding. In many countries, it may be necessary to start with a basic allocative a formula based on the size of population being served before developing more sophisticated approaches. One particular challenge that managers often face is balancing allocated funding for local health priorities with the

management of donor funds earmarked for particular disease priorities. Besides potentially distorting local priorities, donor funding can create some unique tensions: New donor-linked incentives and targets can change the behavior of actors within the health service, and potentially result in an overall reduction of local funding to the health sector if the funding is seen as a subsidy rather than as a supplement to the government budget.

Budgets specify the allocation of financial resources within an organization. These statements of intended expenditure are required in the implementation of planned activities, and should reflect the plans of an institution or service. Thus, budgets are vital management tools. Two major forms of budgets are distinguished:

- *Capital/development:* for buildings, large equipment, and vehicles
- *Recurrent/revenue:* for expenses necessary to run the organization, including salaries and supplies

In addition, an organization may develop separate project budgets, which may combine both of these forms.

A budget consists of three components: what it is spent on (line items such as personnel, medicines, and equipment), for what purpose (service or institution or geographic area), and when. Within health organizations, budgets may also be set up on the basis of programs (such as maternal health) that cover all aspects of care irrespective of where it takes place. Such program budgets are better aids in planning because they allow for programs to have different priorities; however, they are much more difficult to manage.

Budgets may be established in various ways. The most commonly used practice is historical incrementalism, whereby the previous period budget is increased across the board. Such an approach is easy to administer and nonthreatening, but it does not reflect major changes in priorities or new developments. A common variant on this approach involves additional increases to a budget (over and above the historical increment) arising from capital developments such as a new building and its associated service.

A contrasting method is activity-based budgeting, whereby the budget is set based on predicted activity levels. This technique allows for a clear link between the budget and the service objectives. An extreme example is zero-based budgeting, in which no prior assumptions are made, and every item in the budget for each year has to be fully justified.

Zero-based budgeting is more time-consuming and generally more politically sensitive compared to historical incrementalism, as it requires information on both the level of activity and the costs of a unit of activity. These costs may, of course, change depending on the level of activity and on the relationship between fixed and variable costs.

A possible timetable for budgeting is given in **TABLE 15-4**.

Expenditure Drawing and Disbursement

The next element is that of expenditure against a set budget. Each financial management system will have mechanisms to authorize named officials to incur expenditures, within certain predefined limits and conditions, that reflect the level of decentralized authority.

Expenditure Monitoring and Control

Monitoring of expenditures against the budget throughout the year is an important management function. Critical to this process is a good information system. Each management system will have its own set of accounting methods. Such a system needs to be able to give the manager up-to-date information not only on past expenditures, but also on any commitments already made. For managers to understand the real implications of such information, they need to have an indication of the likely expected profile of expenditures throughout the year; that will allow them to discover any variance from the budget early enough that they can take the necessary remedial action. Needs for some resources (such as staff salaries) may be evenly spread out over the course of the year; others, such as medicine purchases, may occur more sporadically (in batches) and may reflect within-year differences in usage due, for example, to seasonal disease incidence. Capital equipment is even more lumpy in terms of expenditure patterns.

A manager needs to be able to compare actual expenditures against the expected expenditure profiles and understand the likely reasons for possible differences. An example of a simple monitoring tool appears in **TABLE 15-5**. It shows a monthly management statement after nine months of the year. The "Budget to Date" column shows what might be expected to have been spent three-fourths of the way through the year. It is assumed here that spending will be equal each month (which is unlikely, of course); more sophisticated estimates could be made. The variance columns show the relative (not actual) over- or underspending.

TABLE 15-4 Annual Budget Cycle for a Three-Level Decentralized Health System

Month	Activity
1	The financial year begins.
5	The ministry of health receives its provisional annual allocation from the central government, together with any special constraints or conditions.
6	The ministry of health issues broad resource allocation guidance to regions on the basis of the provisional allocation.
7	Regions issue broad allocations to districts on a similar basis.
8	Budget holders develop and return proposals showing the following: ■ Review of service targets in line with the plan ■ Estimated expenditures for the previous year ■ Estimated expenditures for the current year ■ Reasons for under- or overspending ■ Budget proposals for the following year, costed and showing how they will meet planned service targets
8	Budgets are totaled and reconciled first at the regional level and then at the ministry of health.
9	Adjustments are made for the following reasons: ■ To reflect national policy ■ To reconcile with other budget proposals ■ To reflect constraints ■ To reconcile with central government allocations
9	Discussions are held with the central government.
10	Adjustments are made with the input of service managers.
11	Informal approval of the budget is granted.
12	The government approves the budget.
1	Budgets are issued to budget holders.

Modified with permission of Oxford University Press from Green, A. (2007). *An introduction to health planning in developing health systems* (3rd ed.). Oxford, UK: Oxford University Press.

Once managers have identified a projected over-expenditure (or under-expenditure), they need to consider the various options open to them to prevent this departure from the budget. Potential actions include the following:

- Seeking additional funds
- Instituting cost-control mechanisms, such as freezing expenditures
- Revising service objectives
- Reallocating funds from one budget (line item) to another

To select the appropriate option, managers need to have a clear understanding of the reasons for the overspending or underspending, so they can suggest the best remedial strategy. It is also important that the implications of any remedial actions be weighed against their effects on the organization's objectives. Cost-control mechanisms, for example, will affect services in terms of either the quantity deliverable, the quality, or the efficiency with which the services are delivered. The first two of these factors will, in turn, affect the ability to achieve the service objectives.

TABLE 15-5 Monitoring Tool Example

				Variance from Budget to Date (+, under-expenditure; −, over-expenditure)		Projected Year-End Expenditure	
Month: 9	Full-Year Budget ($000)	Expenditure to Date ($000)	Budget to Date ($000)	Amount ($000)	Percentage	Projected to Estimated ($000)	Comment
Personnel	600	400	450	+50	+11.1	500	Underspending is due to recruitment difficulties
Medical supplies	150	140	112.5	−28	−24.4	160	Earlier epidemic, likely to result in final overspending
Transport	80	75	60	−15	−25.0	95	Earlier epidemic
Utilities	50	10	37.5	+28	+73.3	50	Delays in invoicing
Other	120	100	90	−10	−11.1	100	
Total	1,000	725	750	+25	+3.3	905	Overall underspending projected

Auditing

The last elements of the financial management process are internal and external auditing. An internal audit is a process of testing and validating the financial control and accountability systems to ensure that they are working correctly. This task is performed by a special department of the organization, which is separate from the normal accountancy group. In external auditing, the accounts of an organization are subjected to independent external scrutiny to confirm that they do not contain errors or hide fraud. For any organization, auditing is an important part of its governance.

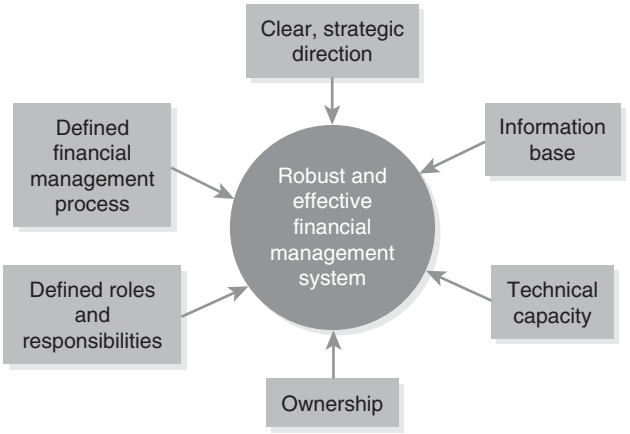


FIGURE 15-13 Requirements for an effective financial management system.

Requirements for an Effective Financial Management System

FIGURE 15-13 identifies a number of critical features of an effective financial management system. If these aspects are lacking, there is a real danger that the system will not perform its expected task—that is, ensuring that the desired strategic direction is followed.

Management of Staff [6]

The management of health staff was for some time a somewhat neglected issue. Indeed, only since the mid-2000s have international policy makers explicitly refocused their attention on human resources for health (WHO, 2006b). In the past, policies largely sought to reform health systems through new financing arrangements or organizational changes, such as decentralization. Yet staff costs are a major part of the health budget, and the quality of staff is vital to ensure the effectiveness of health care (Buchan, 2004; Rondeau & Wager, 2001). Furthermore, the development of staff management systems and skills throughout the health system increases in importance as decentralization and hospital autonomy spread out the management responsibility throughout the health system.

In developing staff management capabilities, several factors must be considered. First, the inequitable distribution of health staff must be addressed. It is often difficult to persuade health staff to work in poorer areas, where incentives such as good living conditions and the possibility of supplementing income through private practice are lacking. Second, some health systems are undergoing important changes such as decentralization, or facing contextual changes such as conflicts, that are affecting how health staff are managed and how their capacity is enhanced (Kolehmainen-Aitken, 2004; Prashanth et al., 2014; Rockers Peter & Bärnighausen, 2013; Roome et al., 2014). This kind of transformation can mean important changes in civil service arrangements, for example (as shown in **TABLE 15-6**). Two other influences on how staff are managed are professional authority and power, and corruption and patronage.

There are many systems and approaches for managing staff, ranging from defining staffing needs and jobs to employing people and developing their performance. The discussion here focuses on a number of key issues and functions.[7]

Staffing Review

An important issue is how the workforce can be used to improve performance in meeting health needs. A review may be undertaken, with key questions focusing on the number of staff employed, possible substitution by technology, and changes in the skill mix and grades currently employed (Strike, 1995). A simple technique for defining staffing needs of a healthcare unit is provided

TABLE 15-6 Decentralization and Options for Civil Service Structure

Civil Service Option	Commentary
Centrally controlled system	Terms for the civil service are centrally controlled and staff are seconded or transferred to local units.
Health-sector centrally controlled system	Health-sector staff are taken out of the national civil service and included in a national system exclusively for health staff.
A national local government system	All local government units operate under the terms and conditions of a national system that is specifically designed for local government staff.
Local public service commissions	A public service commission can be tasked with and empowered to investigate, monitor, and evaluate the organization and administration of the public service, as well as tasked with promoting measures to ensure effectiveness and efficiency.
Decentralized unit	The decentralized unit determines the terms and conditions and employs its own staff.
Mixed system	Following decentralization, existing staff are kept as part of the national civil service, while new staff are employed under the new decentralized system.

Data from Collins, C. (1994). *Management and organisation of developing health systems*. Oxford, UK: Oxford University Press; Kolehmainen-Aitken, R.-L. (2004). Decentralisation's impact on the health workforce: Perspectives of managers, workers and national leaders. *Human Resources for Health*, 2, 5.

[6] Acknowledgment is given to Tim Martineau for his input into this section.

[7] A more comprehensive and in-depth analysis would cover issues such as employee relations, change management, staff communication, and conflict management.

by Kolehmainen-Aitken and Shipp (1990); it involves the definition of job responsibilities and calculation of a "standard workload" for each person (e.g., number of outpatients attended per day). This value is compared with projected workloads in the health facility to estimate staff required. The possibility of using such planning tools is restricted by the degree of autonomy local managers are given, the existence of patronage, the availability of information, the capacity of managers to use the tools, and the willingness of staff to move.

At a facility level, WHO (2010b, p. 2) has developed human resources management tools and guidelines that (1) determine how many health workers of a particular type are required to cope with the workload of a given health facility and (2) assess the workload pressure endured by the health workers in the facility. The tools take into account multiple services that need to be delivered in a facility.

Defining Jobs and Employment

The defining of jobs and staff employment may be viewed as a set of the following logical steps:

1. Job analysis: assessment of which tasks are required and which category and level of staff are required
2. Development of a clear job description: a detailed description of the planed duties and responsibilities, as well as personal specifications (both essential and desirable) for the anticipated candidate
3. Advertising: inviting potential candidates to apply for the position(s)
4. Selection and recruitment: processing of applications and selecting (typically following an interview, test, or another method to assess the candidate's appropriateness) the most suitable candidate(s) for the position against the personal specifications
5. Posting: assigning the recruited candidate to a workspace
6. Induction: initial training and discussions to help candidate understand the organization and start performing the duties

The definition and use of clear principles and rules and an open and transparent process can facilitate an effective and fair employment process.

Staff Payment

The issue of payment is related to the staff grading system and form of public service system (see the earlier discussion). Changes in health systems, such as decentralization, can lead to corresponding changes in the source and level of pay. The form of payment may be based on time spent working, physical results produced, individually assessed performance, or a combination of these factors. Payment by time is simple, whereas performance-related pay is both complicated and controversial (Alimo-Metcalfe, 1994; Kessler, 1994). The latter approach assumes that staff will work more, and will perform better, if their earnings are related to their performance. In addition to practical difficulties in measuring performance, drawbacks of this strategy include its potential to introduce conflict between staff members and to counteract a teamwork approach.

Reliance on material incentives downplays the important motivation that public health staff may derive from the inherent usefulness of their jobs and public service (Perry, Hondeghem, & Wise, 2010; Vandenabeele, 2007). In fact, the extent to which pay actually acts as a motivator will vary according to the context and the type of staff concerned.

Delegation

Delegation involves transfer, trust, responsibility, tasks, and authority. Delegation is a transfer of authority and tasks between two people or groups of people. It entails entrusting part of your own authority to someone else (French & Saward, 1983). Although the person receiving the delegation has the responsibility or obligation to satisfactorily perform his or her duties, the person doing the delegating does not relinquish responsibility for that task. Rather, that individual retains "responsibility for that person's exercise of authority" (French & Saward, 1983, p. 122). This responsibility can also involve authority over resources—an issue that is important to recognize in the act of delegation.

The purpose behind the delegation is likely to affect its form and should be clear. For example, delegation might be used to address work overload or underload, to motivate staff, to bring decisions nearer the point of service delivery, or to gain a better understanding of service needs and achieve more timely decisions and increased flexibility and adaptability. Greater ownership may also be achieved with a more participative management style.

Delegation is not easy, particularly where there are constraints on its effectiveness. In a setting of limited resources, for example, delegation may not transfer authority over those resources. Managers may resist losing authority and control; staff may lack the knowledge, skills, and values to practice delegation. They may also lack motivation where the delegation has no additional incentives and a culture of blame exists in the organization that is not conducive to delegation.

It is important to identify the objectives involved to communicate them to staff, monitor the effectiveness of the delegation, and design the means of delegation. The delegation needs to be clearly identifiable in terms of tasks, authority, control over resources, and its limits. A particular issue is the capacity (e.g., skills and experience), support, and motivation of the person receiving the delegation. Is that person equipped to receive the delegation? A useful concept in making this determination is staff "maturity" (Hersey, Blanchard, & LaMonica, 1978), which encompasses the skills, experience, and willingness to take on new responsibility.

Performance Management

In the field of staff management, performance management is frequently associated with annual confidential reviews or appraisals that are bureaucratic and routine, and that do little to develop performance. Effective performance management requires not only reinvigoration of such yearly events, but also a broader approach of continuous performance management (**TABLE 15-7**).

Performance can be improved by focusing on its various aspects. Grindle (1997), for example, has examined organizational culture and identified which aspects of an organization lead to good performance. These elements include an organizational mission and a strong sense of attachment to it among staff; good management relations, such as fairness and teamwork; positive expectations concerning staff performance; and institutional autonomy in staff management.

Incentives and Motivation

Incentives and motivation represent an area in which there has been considerable theoretical debate and that has provided useful indications about possible causes of staff motivation. Unfortunately, theories on motivation often fall short of providing health managers with a comprehensive framework that would allow them to understand fully the wide range of factors that can motivate or demotivate staff. Furthermore, the list of such factors can be so long as to render it useless as a managerial guide.

Another problem associated with attempts to explain motivation is that different individuals and

TABLE 15-7 Steps in Performance Management

Steps	Commentary
1. Explain and reinforce the logic of performance management.	Managers explain to staff the organization's performance management and supervision practice.
2. Understand job context, content, and relations.	Performance management needs to be based on appreciation of the job content and the often complex relations of authority it involves. When they understand the job context, managers and staff should recognize the various factors that affect performance.
3. Agree on performance criteria.	Managers and staff develop a joint understanding about the meaning of "effective work." This agreed-upon standard accompanies an understanding of the conditions that need to be in place (e.g., logistics support, flow of funds, skills) for actual performance to meet these standards.
4. Compare actual performance with standards.	Managers and staff monitor performance using the agreed-upon performance indicators (both quantitative and qualitative).
5. Discuss and agree on action.	This step involves both feedback on performance with dialogue and agreement on the causes of this performance and possible changes required (including agreement on training needs).
6. Take action and monitor.	Action may need to be taken in areas, such as job context, content, and support. The effects of this action need to be monitored.

Data from Bacal, R. (1999). *Performance management*. New York: McGraw-Hill; McMahon, R., Barton, E., & Piot, M. (1992). *On being in charge: A guide to management in primary health care* (2nd ed.). Geneva, Switzerland: World Health Organization; Weightman, J. (1996). *Managing people in the health service*. London, UK: Institute of Personnel and Development.

groups of health staff respond to different factors (Weightman, 1996). Attempts to generalize across groups are not easy, although a general distinction may be made between "satisfiers" and "dissatisfiers" (Herzberg, Mausner, & Snyderman, 1959).[8] Dissatisfiers are those factors that in themselves do not necessarily motivate workers, but whose absence can demotivate individuals; they are related primarily to attracting people into jobs and ensuring their retention. Typical examples are pay and security. In contrast, the presence of satisfiers does lead to improved motivation and performance; typical examples are career development and job achievement. Notably, the following factors can have a bearing on staff motivation (Henderson & Tulloch, 2008; Ogundeji, Jackson, Sheldon, Olubajo, & Ihebuzor, 2016; Willis-Shattuck et al., 2008):

- Career development
- Job characteristics
- Working and living conditions
- Managerial style and organizational culture
- Job rewards (financial and nonfinancial)

In using such a checklist of motivational factors, managers need to distinguish between the satisfiers and the dissatisfiers. Job characteristics, for example, can be the cause of demotivation for staff if too many duties are imposed on the job holder; if the job lacks variety, scope, and delegated authority; and if the job is badly defined and, therefore, the object of conflict. Even so, job holders may feel that the job is meaningful and contributes to social welfare and solidarity. Staff working in the poorer areas of a country may perceive that they have inadequate working conditions (e.g., office accommodation, information technology) and living conditions (e.g., schools), which are also dissatisfiers.

In contrast, transparent and open management, showing an interest in good supervision and communication, staff involvement in policy making and problem solving, good teamwork, and reinforcement of ethical standards can be positive forces (satisfiers). The pattern of rewards and incentives is important as well, and can include the extent to which the basic needs of health workers are met and are considered equitable, the regularity and security of payment, and nonfinancial rewards such as stability and status of employment. Likewise, the possibility of career development can be important. Building workplace trust relationships, through supervisory support, appreciation and recognition, communication, and feedback are important for promoting and sustaining intrinsic motivation—that is, working because the work is inherently enjoyable, rather than just for a tangible reward (Okello & Gilson, 2015).

Supervising

Like the other management functions, supervision is a process, rather than an end in itself. Supervision supports staff through a process of continuous professional improvement, includes all levels of the organization, and uses information from the staff, service users, and the community. It is defined as "[a] range of measures to ensure that personnel carry out their activities effectively through direct, personal supervisory contact on a regular basis to guide, support, and assist designated staff to become more competent in their work" (Management Sciences for Health [MSH], 2003, p. 3).

Among the roles of the supervisor are the following (Bosch-Capblanch, Liaqat, & Garner, 2011; Collins, 1994):

- Defining tasks to be carried out and planning their implementation
- Providing professional and technical advice
- Problem solving and decision making
- Ensuring performance standards (technical, ethical, and legal) and providing feedback
- Personal counseling and employee motivation
- Providing referral, broker, and advocate services
- Ensuring appropriate support for the job
- Providing training and development support

A distinction may be made between line management authority and supervisory authority (see Figure 15-2). In particular, the two types of authority may be separated to allow the line manager to widen his or her span of control (Rowbottom & Billis, 1987). Depending on the relation between supervisor and line manager, the supervisor can also be involved in disciplinary actions and career development.

Despite its importance, the value of good supervision is often overlooked. Pressures on time and the demoralizing lack of incentives and overall support to managers and health workers may shunt supervision aside, such that it is perceived as a peripheral activity. Yet the importance of effective supervision is well recognized. For example, Trap and associates (2001) evaluated the effect of supervision on pharmaceutical management in Zimbabwe and concluded that it had a positive impact, emphasizing the importance of training in supervision.

[8] For a brief description of Herzberg's approach and its application to a LMIC context, see Dieleman et al. (2003).

Managers have a variety of options at their disposal with which to improve the supervisory process. Both individual and group supervisory techniques may be considered, and a combination of formal and informal methods may be implemented (Collins, 1994).

The style of supervision is also important. The concept of supportive supervision was emphasized by Management Sciences for Health as one of the most effective ways of supporting improvements in staff performance (MSH, 2003). A distinction can be made between supervision that focuses on the supervisee's performance and supervision that focuses on the more personal relationship factors. The style of supervision employed may be contingent on a variety of contextual factors, some of which are discussed in this chapter.

Finally, the supervisory process may be formalized. A three-stage plan for such a process can consist of preparation, supervision, and follow-up.

Teamwork

Teams are widely recognized as a positive influence in the health sector and can improve decision making, problem solving, and innovation; motivate team members; improve communication, collaboration, and support; and allow for training. Teams are used in a variety of health-sector settings and for a variety of reasons (**EXHIBIT 15-4**).

EXHIBIT 15-4 The Importance of Teams for Health and Health Care

Teams can operate at senior management levels within a district health system (DHS). Sometimes referred to as a district health management team (DHMT), such a team is made up of staff with a range of competencies and capabilities who manage under conditions of complexity. Teams can also be used to work throughout the DHS. In South Africa, district clinical specialist teams (nurse–physician dyads working in three key disciplines) were introduced to support clinical governance through clinical supervision and mentorship at the facility level as well as to support the DHMT by supporting organizational collaboration and communication. Teams can also be developed to deliver preventive health services at the interface between the community and the health services platform, such as ward-based outreach teams as well as more specifically in primary care to promote holistic patient care, known as interprofessional teamworking.

Data from Doherty & Gilson, 2015; Kwamie, Agyepong, Dijk, 2015; Mampe, Schneider, & Reagon, 2016; Oborien et al., 2015; Voce et al., 2014; Xyrchis & Lowton, 2008.

Although Exhibit 15-4 suggests some reasons for developing health teams, such groups do have the potential to be associated with problems. Accountability within the team may be dispersed and lost, and team meetings can put pressure on time.

Ultimately, the effectiveness of teams will depend on factors such as definition of the objectives and tasks of the team, team members' willingness to modify their working styles to fit the team, good motivation among team members and leadership, effective team meeting skills, getting team size and members' characteristics right, understanding the changing nature of teams, good team action planning, and achieving a balance between team cohesion and diversity (Noakes, 1992). An increasing body of knowledge is available on how to build better teams. In addition, interventions that focus on team design, training and development, and team leadership can be applied to improve team effectiveness (Katzenbach & Smith, 2005; Kozlowski & Ilgen, 2006).

Managing Staff Development

Staff development is an important aspect of individual-level capacity development and a factor that affects the motivation and performance of staff members. It can be done through various combinations of training, delegation of tasks, and supportive supervision. Although training is often associated with off-work courses, there is scope for incorporating staff development into the management process. We have already noted the importance of induction for new recruits. **TABLE 15-8** looks beyond this step, suggesting informal processes that allow training to become an integral part of the management process. Many of these mechanisms avoid taking workers out of their environment and disrupting service delivery. Moreover, learning in the workplace can be more relevant and realistic.

Unfortunately, interest in these training mechanisms may arise for the wrong reasons, such as cost cutting or the ineffectiveness of poorly managed training courses. Learning environments are not easily created at work, especially where employees face strong pressures to ensure service delivery. Furthermore, when carried out properly, on-the-job training is not an inexpensive option.

It is not enough to think about the development of individual management competencies in isolation from the broader complex contextualized settings within which managers work. Indeed, managers will be able to put acquired individual knowledge and skills into practice only given the right conditions. WHO (2007a) considers three other dimensions to be critical to leadership and management capacity development:

TABLE 15-8 Forms of Training Within Management

Form of Training	Commentary
Self-development	A learning agreement may be used in which the conditions are provided for staff to take the initiative in planning their own training. This experience can be isolating, however, and those workers who are less able and more in need of training may be less successful in self-development. Also, there is no guarantee that the staff view of self-development will coincide with the interests of the service.
Shadowing	In shadowing, a worker learns from another employee through observation. Potential drawbacks are that the "shadower" may be exploited as a "free helper," observations may not be systematically recorded, and the "shadowed" individual may be unconvinced of the usefulness of the procedure. Care needs to be taken that the experience is relevant.
Mentoring	Mentoring may involve an experienced colleague (not the worker's manager or supervisor) taking on roles such as providing advice, setting specific learning tasks, and suggesting new work options. Recognizing the problems that can occur within the context of this relationship, mentoring should be voluntary and parties' roles should be clear.
Supervision	Although there is a danger of overloading this important relationship, there are opportunities for supervisor and supervisee to agree upon and implement plans for skills development and problem solving.
Delegation	Delegation may provide an opportunity for developing staff, although the worker receiving the delegated task should have sufficient capacity to perform it. Good supervision is necessary to take full advantage of the possibilities inherent in delegation.
Secondment	Staff can gain from the experience of temporarily working in another unit or organization. Care needs to be taken to ensure that the secondment experience is relevant to the responsibilities of the original job and that opportunities are given for learning.
Job rotation	Staff may undergo planned movements throughout the organization to develop new knowledge and experience.
Action learning	Learning through participation in completing tasks in a challenging environment may be useful. Such learning must be carefully planned and should not disrupt service delivery. This strategy is particularly useful to ensure a deep and engaging approach to learning.
Group-based work	Such work can involve a range of activities, such as study circles, support groups, professional and occupational groups for developing quality standards, and specific group meetings.
Reflective practice and journaling	This form of staff development involves thinking and writing about what has happened during the day. The purpose is for managers to pause and reflect on how they have coped during the day, how they have behaved, and what they have learned from the day—a practice that promotes reflection and learning.
e-Learning	There are multiple ways in which managers can learn online, ranging from access to readings in shared online folders, to modular or full coursework online, to highly interactive web sessions with multiple participants.

Data from Kerrigan, J. E., & Luke, J. S. (1987). *Management training strategies for developing countries*. Boulder, CO: Lynne Rienner Publishers; Storey, J. (1994). Management development. In K. Sisson (Ed.), *Personnel management: A comprehensive guide to theory and practice in Britain* (pp. 365–396). Oxford, UK: Blackwell Business.

an adequate number and deployment of managers throughout the health system, the existence of functional critical support systems to help in decision making (e.g., information and financial services), and the creation of an enabling working environment (e.g., delegation, stakeholder management, cultural, and economic context). Collins and Green (2014) argue that the values of the health system can be found in these broader dimensions for capacity development. For example, a health system with equity as its value base will regulate the migration of health workers from LMICs, will promote an organizational culture that incorporates a public-sector ethos, and will ensure that managers are treated in a fair and empowering fashion to support self and team development.

Personnel Administration

Most staff management functions outlined previously may be performed by either line managers (and supervisors), personnel specialists, or administrators. Personnel specialists can take on strategic, advisory, and operational roles (Cole, 1988). The administrator category encompasses the mostly routine administrative and procedural activities associated with functions such as recruitment, selection, pay administration, staff contracts, staff grading, promotion, transfer, staff communication, health, safety, and welfare in addition to maintaining a human resources information system (Strike, 1995).

Supplies Management[9]

The management of supplies includes those resources related to health technologies, such as medicines and diagnostics, and nonmedical items (such as food, fuel, and cleaning materials). This section focuses mainly on medicines, although many of the principles apply to other supplies.

As much as one-third of a country's budget for health care can be spent on medicines. Thus, as one would expect, this is an area of political interest (see the *Pharmaceuticals* chapter). It is also a complex area requiring, among other things, the choice of new health technologies at global, national, and local levels (Frost & Reich, 2008), and a process of procurement (including international and national tendering) characterized by complicated technical specifications and strict quality assurance throughout the process.

Effective management of pharmaceuticals requires international-, national-, and local-level management processes. At the international level, standards of quality, safety, and efficacy need to be set and monitored. An example of such a system is the WHO prequalification project, which since 2001 has monitored the compliance of selected HIV/AIDS, tuberculosis, and malaria medicines with unified standards of quality, safety, and efficacy (WHO, 2010c). Another important global public health task is the control of counterfeit medicines in international trade (International Medical Products Anti-Counterfeiting Taskforce [IMPACT], 2010). At the national level, in addition to robust systems for ensuring quality, safety, and efficacy of medicines, the appropriate mechanisms for licensing of medicines and regulating costs are important. Local challenges concern the storage and distribution of medicines in often inhospitable climatic and inaccessible geographic conditions with limited resources.

Problems with the availability of medicines in many services can be a major source of community complaints and irritation for health staff. In part, this issue arises because of the scarcity of material resources and the economic context. Other challenges relate to the existence of an underpaid and insufficient workforce with inadequate training, which may lead to poor management and corrupt practices. Poor management, for example, may include inadequate determination of need and demand, inappropriate selection, long delays in procurement, weak distribution systems, inadequate systems for medicine storage, and unsuitable use.

Supplies management is sensitive to changes in health systems and health policies. Reorganization of this system, for example, may result in separation of the regulator, purchaser, and provider functions. Decentralization requires new definitions of responsibilities in supplies management. Potential economies of scale in central purchasing must be reconciled with the forms of local supplies management and accountability and their associated responses to local needs and flexibility. National medicine policies should include key issues such as the pricing and taxing of essential medicines, the development of legislation and regulatory capacity, the training of health workers and education of the public, and implementation of international policy initiatives such as those developed by the Global Fund to Fight AIDS, Tuberculosis and Malaria (Anderson, Huss, Summers, & Wiedenmayer, 2004).

The supplies management system for pharmaceuticals can be seen as comprising four basic stages:

[9] Acknowledgment is given to Mayeh Omar and Reinhard Huss for their inputs in this section.

selection, procurement, distribution, and use (MSH & WHO, 1997). These four stages, which are linked and cyclical, need to be made operational to ensure that the right quantity and quality of medicines are located and used at the right time and right place. **EXHIBIT 15-5** outlines each of the stages and identifies some of the key issues and techniques involved in this process. A more detailed description of management of medicines is provided in the *Pharmaceuticals* chapter.

An important feature of supplies management is the coordination of the various resources such as health technologies, information, money, staff, transport, and nonmedical items. Several functional subsystems are required for good supplies management, such as health management information, regulation, financial/budget management, and transport management. For example, the supplies process must be linked to budget profiling. Also, efficient supplies management relies on a good management information system to monitor items such as stocks and treatment supplies. Furthermore, an effective supplies system is contingent on appropriate staff training and supervision, as was shown in a study conducted in Zimbabwe (Trap et al., 2001).

Transport Management

The importance of managing for an effective transport system should not be underestimated. Transport is central to ensuring access of patients to health services and the effective working of the referral system through both the emergency ambulance service and the transport of nonemergency patients. Health policies based on improved access and the principle of equity should

EXHIBIT 15-5 Stages of the Supplies Management System

Selection

Selection should consider the needs of the health units and programs, in addition to understanding the health situation of the particular communities concerned, their unmet needs, and the way in which these factors affect medicine requirements. Among the key considerations are the medicine requirements of the different levels of health care, priorities set, targeting of specific groups (e.g., children younger than 5 years), the use of an essential medicines list and generic medicines, and the design and implementation of quality assurance specifications.

The quantities required need to be estimated. Such estimates can be constructed via methods that use (1) population-based data on morbidity and mortality complemented by norms, (2) service-based data on diagnoses (and frequency) complemented by standard treatment norms, or (3) historical consumption data. Each of these methods has both pros and cons. For example, the second method fails to take into account unmet demand; the third method may be the easiest to use, but fails to account for changes that have occurred since the data were collected or are now occurring. Adequate supplies have to be maintained over time, so attention needs to be paid to issues such as consumption patterns over time, lead time for procurement, safety stocks, and reconciliation of medicine needs with available resources.

Procurement

Medicines are obtained through a procurement process, which involves actions such as following purchase procedures (e.g., tendering, negotiating), selecting the supplier (based on criteria such as price and quality through inspections), clarifying the terms of supply and supply periods (fixed or variable intervals), monitoring order status, and receiving and checking the medicines.

Distribution

Distribution can be based on either a "push" (kits) system or a "pull" (inventory) system. In the former type of system, quantities of medicines are sent at regular intervals, based on estimates of anticipated usage. The latter type of system requires the health unit to order medicines according to need. Use of this approach assumes, for example, the existence of an adequate stock control system and good communications for ordering and delivering. Among the issues to be considered are the simplicity, regularity, and reliability of the push system, in addition to how it compares with the greater sophistication and adaptability of a pull system.

Use

The use stage involves "diagnosis," "prescribing," "dispensing," and "patient compliance." A variety of indicators can be developed to monitor and evaluate medicine use.

take full note of the role of transport systems in reducing the direct costs of accessing health care, as those costs can impose economic consequences on households (McIntyre, Thiede, Dahlgreen, & Whitehead, 2006; Thiede, Akweongo, & McIntyre, 2007).

The management and operation of primary healthcare units and the implementation of community health programs require transport. Community work that involves nonresident health workers also relies heavily on good transport. Likewise, mobility and transport are vital for effective management such as supervision and operation of a supplies system. Finally, transport consumes a significant portion of the health services budget.

Although there are undoubtedly good transport practices that can be applied, problems in transport management are also commonplace, including lack of management skills and systems. For example, there may be an insufficient number of drivers, inappropriate or poorly maintained vehicles, no supervision, no vehicle scheduling or usage information, or poor maintenance. A tendency to treat transport as a free good may be one cause of a lack of cost-consciousness. Donor involvement may lead to a multiplicity of vehicle makes, making it difficult to maintain an effective store of spare parts and capacity to perform repairs. Corruption and the appropriation of transport for private use may also diminish the effectiveness of the transport function. **EXHIBIT 15-6** outlines measures to improve transport management and planning.

EXHIBIT 15-6 Measures to Improve Transport Management

- Include consideration of transport in the health policy, planning, and programming process.
- Determine the organizational arrangements and responsibilities for transport management and planning.
- Ensure financial management arrangements for transport.
- Develop vehicle programming.
- Develop fleet management, including vehicle maintenance and repair and operational norms for transport use.
- Manage and develop human resources for transport.
- Use information for improving transport performance.
- Clarify the contributions made by donors to transport management and planning.

Data from Collins, C. D., Myers, G., & Nicholson, N. (1992). A successful transport scenario for the health sector in developing countries. *World Hospitals, 28*(3), 9–14; Transaid. (2001). *Transport management manual*. London: Transaid Worldwide.

Attempts to develop a comprehensive policy for transport management have been made in countries such as South Africa (Department of Health and Developmental Social Welfare, 1999). Transaid, a transport international NGO working in the health sector in LMICs, has developed a transport management system geared toward service delivery organizations that includes detailed tools for transport management. Two key elements within its system are operational vehicle planning and using information (Transaid, 2008, 2017),

Ideally, the operational planning of transport will be coordinated by a dedicated transport officer, though in many cases this function can be centralized to ensure efficiency of operations. The discussion here assumes that transport in the organization is pooled and that appropriate decisions are made on how to use it best. Users of transport, such as health programs and units, first identify their transport needs. Subsequent decisions are then made based on, for example, service delivery and urgency. Vehicles are programmed, with every effort being made to meet needs and priorities, effectively monitor the process, and ensure the transparency of the process.

The use of information for improving performance is a theme throughout resource management and is particularly evident in transport. Managers responsible for transport need to pay attention to developing—among other documentation—log sheets, vehicle maintenance records, fuel consumption reports, budget management instruments, and reports on accidents and incidents. Key indicators of performance need to be developed relating to the distance traveled (per vehicle), utilization of fuel, vehicle running costs, vehicle availability, and utilization, together with information on the safety of each vehicle and the extent to which the vehicle meets service needs.

Information for Management and Management of Information

One resource that the manager is both responsible for and relies heavily on is information. This section looks at issues related to the management of information, including the development of an organizational culture of evidence-informed management. Information systems provide "information support to the decision-making process at each level of an organization" (WHO, 2004, p. 3). A health management information system (HMIS) is "an information system specially designed to assist in the management and planning of health programmes, as opposed to delivery of care" (WHO, 2004, p. 3).

Types of Information

Managers are constantly bombarded with, and attempting to make sense of, information. Most tend to consider information for use in management as that which is explicit, and often the information gold standard is seen as quantified data. In practice, however, most managers continuously use a considerable amount of information without being aware of it—often information derived from their own experience and frequently of a qualitative nature (Scott, Dinginto, & Xapile, 2015). It is important to recognize and value this type of input, while at the same time being aware that such information may be difficult to validate, in part because of its implicit nature. Managers need to identify points at which decisions are sensitive to critical information and, where necessary, triangulate the information. Data collection and analysis also carry substantial costs; thus, a good manager is able to request and use only the minimal amount of information at the minimal level of accuracy required for that decision.

Information may be of many types and describe a wide range of issues of importance to a manager. Indeed, several types of information systems were described in the preceding sections, including those dealing with budgetary information, information on transport usage, and personnel information. Within these health information systems, three domains of information can be distinguished: information on determinants of health (e.g., socioeconomic factors), information on a population's health status (e.g., mortality, morbidity, well-being), and information related to the functioning of a health system (inputs, processes, and outputs—for example, health services) (Health Metrics Network [HMN], 2008).

Sources of Information

The sources of health-related information can be classified into six broad categories (HMN, 2008; Stanfield, Walsh, Prata, & Evans, 2006):

- National census
- Population surveys
- Public health surveillance
- Vital events monitoring (civil registration)
- Health service records (statistics)
- Resource tracking subsystem

The following four categories are particularly important from the management perspective:

- Routine ongoing data collection (such as immunization records or resource tracking)
- Specific periodic or one-off surveys (such as community health surveys)
- Documented experiences and tacit knowledge of managers
- Comparative information from other health systems

Each source has both strengths and weaknesses in terms of accuracy, costs, and relevance to decision processes. A weakness of some management information systems (MISs) is that too much emphasis is placed on routine data collection systems when less frequent data collection would suffice. There is a temptation to institutionalize all data requirements, which can lead to an unwieldy formal MIS and perhaps a devaluing of the overall system.

Users and Providers of Information

Information systems include a number of components, with the information process being the central element. It comprises the following stages (HMN, 2008; Lippeveld, Sauerborn, & Bodart, 2000):

1. Identification of information needs
2. Collection of relevant data
3. Analysis of data to provide information
4. Use of information
5. Timely feedback to providers of information

Other important elements of information systems include the organizational and systems procedures and rules (e.g., the frequency of reporting) and the availability of resources (e.g., human resources, software and hardware, consumables).

Weaknesses can occur in an MIS at any stage of the information process. In some cases, data collected may not be relevant. For example, a manager seeking to understand why there is low utilization of health care is unlikely to get answers from service-based information, which cannot provide data on access barriers that prospective users may be facing in communities and which may underestimate the prevalence of unreported conditions. One frequent failing in information systems is a lack of feedback to data providers, which can lead to a downward information spiral in which the accuracy of data collected declines due to poor motivation and, as a result, the data are used even less.

Sometimes, inappropriate amounts and types of information may be found at different levels in the management system. **FIGURE 15-14** illustrates a common failing in which a similar amount of data flows up all levels of the system; ideally, the higher management levels should operate on increasingly selective amounts of key data.

Analysis of data and its transformation into information also raise issues in terms of the level at

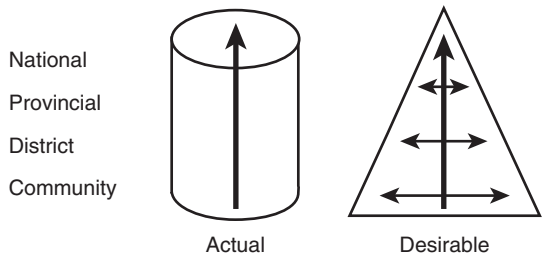


FIGURE 15-14 Data flow.

Reproduced with permission from Wilson, R. G., Echols, B. F., Bryant, J. H., & Abrantes, A. (Eds.). *Management information systems and microcomputers in primary health care.* Geneva, Switzerland: Aga Khan Foundation.

which these tasks are best done and by whom. A general rule of thumb is the closer the person doing the analysis/transformation to the collection of the data, the better.

Typically, HMIS represents the mainstream information system, covering all services. However, individual programs and projects (e.g., maternal–child health or mental health) may establish their own information systems, which may contribute to fragmentation of the health information system.

Information Technology

The rapid development of information technology provides both opportunities and challenges for health systems. Information technology can significantly increase efficiency of data collection, processing, analysis, and presentation—all of which may potentially contribute to improved quality and utilization of information in management and planning decisions. Information technology has developed rapidly, and organizations must be ready to adapt to these changes. For example, more than a decade ago Ash et al. (2004) reported that in the United States, the Netherlands, and Australia, the implementation of patient care information systems (PCIS) had the potential to foster errors related to the entry and retrieval of information, and to hinder the communication and coordination process that the PCIS aims to support. Nowadays, an increasing number of studies report the opposite effects. The positive effects that followed the introduction of the "SMS for life" plan in Tanzania offer a case in point. Through weekly mobile text messages and electronic mapping technology, this program helped improve accuracy of stock counts and timely replenishments of antimalarial medicines. The plan, which was implemented through a collaborative partnership of public and private institutions, contributed to greater visibility of antimalarial stock levels and supported more efficient stock management (Barrington, Wereko-Brobby, Ward, Mwafongo, & Kungulwe, 2010).

Furthermore, mobile technology provides opportunities to improve the range and quality of health services due to its increasing accessibility at the grassroots level. For example, a recent systematic review identified that, in addition to collecting health data, community health workers have increasingly used mobile technology to receive alerts and reminders, facilitate health education, and conduct person-to-person communication across various programs, including those dealing with maternal and child health, HIV/AIDS, and sexual and reproductive health (Braun, Catalani, Wimbush, & Israelski, 2013). This is consistent with a more general point, reported in an earlier systematic review, that health information technology can facilitate improvements in quality (through increased adherence to guidelines, decreased medication errors, and enhanced disease surveillance) and efficiency benefits (decreased utilization of care) (Chaudhry et al., 2006). The *Innovation, Technology, and Design* chapter provides more information on health and technology.

Key Themes in Management and Planning

The previous sections reviewed two key management and planning activities: organizing and resources management. Permeating these activities are various cross-cutting issues that are brought together and directly addressed in this section. How these cross-cutting issues are approached will have an impact on the way management functions are carried out within the health system. This section focuses on the following issues, which in many respects apply to both health management and health planning:

- Styles of leadership
- Accountability
- Evidence-informed approach
- Systems approach
- Linking and working together
- Sustainability
- Contradictions, tensions, and change

Styles of Leadership

Many of the issues related to leadership have been mentioned earlier in connection with supervision. Leadership is understood as a process of "providing direction to, and gaining commitment from, partners and staff, facilitating change and achieving better health services through efficient, creative and responsible deployment of people and other resources" (WHO, 2007b, p. 1). Effective leadership is concerned

with setting and maintaining an appropriate organizational culture and values that all (or at least a majority) of the organization's members will accept and subscribe to and that will help the organization achieve its organizational vision and mission. More than 400 definitions of leadership exist in the literature (Aarons, 2006; Edmonstone & Western, 2002; King & Cunningham, 1995; WHO, 2007b), and numerous frameworks have been developed for understanding leadership styles and approaches. Examples include a four-way typology of leadership styles denoting the distinction between high and low degrees of political and business orientations (Goodwin, 2000) as well as the various underlying theories of leadership (King & Cunningham, 1995). One framework of particular relevance to management is based on the distinction between the two broad leadership styles: *transformational leadership*, which inspires and motivates followers, and *transactional leadership*, which tends to rely on reinforcement and exchanges (Aarons, 2006; Stordeur, Vandenberghe, & D'hoore, 2000).

EXHIBIT 15-7 presents a number of functions of a leader. How managers practice effective leadership while carrying out these functions will depend on various factors. Personal qualities are certainly important—for example, one would expect a good leader to show personal understanding, confidence, and principled conduct. Nevertheless, these personal qualities are difficult to define precisely and measure. The style of leadership is also important. These styles can be viewed as varying, for example, from authoritarian to democratic, and from technically based to more personal-based leadership. However, leadership style needs to be adapted to the circumstances in which it is practiced, including the nature of the task. For example, problems with open-ended solutions may require a more flexible leadership style than more closed or finite problems; problems requiring an immediate solution may lead to a tighter style of leadership.

The category of staff over which leadership is exercised is also important. As noted earlier in this chapter, Hersey, Blanchard, and LaMonica (1978) highlighted task and relationship behavior as key variables of leadership activities. They highlighted that the maturity of staff is influenced by different combinations of four variables: capacity to set ambitious but realistic goals, willingness to take responsibility, and staff education and experience as both individuals or as a group. Managers adapt their style of supervision according to the level of maturity of their staff. For example, staff with a low level of maturity need the manager to demonstrate a high level of task leadership. As the maturity of the staff increases, the level of task leadership can decrease and the level of relationship behavior can increase.

The various styles of leadership in management are outlined in the following subsections. Note that these styles are not mutually exclusive, but rather can coexist within a single organization.

EXHIBIT 15-7 Functions of a Leader

- Scanning for changes and new developments
- Structuring and focusing the situation: determining what is a key priority, making it clear where the group is going and which action has to be taken
- Controlling group behavior: creating and enforcing appropriate rules for guiding the behavior of group members
- Speaking for the group: sensing and articulating (both internally and externally) the objectives and feelings of the group
- Uniting the group to achieve its goals and potential: mobilizing and coordinating group resources and decision making toward a common goal
- Inspiring: being a role model who serves and encourages

Coleman (1969), as quoted in Smith et al., 1982.

Proactive

A proactive approach reflects a key difference between management and administration and lies at the heart of the pursuit of objectives. When this style is applied, management performance should not be measured based on adherence to rules and regulations, as in administration, but rather on the extent of movement toward objectives. Such movement requires a proactive approach and reflects the ability to predict constraints and new opportunities, influence the environment, develop coalitions of support, foresee future problems and take early preventive action, and monitor performance so as to take corrective action.

There is a potential danger that managers' performance may be (perceived to be) judged on the mere appearance of proactivity. This circumstance can lead to shows of being proactive, such as displays of haste, workaholic behavior, and constant change. In some contexts, the necessity to demonstrate a proactive approach may lead to "change for change's sake," with little consideration of the underlying rationale. This is not good management.

Risk Taking

Risk taking reflects another important difference between administration and management. It expresses the inventiveness of managers, meaning their use of an

approach that does not accept existing ways of doing things are the only way. New methods and approaches involve an element of risk, however, so they represent a source of danger for managers. There are necessary limits to the extent to which managers should take risks; this point is particularly evident in the field of health. The notion of reasonable and acceptable risk needs to be understood.

How staff performance is managed also reflects the organization's attitude toward risk taking. For example, if staff are routinely criticized for failure and not praised or rewarded for success, then few staff will be willing to take reasonable and acceptable risks. Effective and supportive communication and supervision, together with the existence of clear and agreed-upon ethical and technical work standards, can create a more conducive environment of acceptable risk taking.

Problem Solving

One function of management is to seek solutions to problems arising from unforeseen changes in circumstances either within the organization or externally. An attribute of good management is the ability to deal creatively with new issues. One danger of paying too much attention to "problems," however, is neglect of future emerging issues at the expense of current firefighting. Problem-solving management needs to be balanced with a more long-term set of considerations.

Accountability

Accountability is defined as "the obligation of individuals or agencies to provide information about, and/or justification for, their actions to other actors, along with the imposition of sanctions for failure to comply and/or to engage in appropriate action" (Brinkerhoff, 2004, p. 372). It can be interpreted as a health system value as well as a principle of good governance (Collins & Green, 2014; Siddiqi et al., 2009; Travis, Egger, Davies, & Mechbal, 2002), and its importance permeates the health system. Accountability is essential to performance management, underlines the seriousness of health and health care and the concern for people's lives, and is crucial to exercising checks and balances on the use of resources. Networks of different (but overlapping) and changing forms of accountability run through the management of the health system, as summarized in **TABLE 15-9**.

Managers need to recognize and respond to each form of accountability. In some cases, they are

TABLE 15-9 Different and Changing Forms of Accountability

Form of Accountability	Commentary
Managerial	Managerial authority sees subordinates as accountable to the hierarchically superior manager. This perspective is typically expressed in the main line managerial relations in the organizational chart.
Professional	Professions are hierarchically structured, so that junior staff are technically accountable to those higher up on the professional ladder. Members of the profession may also be accountable to a professional body for their actions and behavior.
Political	Political representatives can hold staff accountable for their actions and behavior—for example, under devolution, where health staff are accountable to an elected local council.
Community	Health staff may be formally accountable to the community through mechanisms that monitor areas such as availability of medicines, personal relations with health staff, and punctuality of staff. Community accountability mechanisms in LMICs include health center and clinic committees, village health committees, patients' rights charters, and citizen report cards (McCoy, Hall, & Ridge, 2012; Molyneux, Atela, Angwenyi, & Goodman, 2012).
User	User accountability differs from the community form in that health staff are accountable to users of a particular service.
Market	Users are viewed as consumers of a particular service, for which they pay. Health staff are accountable to the consumers of the service through the market. Consumers who are dissatisfied with the service may go elsewhere.

reporting to authorities; in others, they are themselves the authority. Each form of accountability imposes rights and obligations on the management process. At the same time, individuals involved in management take part in the design of accountability relations and need to balance the various aspects of the network of complex, changing, and overlapping relations. In so doing, they are influenced by the values of the health system and by demands for new forms of accountability. The PHC approach, for example, puts the onus on community-based accountability, whereas the contemporary reform of health systems toward devolution emphasizes politically devolved forms of accountability.

Managers need to strike a balance between bureaucratic accountability (at different levels of the health system) and external accountability (to the community) to promote responsiveness to patients (Cleary, Molyneux, & Gilson, 2013). **EXHIBIT 15-8** identifies the key determinants of accountability of health services providers to the public. Note that this exhibit uses the term "user" rather than "consumer" to reflect that these determinants of accountability can be considered in both the public service and the private market context.

Evidence-Informed Approach

As stressed earlier in this chapter, management differs from administration in that it requires individual initiative rather than reliance on rules and regulations. The success of a manager in exercising such initiative is based on a combination of factors, including his or her technical skills and judgment. Both of these also rely on the manager's ability to draw on existing evidence (from HMIS, research, and other sources) when making managerial decisions. Different models have been proposed for utilization of evidence in decision-making processes, including knowledge-driven, problem-solving, interactive, political, tactical, and enlightenment models (Bowen & Zwi, 2005; Hanney, Gonzalez-Block, Buxton, & Kogan, 2003). Most health systems apply a combination of those models.

For the use of evidence to occur, various preconditions need to be met. First, the organizational culture needs to seek, generate, and accept the use of evidence in its decision making. Earlier we mentioned the importance of the values that underpin management and planning decisions; evidence-informed decision making itself can be seen as both a value and a process emanating from other values such as equity. A major criticism of many organizations relates to their failure to use evidence. This omission may result from either a genuine lack of evidence or a resistance to research-generated evidence that is not perceived as relevant—a phenomenon called the "know–do" gap (Campbell et al., 2009; Hanney et al., 2003; Mirzoev et al., 2017). In addition, some of this resistance may be generated by stakeholders with a vested interest in maintaining a particular position that evidence might

EXHIBIT 15-8 Key Determinants of Accountability of Service Providers

While there are many types of accountability, accountability by providers to communities and users is gaining significant traction in international discourse. Managers are stewards of the health system and must ensure that health providers are accountable within the health system and are responsive to community needs. Drawing on Berlan and Shiffman (2012), factors that may shape health provider accountability to users can include the following:

Health System Factors

- Appropriate oversight mechanisms are needed: The actors who supervise health providers may influence whose preferences and needs are prioritized.
- The source of revenue within the healthcare financing system can influence whose needs are prioritized (the rich versus the poor; the funder versus the community).
- The nature of competition between providers may lead to greater focus on user preferences to obtain patients, although this needs to be considered within a broader context of the public–private mix of providers in a country.

Social Factors

- Users often lack information and power to influence provider behavior. For this reason, a focus on empowering users through information provision and access to a broader collective can assist them in holding providers accountable.
- Professional and societal norms can influence the level of accountability that providers feel toward users. Such norms can include provider gender bias, class bias, and professional norms.

Data from Berlan, D., & Shiffman, J. (2012). Holding health providers in developing countries accountable to consumers: A synthesis of relevant scholarship. *Health Policy and Planning. 27*, 271–280.

challenge. Managers have a responsibility to generate an organizational culture that respects evidence-based approaches.

Second, a degree of "user-friendliness" of evidence is needed that is appropriate to the decisions being made (Hanney et al., 2003; Mirzoev et al., 2017; Oxman, Lavis, Fretheim, & Lewin, 2009). Facility managers and even district health planners are unlikely to find the time to read lengthy reports of monitoring and evaluation data sets. Instead, such individuals are more likely to read, understand, and eventually use succinct summaries in their planning and management decisions—that is, brief synopses that are written in nontechnical language and that clearly highlight key facts or issues that need addressing or consideration in the plans or management decisions. Although they certainly offer benefits in terms of their brevity, the level of accuracy may be compromised in such summaries. As noted earlier, there are clearly costs involved in seeking evidence, and these costs are directly related to the level of validity and accuracy of generated evidence as well as "packaging of the evidence" in an easy-to-use format to inform planning and management decisions. In turn, a decision maker needs to request only the minimum set of evidence necessary for any particular decision without compromising the quality of the overall evidence.

In the end, we return to the theme of the manager and the planner as practitioners of a value-driven art rather than a purely technical science. An indefinable quality of a good manager, which in part results from experience, is the ability to sift through information, select the appropriate information, and act appropriately on it. While management information systems, as a source of evidence, are an important source of inputs into the decision-making process for managers, Collins and Green (2014) note that decisions and activities decided upon by managers are, in fact, premised on an intersection of evidence, values, and judgment. Managers must combine the evidence (scientific and nonscientific) with which they are presented with their personal and health system values that they seek to uphold and achieve in the future. Management judgment is then applied to weigh tradeoffs and potential conflicts in decisions and activities. Policy, planning, and management decisions often require the agreement of many stakeholders, all with their own personal and collective value sets. Managers and planners should be aware that these stakeholders may or may not share their own values base, and should be aware of their own value sets, which shape their practices and management decisions.

Systems Approach

The advantages of the division of labor often prompt organizations and health staff to parse broader functions into specialized tasks. Although understandable, this approach runs the danger of producing fragmentation, narrow perspectives, and inflexible working patterns. In the management process, staff need to complement their own specialized responsibilities with a broader, more flexible approach—that is, with a *systems approach*.

Embrace of the systems approach means that staff, in addition to handling their own specialized tasks in management (such as human resources or logistics management), need to acquire a broad range of knowledge, skills, and aptitudes. This wide scope is important for the management process because managers need to combine and use different resources. It is also important in that it allows for more flexible working patterns, allowing staff to be more multifunctional and deal more effectively with short-term changes in the organization's work.

In addition, adoption of the systems approach means that staff are clear in relating their own positions in the organization to the overall organizational objectives. Thus, the manager of the TB program is not just concerned with a reduction in TB morbidity and mortality, but also with how overall community morbidity and mortality can be reduced. In this case, staff appreciate how their own actions relate to the work performed by other parts of the organization. Perhaps extra resources for one department or program are taking resources away from an area of greater priority. Or maybe having extra staff at a health center located in a high-income area may means staff are not assigned to health centers in poorer areas. Thus, the systems approach means appreciating the opportunity costs of managerial action.

The concept of systems thinking was recognized by the WHO Alliance for Health Policy and Systems Research in its flagship publication that identified the 10 steps in this process (De Savigny & Adam, 2009). The application of systems thinking in LMICs is deepening our knowledge of managing and leading in complex health systems (Adam, 2014; Gilson, Elloker, Olckers, & Lehmann, 2014; Kwamie, van Dijk, & Agyepong, 2014; Prashanth et al., 2014).

Linking and Working Together

The call for more collaboration is a frequent recommendation in management reports and guidance documents for health planners. Collaboration, however,

remains an elusive feature of management and planning, although attempts have been made to define its different manifestations, such as coordination and cooperation (e.g., Wang, Collins, Vergis, Gerein, & Macq, 2007). The forces behind individualism, exclusive group interests, and group interorganizational and intraorganizational conflicts are difficult to overcome. Yet the health system, if it is to function effectively and efficiently, needs to push collaborative strategies to the forefront of managerial action. Health staff, in their various organizations, need to work together. This includes the development of integrated management of patient care, team approaches within an organization, interdepartmental cooperation, public–private partnerships, and interorganizational and intersectoral linkages (Elloker, Olckers, Gilson, & Lehmann, 2012–2013; Langlios, Montekio, Young, Song, & Alcalde-Rabanal, 2016).

Working together through an intersectoral approach has long been advocated. Health depends on a wide range of factors located outside the healthcare system, including those known as social determinants of health (see the *Understanding and Acting on Social Determinants of Health and Health Equity* chapter). The challenge for management and planning is to devise effective approaches, actions, mechanisms, structures, and attitudes that can lead to greater ability of healthcare personnel to work together and coordinate their activities with their counterparts from other sectors such as education or housing. Ways of working across boundaries include the following:

- *Coordinating*. Two or more units agree on joint objectives, dedicate resources, and develop a joint plan, organization, and program (Rogers & Whetten, 1982).
- *Cooperating*. Two or more organizations keep their own separate but compatible objectives and agree to help—or at the least avoid hindering—each other (Rogers & Whetten, 1982). For example, health and agricultural organizations might agree to support the cultivation of subsistence crops and to protect fish-bearing rivers to ensure adequate nutrition of rural communities.
- *Community supporting*. Communities, which can view needs from a more integrated perspective than can the public sector, might support policies designed to strengthen community involvement in decision making and planning, so as to encourage more community demands and support for integrated action by the public sector.
- *Nesting*. Resources or activities are located in other organizations that support an intersectoral approach (Wang et al., 2007). For example, the MOH might strategically support initiatives of an intersectoral character in local governments, community associations, and NGOs by locating or nesting support in strategic points in these organizations. In return for action of an intersectoral nature, these bodies would receive resources and technical support.
- *Advocating*. The MOH often takes on the role of advocating for a health perspective to be used across the various sectors of the public sector—for example, through control over tobacco consumption, the use of car seat belts and motorcycle helmets, and elimination of fire hazards. Ideally a Health in All Policies approach should be adopted to minimize unintended impacts of public policy (WHO, 2014).
- *Regulating*. As discussed in more detail in the *Design of Health Systems* chapter, the MOH has authority to ensure compliance from other individuals and organizations. A particular challenge has arisen from the growth of the private sector in many LMICs (Lê et al., 2015; Sheikh et al., 2015). The MOH and other health-related organizations may be vested with legitimate regulatory authority to demand certain behaviors leading to health improvement. This role requires the development of legal expertise within the MOH, the focus of regulatory action on health objectives, the creation of regulatory mechanisms to require certain behavior, and the development of a sustainable administrative structure to implement regulations. In turn, political commitment is needed to back up regulatory legitimacy, while proactively initiating dialogue as well as taking intellectual leadership on the part of the MOH. Constraints on regulation might include a lack of state legitimacy, a shortage of financial and human resources to enforce the regulations, and regulatory capture by vested interests.
- *Shifting Authority Upward or Downward*. Authority for setting a cross-cutting strategic direction can be shifted upward in the hierarchy of government, moving above the organizational and sectoral divisions. A powerful cross-governmental commission for social development involving related ministries, for example, could provide intersectoral decision making as well as national planning and resource allocation. Authority can also be shifted downward to multifunctional devolved units at the local level. As discussed in more detail in the *Design of Health Systems* chapter, devolution, as a form of decentralization, opens up the potential for an intersectoral approach that links healthcare systems and other departments such as education and social services, although

the divisions between departments within local government can still constitute strong constraints.

- *Intersectoral collaboration*. This type of collaboration is defined as "the joint action taken by health and other government sectors, as well as representatives from private, voluntary, and nonprofit groups, to improve the health of populations" (Public Health Agency of Canada, 2016). In addition to requiring political commitment and support for the integrated system, successful intersectoral collaboration can be facilitated by engaging with partners early and establishing shared values and purpose. This includes deciding on concrete objectives and results as well as making clear the resource implications (personnel and money) for each party. It is also important to invest in the alliance-building process and to identify relevant champions who can support and drive the collaboration forward. Leadership, accountability, and rewards should be shared fairly among all parties (Public Health Agency of Canada, 2016).

Management Systems

Collaboration can be built into the management systems of organizations based on how they use resources. Agreements on consultation and use of information, for example, could be integrated into the system of district health planning and other sector planning. Information systems can also take on an intersectoral character (de Kadt, 1988).

Contracting

Contracts between organizations can be a useful means for developing collaboration. As shown in **FIGURE 15-15**, this relationship involves a commissioning body or purchaser, the contract, and the provider. The existence of a contract assumes some form of separation between the purchaser or commissioner (who has the funds and is the primary agent for determining the needs to be met by the contract) and the provider. The contract itself can include issues relating to services, object, and payment, and can be executed in a competitive or noncompetitive environment.

The elements involved in the contracting process suggest that variations in these components can lead to numerous contractual forms (**TABLE 15-10**). The different types of contracting may be appropriate for different environments.

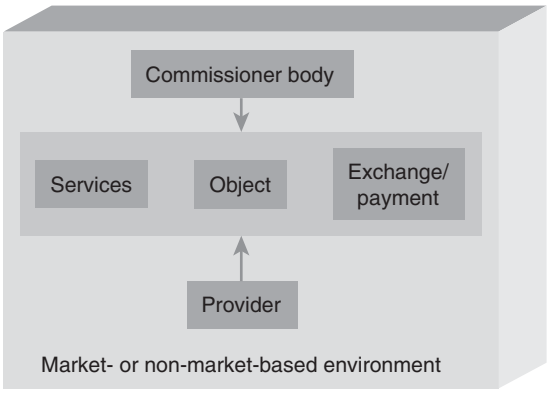


FIGURE 15-15 Elements in contracting.

TABLE 15-10 Variables in the Analysis of Contracting Primary Care Services

Variable	Comment
Object of the contract	Ranging from type of service to geographic coverage
Type of contracting	Ranging from formal and legalistic to more flexible contracting
Type of commissioner	Could include centralized or decentralized agencies with different degrees of organizational autonomy and financing (e.g., taxes, insurance systems, fee based)
Type of provider	Public-sector organization, NGO, or private for-profit organization, each with its own variations
Type of exchange or payment	Based on a definition of services and payment system, which could be, for example, performance based
Market or nonmarket based	Contracts can be based on competitive tendering, with services provided in a market environment, or without competitive tendering, with service providers operating in a noncompetitive environment

One can distinguish between a time-based versus a lump-sum nature of the payments in contracts. As an alternative means of classifying contracts, Palmer (2000, p. 823) distinguishes between classical contracts ("discrete transactions between people who will never see each other again"), neoclassical contracts ("less discrete and therefore contain techniques for flexibility within the terms of the contract, such as third-party determination of performance"), and relational contracts (the terms of the contract are not the key element; rather, the whole relationship between the parties over time is key, and importance is put on harmony and keeping the relationship). Palmer suggests that LMICs often rely on classical and neoclassical contracting, despite the possibility that relational contracting might well be more promising in that environment.

Sustainability

An important theme in management is that of the sustainability of services. Lafond (1995, p. 63) sees sustainability as "the capacity of the health sector to function effectively over time, with a minimum of external input." The mention of "external input" is of particular relevance to poor countries that are heavily reliant on funds from international agencies. Olsen's (1998) definition emphasizes the longer term: "A health service is sustainable when operated by an organisational system with the long-term ability to mobilize and allocate sufficient and appropriate resources (manpower, technology, information, and finance) for activities that meet individual or public health needs/demands" (p. 289).

The relevance of sustainability to health management and planning rests on two key points. First, health care is rarely time limited. Disease control, for example, requires continued interventions over time. As one case in point, the continuity of action in TB control is important to limit the spread of disease. In particular, continued medicine supply and continued treatment are important to avoid the development of drug-resistant TB.

Second, the importance of sustainability draws on four problematic and interrelated features of the public sector, particularly in LMICs:

- The extremely limited resource base of poor societies, coupled with the limited authority of national governments, leads to intense resource scarcity in the public sector. This limitation, in turn, underlies the problem of maintaining a secure and constant flow of resources to finance government programs and produces high levels of dependence on international aid.
- The historical focus of international aid to LMICs on capital investment, with less support being given to recurrent expenditure, leads to difficult problems of maintaining programs of action. This issue is compounded by the vertical and cyclical nature of foreign aid projects.
- Corruption has a marked impact on the continuity and effectiveness of government given the high level of resource scarcity.
- High staff turnover may lead to limited effectiveness and reduced sustainability of government programs (Collins et al., 2000).

In practical terms, managers need to develop the appropriate health systems capacity, and in particular the organizational culture based on the chosen values, vision, and mission, and then ensure its sustainability. By "ensuring sustainability," health planners usually mean drawing realistic plans and ensuring consistency between longer-term policies or strategic plans and shorter-term operational plans and management practices (Mirzoev & Green, 2017).

Sustainability is not actually an end in itself, but rather a means to a wider end. It is colored by the policy objectives held by the stakeholders. This influence introduces two possibilities:

- *Appropriate sustainability* occurs when a particular activity has continuity over time and has outputs that maintain their value. For example, policy objectives often specify the achievement of improved and equitable health care, with certain types of action deemed essential to achieve this outcome.
- *Inappropriate sustainability* occurs when programs and activities need to be stopped at a particular point, but instead continue. For example, as the incidence of a particular health problem declines, related interventions may need to be reduced. Smallpox vaccination is a classic example.

Of course, what is seen as appropriate sustainability by one stakeholder may be viewed as inappropriate sustainability by another. Stakeholders in a particular government activity will have different interpretations of a program, largely determined by how it affects their own particular interests. **TABLE 15-11** highlights the possible interests of different stakeholders in sustainability.

To capture the determining features of sustainability, the analysis of the health system needs to recognize the broad-ranging character of sustainability and its multiple determinants. The factors influencing sustainability will also vary in different circumstances.

TABLE 15-11 Stakeholders' Potential Interests in Sustainability

Stakeholder	Potential Interests in Sustainability
Donors	Shifting resource dependence away from donors and increasing local contributions
Ministry of health	Ensuring regularity of donor resources and/or shifting resource dependence away from donors and increasing local contributions
Staff	Ensuring continuity in employment and career progression
Users	Ensuring continuity of improved service delivery
Nonusers	Expansion of services on a regular basis to cover their needs
Private-sector contractors	Shift to public financing and private provision under favorable conditions

Conclusion and Challenges for Managers and Planners

Management and planning are fundamentally concerned with the relationship between resources and objectives. They both deal with scarce resources, operate within various technical and political constraints, are strongly influenced by values, can operate in a political manner, and have a strong interrelationship with the changing social, political, economic, and international context. Two interrelated functions of management and planning have been emphasized in this chapter: Organizing and management of resources. Cross-cutting issues permeating these management activities include styles of leadership; accountability; an evidence-informed approach; a systems approach; linking and working together; and sustainability.

The future direction of health management and planning in LMICs is far from clear. Certainly, the goal of universal health coverage will require significant improvements in health management and planning. However, such improvements cannot be divorced from some challenging questions.

What Is the Future of Public-Sector Management and Planning?

Criticisms of state and public-sector management and planning have ranged from simple market orthodoxy to a more nuanced approach that recognizes their needs and social responsibilities. All too often, the prevailing view of corruption, bureaucratic expansion, administrative fragmentation, and patronage within health systems has made it easy for many to levy criticisms. Nevertheless, two points need to be recognized.

First, public-sector management and planning typically operate in a context of low funding, intense and changing needs, and an unstable environment. Of course, the degree to which the resources are constrained varies across countries. For example, the rapid growth of many Asian economies provides—at least in theory—a platform for increased investments in these countries' public health sectors. Changing health needs (e.g., due to epidemiologic and demographic transitions) and unstable political and economic environments and crises put additional strains on management of the public health sectors in many countries.

Second, there is a need to recognize the specific characteristics of public-sector management and planning. Public-sector and private-sector management have much to learn from each other. The future development of public sector management, and the values that underpin it, need to take into account three considerations:[10]

- The public sector seeks to meet public needs and interests, although different social and political groups have different interpretations of these needs and interests. Resolution of these discrepancies requires open and transparent

[10] These three points draw on the work of Stewart and Ranson (1994).

participation and negotiation regarding the definition of public interest and agreement on the underlying values. Public-sector bodies need to define their objectives in relation to these public needs and interests.

- The multifaceted nature of public needs means that public-sector organizations can achieve their objectives only by engaging in collaborative strategies with other bodies, whether public-sector organizations, NGOs, or private for-profit organizations, and across all health-related sectors, such as education or housing. This also requires collaborative strategies between the public sector and communities.
- Social needs are not ephemeral in nature but rather are long-term issues, and they require public-sector management to recognize and respond to their persistence. Planning, sustainability, continuity of service provision, capacity building, systems development, and planning for professional development are among the key features of public management.

Are Managers Being Asked to Manage the Unmanageable?

That managers in health systems have to deal with contradictions, tensions, and change is not new. However, the immensity of these challenges becomes evident as we consider the paucity of resources available even in middle-income countries and the scale of the health needs stressed elsewhere in this text. In dealing with these challenges, there is a gap, which varies from one country to another in terms of both its size and its nature, between the existing capacity of management to meet those challenges and the potential that exists to do so through management strengthening. It is a matter of concern, however, that an already difficult situation is being made worse through sharpening contradictions and weakened means to respond to populations' needs.

First, the tension between public interests and private gain is increasing as a result of the growth of public–private partnerships and the trend toward self-funding public bodies. Second, the public sector faces threats because of the poor rewards for public servants, the downgrading of the role of the public sector, and a hollowing out of its rationale and process. The culture of individualism, internal markets, and self-funding of health units and programs all stand in opposition to the public sector's perspective of social interests, collaborative strategy, longer-term framework, and public ethos. In the words of Flynn (1997, p. 232):

> What . . . [was] . . . needed is a change in attitude towards the public sector. If spending could be based on need and a realistic assessment of what is affordable rather than a constant state of crisis and if management arrangements could be based less on distrust and fear and more on cooperation, then public services could make a valuable contribution to the economic health and quality of life of society.

Twenty years on, although the narrative of the New Public Management persists, it has become decidedly more nuanced. The initial illusion of possible vast increases in the efficiency of the public sector (through adopting private-sector principles) has since been overshadowed by the recognition of the continuous "firefighting" mode in which the public sector operates, which requires continuous adaptation to multiple crises. The nature of these crises is fundamentally different from the pressures that the private (particularly for-profit) sector faces. For example, while populations in many countries have become more affluent, health inequalities have widened dramatically in rapidly developing countries such as China. Large population displacements across the globe, for various reasons, have put additional strains on the public health systems of the recipient countries. Even as the likelihood of profit generation has increased, the quest to address the social goals of equity and universal health coverage has become more of a challenge. The concept of health systems resilience has emerged in the last several years, signaling the need to prepare health systems (and particularly the public sector) and their managers for both everyday shocks and major crises (Gilson et al., 2017; Kruk et al., 2017).

Perhaps as a consequence of the ever-changing contextual pressures, although the fundamental principle of focusing on social value remains the dominant theme within the public health sector, the role and resultant shape of that sector has changed. The introduction of internal public-sector markets has brought principles of competition to the public health facilities in countries such as Vietnam (Lê et al., 2015) and China, raising questions regarding the balance between the social goals of the public sector and the pressures to remain financially sustainable. In other rapidly developing countries (e.g., Thailand and Brazil), the public sector's role and primary focus on

social goals have been maintained, illustrating the public sector's potential resilience to pressures of marketization and calls for increased efficiency. Such calls can, for example, counterbalance the pursuit of health equity and universal health coverage.

What Are the Challenges Facing Public-Sector Health Managers?

This chapter has provided a brief overview of the key issues that a manager or a planner working in the public health sector needs to address. Management and planning, we have argued, is a key, but often undervalued, component of a health system. As managers and planners seek to establish the structures, resources, and processes necessary for meeting the SDG, which emerging challenges are they likely to face?

First, public health planners and managers will need to grapple with the increasingly diverse structures of the health system and new approaches to funding. The ever-more-intertwined public and private institutional complexity and changing relations between the system center and the local levels, combined with new funding flows and public–private partnerships, suggest that decision makers need to seek new ways to meet these challenges.

Second, although LMICs continue to face major financial resource constraints, the crisis of the health worker gap must be seen, for many countries, as an even greater constraint. This trend, coupled with the low staff morale in many health systems, presents major challenges for planners and managers, who will need to find new ways to cope with shortages of healthcare workers.

Third, managers face an increasingly vocal and empowered citizenry, conscious of the global rights agenda. This trend will increasingly—and rightly so—challenge established values and principles (such as command and control) and will ultimately place greater demands for accountability on the health system. Planners and managers will need to think strategically in meeting this demand.

Fourth, new technology is continually emerging and increasingly being used in various aspects of health systems, including provision of health services, planning, and management. Its utility is often counterbalanced by its high costs, with managers being called upon to respond to and manage the potential of, and access to, the technology.

Fifth, the increasing number of fragile economies in the countries across the world, combined with heightened interest in building resilient health systems that can anticipate and withstand everyday challenges as well as major shocks, create specific challenges for planners and managers. These challenges are manifold: Examples include responding to an ever-changing context including the resource environment, ensuring consistency between short- and longer-term perspectives, and ensuring sustainability.

The nature, and indeed excitement, of management and planning lies in identifying and adequately responding to such emerging and ever-changing challenges. Management as a scientific art must continuously develop new approaches and tools to accomplish this feat.

Discussion Questions

1. Which approaches to health management exist in a country known to you? Which factors affect any differences in approaches to health management at different levels of the country's health system, and why?
2. Consider which kind of health planning takes place in a country known to you. How successful is it? Which major health needs exist that are not met? Can you identify an expenditure that is of lower priority than these unmet identified needs? How could health planning be improved?
3. How are healthcare priorities decided within the health sector of a country known to you? What is the role of the manager in this process? Which values do these priorities and plans reflect?
4. Describe the financial management system of a country known to you. How are resources allocated to lower levels? Are there any differences in allocation between different types of resources? How is resource allocation linked to the planning process?
5. Which system of decentralization exists in your country? What are its strengths and weaknesses for the health sector?
6. Which practical steps could be taken to improve intersectoral collaboration for health in your country?
7. Consider the way in which local-level health staff are managed. Which steps could be taken to improve the way in which these staff are managed?
8. What information is included in the health information system(s) of a country known to you? How effective is it in providing timely, accurate, and relevant information to inform health management and planning decisions?

References

Aarons, G. A. (2006). Transformational and transactional leadership: Association with attitudes toward evidence-based practice. *Psychiatric Services*, *57*(8), 1162–1169.

Adam, T. (2014). Advancing the application of systems thinking in health. *Health Research Policy and Systems, 12,* 50.

Alimo-Metcalfe, B. (1994, October 20). The poverty of PRP. *Health Services Journal*, 22–23.

Anderson, S., Huss, R., Summers, R., & Wiedenmayer, K. (2004). *Managing pharmaceuticals in international health*. Basel, Switzerland: Birkhaeuser Verlag.

Arnesen, T., & Kapiriri, L. (2004) Can the value choices in DALYs influence global priority-setting? *Health Policy*, *70*(2), 137–149.

Arredondo, A., Orozco, E., & De Icaza, E. (2005). Evidence on weaknesses and strengths from health financing after decentralization: Lessons from Latin American countries. *International Journal of Health Planning and Management*, *20*(2), 181–204.

Ash, J. S., Berg, M., & Coiera, E. (2004). Some unintended consequences of information technology in health care: The nature of patient care information system-related errors. *Journal of the American Medical Informatics Association, 11*(2), 104–112. doi:10.1197/jamia.M1471

Atkinson, S., Medeiros, R. L., Oliveira, P. H., & de Almeida, R. D. (2000). Going down to the local: Incorporating social organisation and political culture into assessments of decentralised health care. *Social Science and Medicine*, *51*, 619–636.

Bacal, R. (1999). Performance management. New York: McGraw-Hill.

Barasa, E. W., Molyneux, S., English, M., & Cleary, S. (2015). Setting healthcare priorities at the macro and meso levels: A framework for evaluation. *International Journal of Health Policy and Management, 4*(11), 719–732.

Barrington, J., Wereko-Brobby, O., Ward, P., Mwafongo, W., & Kungulwe, S. (2010). SMS for Life: A pilot project to improve anti-malarial drug supply management in rural Tanzania using standard technology. *Malaria Journal, 9*(1), 1–9

Bednarz, D. (2010). Impacts of the economic crisis on public health, part II: Paradigms and the right questions. *Energy Bulletin*. Retrieved from http://www.energybulletin.net/node/52172

Berlan, D., & Shiffman, J. (2012). Holding health providers in developing countries accountable to consumers: A synthesis of relevant scholarship. *Health Policy and Planning, 27,* 271–280.

Bosch-Capblanch, X., Liaqat, S., & Garner, P. (2011) Managerial supervision to improve primary health care in low- and middle-income countries. *Cochrane Database of Systematic Reviews.* doi:10.1002/14651858.CD006413.pub2

Bowen, S., & Zwi, A. B. (2005) Pathways to "evidence-informed" policy and practice: A framework for action. *PLoS Medicine*, *2*(7), e166.

Braun, R., Catalani, C., Wimbush, J., & Israelski, D. (2013). Community health workers and mobile technology: A systematic review of the literature. *PLoS One, 8*(6), e65772.

Brinkerhoff, D. (2004). Accountability and health systems: Toward conceptual clarity and policy relevance. *Health Policy and Planning, 16*(6), 371-379.

Brugha, R., & Varvasovszky, Z. (2000). Stakeholder analysis: A review. *Health Policy and Planning*, *15*(3), 239–246.

Buchan, J. (2004). What difference does ("good") HRM make? *Human Resources for Health*, *2*, 6.

Buse, K., & Harmer, A. M. (2007). Seven habits of highly effective global public–private health partnerships: Practice and potential. *Social Science & Medicine*, *64*(2), 259–271.

Campbell, D., Redman, S., Jorm, L., Cooke, M., Zwi, A., & Rychetnik, L. (2009). Increasing the use of evidence in health policy: Practice and views of policy makers and researchers. *Australia and New Zealand Health Policy*, *6*(1), 21.

Chaudhry, B., Wang, J., Wu, S., Maglione, M., Mojica, W., Roth, E., . . . Shekelle, P. G. (2006). Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care. *Annals of Internal Medicine, 144*(10), 742–752.

Cleary, S. M., Molyneux, S., & Gilson, L. (2013) Resources, attitudes and culture: An understanding of accountability mechanisms in primary health care settings. *BMC Health Services Research*, 13.

Cole, G. A. (1988). *Personnel management: Theory and practice*. London, UK: DP Publications.

Collins, C. (1994). *Management and organisation of developing health systems*. Oxford, UK: Oxford University Press.

Collins, C., & Green, A. (2014). *Valuing health systems: A framework for low and middle income countries.* New Delhi, India/Los Angeles, CA: Sage.

Collins, C., Omar, M., & Hurst, K. (2000). Staff transfer and management in the government health sector in Balochistan, Pakistan: Problems and context. *Public Administration and Development*, *20*, 207–220.

Collins, C. D., Myers, G., & Nicholson, N. (1992). A successful transport scenario for the health sector in developing countries. *World Hospitals, 28*(3), 9–14.

Couttolenc, Bernard F. (2012). *Decentralization and governance in the Ghana health sector.* Washington, DC: World Bank. doi:10.1596/978-0-8213-9589-9

de Kadt, E. (1988). Making health policy management intersectoral: Issues of information analysis and use in less developed countries. *Social Science and Medicine*, *29*(4), 503–514.

Department of Health and Developmental Social Welfare (DHDSW). (1999). *Transport management manual*. South Africa: DHDSW, Save the Children, & Transaid.

De Savigny, D., & Adam, T. (Eds.). (2009). *Systems thinking for health systems strengthening*. Geneva, Switzerland: Alliance for Health Policy and Systems Research, World Health Organization.

Dieleman, M., Cuong, P. V., Anh, L. V., & Martineau, T. (2003). Identifying factors for job motivation of rural health workers in North Viet Nam. *Human Resources for Health*, *1*, *10*.

Dobrow, M. J., Goel, V., & Upshur, R. E. G. (2004). Evidence-based health policy: Context and utilisation. *Social Science and Medicine, 58*(1), 207–217.

Doherty, J., & Gilson, L. (2015, August 17–19). Workplace-based learning for health system leaders: Practical strategies for training institutions and governments. RESYST. Retrieved from http://resyst.lshtm.ac.uk/resources/workplace-based-learning-report

Edmonstone, J., & Western, J. (2002). Leadership development in health care: What do we know? *Journal of Management in Medicine, 16*(1), 34–47.

Elloker, S., Olckers, P., Gilson, L., & Lehmann, U. (2012–2013). Crises, routines and innovations: The complexities and possibilities of sub-district management. *South African Health Review,* 161–173.

Flynn, N. (1997). *Public sector management* (3rd ed.). London, UK: Harvester Wheatsheaf.

French, D., & Saward, H. (1983). *A dictionary of management*. London, UK: Pan.

Frost, L. J., & Reich, M. R. (2008). *Access: How do good health technologies get to poor people in poor countries?* Cambridge, MA: Harvard University Press.

Fryatt, R., Mills, A., & Nordstrom, A. (2010). Financing of health systems to achieve the health Millennium Development Goals in low-income countries. *The Lancet*, *375*(9712), 419–426.

Gaitonde, R., Oxman, A. D., Okebukola, P. O., & Rada, G. (2016). Interventions to reduce corruption in the health sector. *Cochrane Database of Systematic Reviews, 8,* CD008856. doi:10.1002/14651858.CD008856.pub2

Gilson, L., Barasa, E., Nxumalo, N., Cleary, S., Goudge, J., Molyneux, S., Tsofa, B., & Lehmann, U. (2017). Everyday resilience in district health systems: Emerging insights from the front lines in Kenya and South Africa. *BMJ Global Health, 2*(2), 1-15.

Gilson, L., Elloker, S., Olckers, P., & Lehmann, U. (2014). Advancing the application of systems thinking in health: South African examples of a leadership of sensemaking for primary healthcare. *Health Research Policy and Systems, 12,* 30.

Goodwin, N. (2000). Leadership and the UK health service. *Health Policy*, *51*(1), 49–60.

Green, A. (2007). *An introduction to health planning in developing health systems* (3rd ed.). Oxford, UK: Oxford University Press.

Green, A., Ali, B., Naeem, A., & Ross, D. (2000). The allocation and budgetary mechanisms for decentralized health systems: Experiences from Balochistan, Pakistan. *Bulletin of the World Health Organization*, *78*(8), 1024–1035.

Green, A., & Collins, C. (2003). Health systems in developing countries: Public sector managers and the management of contradiction and change. *International Journal of Health Planning and Management*, *18*, 67–78.

Grindle, M. S. (1997). Divergent cultures? When public organisations perform well in developing countries. *World Development*, *25*(4), 481–495.

Hanney, S., Gonzalez-Block, M., Buxton, M., & Kogan, M. (2003). The utilisation of health research in policy-making: Concepts, examples and methods of assessment. *Health Research Policy and Systems*, *1*(1), 2.

Health Metrics Network (HMN). (2008). *Framework and standards for country health information systems* (2nd ed.). Geneva, Switzerland: HMN, World Health Organization.

Henderson, L., & Tulloch, J. (2008). Incentives for retaining and motivating health workers in Pacific and Asian countries. *Human Resources for Health*, *6*(1), 18.

Hersey, P., Blanchard, K. H., & LaMonica, E. L. (1978). A situational response to supervision: Leadership theory and the supervising nurse. In J. S. Rackich & K. Darr (Eds.), *Hospital organization and management: Text and readings* (pp. 184–191). New York, NY: Spectrum.

Herzberg, F., Mausner, B., & Snyderman, B. B. (1959). *The motivation to work*. New York, NY: John Wiley & Sons.

Hudson, J., & Lowe, P. (2004). *Understanding the policy process: Analysing welfare policy and practice.* Bristol, UK: Policy Press.

Hutubessy, R., Chisholm, D., Edejer, T., & World Health Organization (WHO). (2003). Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. *Cost Effectiveness and Resource Allocation*, *1*, 8.

International Medical Products Anti-Counterfeiting Taskforce (IMPACT). (2010). Retrieved from http://apps.who.int/medicinedocs/en/d/Js20967en/

Kapiriri, L., Norheim, O. F., & Heggenhougen, K. (2003). Using burden of disease information for health planning in developing countries: The experience from Uganda. *Social Science & Medicine*, *56*(12), 2433–2441.

Karki, D., Mirzoev, T., Green, A., Newell, J., & Baral, S. (2007). Costs of a successful public– private partnership for TB control in an urban setting in Nepal. *BMC Public Health*, *7*(1), 84.

Katzenbach, J. R., & Smith, D. K. (2005, July–August). The discipline of teams. *Harvard Business Review*, 1–9.

Kerrigan, J. E., & Luke, J. S. (1987). Management training strategies for developing countries. Boulder, CO: Lynne Rienner Publishers.

Kessler, I. (1994). Performance pay. In K. Sisson (Ed.), *Personnel management: A comprehensive guide to theory and practice in Britain* (pp. 465–494). Oxford, UK: Blackwell Business.

King, K., & Cunningham, G. (1995). Leadership in nursing: More than one way. *Nursing Standard*, *10*(12–14), 3–14.

Koduah, A. (2016). *National level maternal health decisions: Towards an understanding of health policy agenda setting and formulation in Ghana.* PhD thesis, Wageninghen University, Wageningen, Netherlands.

Kolehmainen-Aitken, R.-L. (2004). Decentralisation's impact on the health workforce: Perspectives of managers, workers and national leaders. *Human Resources for Health*, *2*, 5.

Kolehmainen-Aitken, R.-L., & Shipp, P. (1990). "Indicators of staffing need": Assessing health staffing and equity in Papua New Guinea. *Health Policy and Planning*, *5*(2), 167–176.

Kozlowski, S. W. J., & Ilgen, D. R. (2006) Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest, 7,* 3.

Kruk, M. E., Ling, E. J., Bitton, A., Cammett, M., Cavanaugh, K., Chopra, M., . . . Warnken, H. (2017). Building resilient health systems: A proposal for a resilience index. *BMJ,* 357(j2323).

Kwamie, A., Agyepong, I., & van Dijk, H. (2015). What governs district manager decision making? A case study of complex leadership in Dangme West District, Ghana. *Health Systems & Reform, 1*(2), 167–177. doi:10.1080/23288604.2015.1032475

Kwamie, A., van Dijk, H., & Agyepong, A. (2014). Advancing the application of systems thinking in health: Realist evaluation of the Leadership Development Programme for district manager decision-making in Ghana. *Health Research Policy and Systems, 12,* 29.

Lafond, A. K. (1995). Improving the quality of investment in health: Lessons on sustainability. *Health Policy and Planning*, *10*, 63–76.

Langlios, E. V., Montekio, V. B., Young, T., Song, K., & Alcalde-Rabanal, J. (2016). Enhancing evidence informed policymaking in complex health systems: Lessons from multi-site collaborative approaches. *Health Research Policy and Systems, 14,* 20.

Lê, G., Ha, B. T. T., Mirzoev, T., & Nguyen, N. T. (2015). Negotiating compliance: The case of autonomous hospitals in Vietnam. *Development Policy Review, 33*(3), 373–390.

Lehmann, U., & Gilson, L. (2015). Action learning for health system governance: The reward and challenge of co-production. *Health Policy and Planning, 30*(8), 957–963.

Lippeveld, T., Sauerborn, R., & Bodart, C. (2000). *Design and implementation of health information systems*. Geneva, Switzerland: World Health Organization.

Mampe, T., Schneider, H., & Reagon, G. (2016, December). Effectiveness of ward based outreach teams in the North West Province: An Evaluation. University of the Western Cape, South Africa. Retrieved from https://www.researchgate.net/publication/313067745_Effectiveness_of_Ward_Based_Outreach_Teams_in_the_North_West_Province_an_Evaluation

Management Sciences for Health (MSH). (2003). *Supervision guidelines*. Boston, MA: Author. Retrieved from https://www.k4health.org/toolkits/leadershipmgmt/supervision-guidelines

Management Sciences for Health (MSH) & World Health Organization (WHO). (1997). *Managing drug supply* (2nd ed.). West Hartford, CT: Kumarian Press.

Martin, D., & Singer, P. (2003). A strategy to improve priority setting in health care institutions. *Health Care Analysis*, *11*, 59–68.

McCoy, D. C., Hall, J. A., & Ridge, M. (2012). A systematic review of the literature for evidence on health facility committees in low- and middle- income countries. *Health Policy and Planning, 27,* 449–466.

McIntyre, D., Thiede, M., Dahlgreen, G., & Whitehead, M. (2006). What are the economic consequences of households of illness and of paying for health care in low- and middle-income country contexts? *Social Science and Medicine, 62,* 858–865.

McMahon, R., Barton, E., & Piot, M. (1992). On being in charge: A guide to management in primary health care (2nd ed.). Geneva, Switzerland: World Health Organization.

Mikkelsen-Lopez, I., Wyss, K., & de Savigny, D. (2011). An approach to addressing governance from a health system framework perspective. *BMC International Health and Human Rights, 11*(1), 1–11.

Mills, A. (1998). To contract or not to contract? Issues for low and middle income countries. *Health Policy and Planning*, *13*(1), 32–40.

Mills, A., Brugha, R., Hanson, K., & McPake, B. (2002). What can be done about the private health sector in low-income countries? *Bulletin of the World Health Organization, 80,* 325–330.

Mirzoev, T., Das, M., Ebenso, B., Uzochukwu, B., Rawat, B., Blok, L., . . . Huss, R. (2017). Contextual influences on the role of evidence in health policy development: What can we learn from six policies in India and Nigeria? *Evidence & Policy: A Journal of Research, Debate and Practice, 13*(1), 59–79.

Mirzoev, T. N., & Green, A. T. (2017). Planning, for public health policy. In S. R. Quah & W. C. Cockerham (Eds.), The international encyclopedia of public health (2nd ed., Vol. 5, pp. 489–499). Oxford, UK: Academic Press.

Molyneux, S. Atela, M., Angwenyi, V., & Goodman, C. (2012). Community accountability at peripheral health facilities: A review of the empirical literature and development of a conceptual framework. *Health Policy and Planning, 27,* 541–554.

Nancholas, S. (1998). How to do or not to do a logical framework. *Health Policy and Planning*, *13*(2), 189–193.

Noakes, J. (1992, June 18). Team spirit. *Health Service Journal*, 26.

Oborien, K., Harris, B., Eyles, J., Orgill, M., McIntyre, D., Chimbindi, N., & Goudge, J. (2015). Understanding roles, enablers and challenges of District Clinical Specialist Teams in strengthening primary health care in South Africa. In A. Padarath, J. King, & R. English (Eds.), *South African health review 2014/15.* (pp. 45–56) Durban, South Africa: Health Systems Trust.

Ogundeji, Y. K., Jackson, C., Sheldon, T., Olubajo, O., & Ihebuzor, N. (2016). Pay for performance in Nigeria: The influence of context and implementation on results. *Health Policy and Planning, 31*(8), 955–963.

O'Hare, B. (2015). Weak health systems and Ebola. *The Lancet Global Health, 3*(2), e71–e72.

Okello, D. R., & Gilson, L. (2015). Exploring the influence of trust relationships on motivation in the health sector: A systematic review. *Human Resources for Health, 13,* 16.

Olsen, I. T. (1998). Sustainability of health care: A framework for analysis. *Health Policy and Planning*, *13*(1), 287–295.

Oxman, A., Lavis, J., Fretheim, A., & Lewin, A. (2009). SUPPORT Tools for evidence-informed health Policymaking (STP) 16: Using research evidence in balancing the pros and cons of policies. *Health Research Policy and Systems, 7*(suppl. 1), S16.

Palmer, N. (2000). The use of private-sector contracts for primary health care: Theory, evidence and lessons for low-income and middle-income countries. *Bulletin of the World Health Organization*, *78*(6), 821–829.

Pearson, M. (2002). *Allocating public resources for health: Developing pro-poor approaches*. London, UK: DFID Health Systems Resource Centre.

Perry, J. L. Hondeghem, A., & Wise, L. R. (2010, September–October). Revisiting the motivational bases of "Public Service: Twenty Years of Research and an Agenda for the Future." *Public Administration Review, 70*(5), 681–690.

Prashanth, N. S., Marchal, B., Hoeree, T., Devadasan, N., Macq, J., Kegels, G., & Criel, B. (2012). How does capacity building programme for district of health managers work? A realist evaluation study protocol. *BMJ Open, 2*(2), 1–11.

Prashanth, N. S., Marchal, B., Hoeree, T., Devadasan, N., Macq, J., Kegels, G., & Criel, B. (2014). Advancing the application of systems thinking in health: a realist evaluation of a capacity building programme for district managers in Tumkur, India. *Health Research Policy and Systems, 12:*42.

Public Health Agency of Canada, Canadian Best Practices Portal. (2016, April 20). Key Element 6: Collaborate across sectors and levels. Retrieved from http://cbpp-pcpe.phac-aspc.gc.ca/population-health-approach-organizing-framework/key-element-6-collaborate-sectors-levels/

Pyone, T., Smith, H., & van den Broek, N. (2017). Frameworks to assess health systems governance: A systematic review. *Health Policy and Planning, 32*(5), 710–722.

Ramesh, M., Wu, X., & He, A. J. (2014). Health governance and healthcare reforms in China. *Health Policy and Planning, 29*(6), 663–672.

Reichenbach, L. (2002). The politics of priority setting for reproductive health: Breast and cervical cancer in Ghana. *Reproductive Health Matters, 10*(20), 47–58.

RESYST. (2014, October) What is strategic purchasing for health? Retrieved from http://resyst.lshtm.ac.uk/sites/resyst.lshtm.ac.uk/files/docs/reseources/Purchasing%20brief.pdf

Ricketts, J. A. (2010). The making of Jamaica's "National Policy for Persons with Disabilities 2000": Macro, meso and micro factors. *Disability & Society, 25*(5), 551–564.

Rockers Peter, C., & Bärnighausen, T. (2013). Interventions for hiring, retaining and training district health systems managers in low- and middle-income countries. *Cochrane Database of Systematic Reviews*. doi:10.1002/14651858.CD009035.pub2

Rogers, D. L., & Whetten, D. A. (1982). *Interorganizational coordination: Theory*, *research*, *and implementation*. Ames, IA: Iowa State University Press.

Rondeau, K. V., & Wager, T. H. (2001). Impact of human resource management practices on nursing home performance. *Health Service Management Research*, *14*(3), 192–202.

Roome, E., Raven, J., & Martineau, T. (2014). Human resource management in post-conflict health systems: Review of research and knowledge gaps. *Conflict and Health, 8,* 18.

Rowbottom, R., & Billis, D. (1987). *Organisational design: The work level approach*. Aldershot, UK: Gower.

Schick, A. (1998). Why most developing countries should not try New Zealand's reforms. *World Bank Research Observer*, *13*(1), 123–131.

Schmidt, H., Gostin, L. O., & Emanuel, E. J. (2015). Public health, universal health coverage, and Sustainable Development Goals: Can they coexist? *The Lancet, 386,* 928–930.

Scott, V., Dinginto, N., & Xapile, Z. (2015). Operational health service management: Understanding the role of information in decision-making. In A. Padarath, J. King, & R. English (Eds.), *South African health review 2014/15* (pp. 141–150). Durban, South Africa: Health Systems Trust. Retrieved from https://www.profnetmedical.co.za/media/1184/south-african-health-review-2014-15-part-11compressed.pdf

Sheikh, K., Saligram, P. S., & Hort, K. (2015). What explains regulatory failure? Analysing the architecture of health care regulation in two Indian states. *Health Policy and Planning, 30*(1), 39–55.

Siddiqi, S., Masud, T. I., Nishtar, S., Peters, D. H., Sabri, B., Bile, K. M., & Jama, M. A. (2009). Framework for assessing governance of the health system in developing countries: Gateway to good governance. *Health Policy*, *90*(1), 13–25.

Stanfield, S., Walsh, J., Prata, N., & Evans, T. (2006). Information to improve decision making for health. In D. Jamison, J. Breman, A. Measham, G. Alleyne, M. Claeson, D. Evans, . . . P. Musgrove (Eds.), *Disease control priorities in developing countries* (2nd ed., pp. 1017–1030). Washington, DC/New York, NY: World Bank & Oxford University Press.

Stewart, J., & Ranson, S. (1994). Management in the public domain. In D. McKevitt & A. Lawton (Eds.), *Public sector management theory*, *critique and practice* (pp. 54–70). London, UK: Sage.

Stordeur, S., Vandenberghe, C., & D'hoore, W. (2000). Leadership styles across hierarchical levels in nursing departments. *Nursing Research*, *49*(1), 37–43.

Storey, J. (1994). Management development. In K. Sisson (Ed.), Personnel management: A comprehensive guide to theory and practice in Britain (pp. 365–396). Oxford, UK: Blackwell Business.

Strike, A. J. (1995). *Human resources in health care: A manager's guide*. Oxford, UK: Blackwell Science.

Thiede, M., Akweongo, P., & McIntyre, D. (2007). Exploring the dimensions of access. In D. McIntyre & G. Mooney (Eds.), *The economics of health equity* (pp. 103–123). New York, NY: Cambridge University Press.

Transaid. (2001). *Transport management manual*. London: Transaid Worldwide.

Transaid. (2008). Introduction to Transaid's transport management system manual. Draft 3: August 2008. Retrieved from http://allindiary.org/download/introduction-to-transaids-transport-management-system-manual-2008/

Transaid. (2017). Home page. Retrieved from http://www.transaid.org/

Trap, B., Todd, C. H., Moore, H., & Laing, R. (2001). The impact of supervision on stock management and adherence to treatment guidelines: A randomised control trial. *Health Policy and Planning*, *16*(3), 273–280.

Travis, P., Egger, D., Davies, P., & Mechbal, A. (2002). *Towards better stewardship: Concepts and critical issues*. Geneva, Switzerland: World Health Organization.

Vandenabeele, W. (2007). Toward a public administration theory of public service motivation. *Public Management Review, 9*(4), 545–556.

Vian, T., & Bicknell, W. J. (2014). Good governance and budget reform in Lesotho public hospitals: Performance, root causes and reality. *Health Policy and Planning, 29*(6), 673–684.

Vian, T., Brinkerhoff, D. W., Feeley, F. G., Salomon, M., & Vien, N. T. K. (2012). Confronting corruption in the health sector in Vietnam: Patterns and prospects. *Public Administration and Development, 32*(1), 49–63.

Voce, A., Bhana, R., Monticelli, F., Makua, M., Pillay, Y., Ngubane, G., & Kauchali, S. (2014). District clinical specialist teams. In A. Padarath, J. King, & R. English (Eds.), *South African health review 2014/15* (pp. 45–58). Durban, South Africa: Health Systems Trust.

Wang, Y., Collins, C., Vergis, M., Gerein, N., & Macq, J. (2007). HIV/AIDS and TB: Contextual issues and policy choice in programme relationships. *Tropical Medicine & International Health*, *12*(2), 183–194.

Weightman, J. (1996). *Managing people in the health service*. London, UK: Institute of Personnel and Development.

Willis-Shattuck, M., Bidwell, P., Thomas, S., Wyness, L., Blaauw, D., & Ditlopo, P. (2008). Motivation and retention of health workers in developing countries: A systematic review. *BMC Health Services Research*, *8*(1), 247.

World Health Organization (WHO). (2000). *The world health report 2000: Health systems: Improving performance.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2004). *Developing health management information systems: A practical guide for developing countries.* Manila, Philippines: Author. Retrieved from http://whqlibdoc.who.int/publications/2004/9290611650.pdf

World Health Organization (WHO). (2006a). *A guide to WHO's role in sector-wide approaches to health development*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2006b). *The world health report 2006: Working together for health.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2007a). *Everybody's business. Strengthening health systems to improve health outcomes. WHO's framework for action*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2007b). *Towards better leadership and management in health: Report on consultation on strengthening leadership and management in low-income countries.* Department for Health Policy and Development Services. Geneva, Switzerland: Author.

World Health Organization (WHO). (2010a). *Monitoring the building blocks of health systems: A handbook of indicators and their measurement strategies.* Geneva, Switzerland: Author.

World Health Organization (WHO). (2010b). Human resources for health (HRH) tools and guidelines. Retrieved from http://www.who.int/hrh/tools/management/en/

World Health Organization (WHO). (2010c). The WHO prequalification project. Retrieved from https://allcountries.org/health/the_who_prequalification_project.htmlWorld Health Organization

World Health Organization (WHO). (2014). *Health in all policies: Framework for country action.* Geneva, Switzerland: Author. Retrieved from http://www.who.int/healthpromotion/frameworkforcountryaction/en/

World Health Organization (WHO). (2017). *Together on the road to universal health coverage: A call to action.* Geneva, Switzerland: Author. Retrieved from http://apps.who.int/iris/bitstream/10665/258962/1/WHO-HIS-HGF-17.1-eng.pdf?ua=1

Xyrchis, A., & Lowton, K. (2008). What fosters or prevents interprofessional teamworking in primary and community care? A literature review. *International Journal of Nursing Studies, 45*(1), 140–153.

CHAPTER 16

Pharmaceuticals

Kara Hanson, Stuart Anderson, Henry Lishi Li, Barbara McPake, Benjamin Palafox, Giuliano Russo, and Rima Shretta

Introduction

Pharmaceuticals are a critical element of the health system. They consume, on average, 25% of a health system's total expenditures (Lu, Hernandez, Abegunde, & Edejer, 2011) and are a potent symbol of effective health system functioning: When medicines and vaccines are not available in health facilities, the confidence of both patients and providers is undermined. Their central role in health systems is reflected in the Sustainable Development Goal target for universal health coverage, which highlights the importance of access to safe, effective, quality, and affordable essential medicines and vaccines for all, and captures the need for households to be protected from the impoverishing effects of health expenditures, of which expenditures on medicines represent an important part. Pharmaceutical access is influenced by all the core health system building blocks (see also the *Design of Health Systems* chapter):

- Service delivery, which affects how patients access medicines and whether they are appropriately prescribed and used
- Information systems, which are responsible for monitoring appropriate use and effective supply management
- Financing systems, which need to generate sufficient resources and protect households against out-of-pocket payments
- Effective governance and stewardship, capable of stimulating pharmaceutical research and development and regulating the introduction of new products; ensuring the quality, safety, and appropriate distribution of existing products; and providing transparent procurement

Despite these products' importance, large numbers of people around the world lack access to essential medicines. Improving access to pharmaceuticals is, therefore, a global public health priority, as evidenced by the number of international health movements that have included it as part of their central goals. The first Model Essential Drugs list, created in 1977, coincided with the development of the primary healthcare movement. The 1987 Bamako Initiative committed to improving the availability of essential medicines across sub-Saharan Africa. From 1990 onward, a range of global health initiatives, including the Global Fund to Fight for AIDS, Tuberculosis and Malaria; Stop TB; and GAVI took up the challenge of improving access to essential medicines and vaccines, providing both financing and support in procurement to ensure quality standards. In the present era of universal health coverage, there is renewed focus on securing sufficient funds to meet the healthcare needs of populations through the expansion of public health spending, but also recognition of the need for these resources to be used efficiently. For medicines, meeting that goal requires a focus on these products' selection,

procurement, distribution, and appropriate use. Antimicrobial resistance poses a major challenge to the existing pharmaceutical business model and highlights the tensions between enabling access to vital medicines while simultaneously seeking to restrain their use so as to conserve their effectiveness. Moreover, as the epidemiologic transition unfolds in all parts of the world and the burden of noncommunicable disease increases, new challenges arise. Chronic conditions often require lifelong use of medicines, so incorporating these medications into care benefit packages and ensuring their regular availability is critical for treating such conditions.

Three key components of access are addressed in this chapter: availability (addressing both "upstream" issues of product discovery and development and "downstream" challenges of national pharmaceutical supply systems), affordability, and safe and effective use of medicine. Two current issues in pharmaceutical policy—growing resistance to antimicrobials, and the need to ensure adequate availability, affordability, and use of medicines for chronic conditions—are presented at the end of the chapter as case studies to illustrate these issues.

Throughout the chapter, we strive to maintain consistency in terminology, while recognizing that the preferred names of the products that are discussed have changed over time. Currently, *medicine* is the preferred term over *drug* for any substance that is used to modify physiological systems or pathological states for the benefit of the recipient, as the latter is now widely regarded as referring to illicit substances. *Pharmaceutical* is a broader term that generally encompasses both medicines (prescription and over-the-counter) and vaccines in their finished form. Although *medicine* and *pharmaceutical* are often used interchangeably, the latter may also encompass other products of the pharmaceutical industry—for example, the active pharmaceutical ingredients (API) used in production, other biologicals such as blood products and insulin, veterinary medicines, and diagnostic products such as those used for blood typing and test kits for the human immunodeficiency virus (HIV) and malaria parasites (Anderson & Huss, 2004; World Health Organization [WHO], 2002).

Other important and germane definitions and concepts include *essential medicines*, which are defined as the medicines that satisfy the priority healthcare needs of the population, and that are selected based on public health relevance, evidence on efficacy and safety, and comparative cost-effectiveness. *Generics* are pharmaceutical products that are intended to be interchangeable with the original innovator product and are marketed after the expiry of an original product's patent or other exclusivity rights. Generic products may be marketed either unbranded ("commodity generics") or under a new brand name ("branded generics") (Anderson & Huss, 2004; WHO, 2010b).

Pharmaceutical Availability: Upstream Issues

One of the reasons why many people worldwide do not have access to medicines may be that the right sorts of products have not been developed yet, or that they are marketed at prices unaffordable for many. The following sections present the so-called upstream availability issues, discussing in turn the inherent complexity for the pharmaceutical industry of inventing new medicines, the performance and flaws of the current research and development (R&D) system, and the alternative strategies that have been developed to address these shortcomings.

A starting point for securing access to pharmaceuticals in high-income countries as well as low- and middle-income countries is the development of the right products—that is, products that are safe, effective, and acceptable and that address major causes of disease burden. Pharmaceutical R&D mostly takes place in high-income countries, and is driven by a profit logic. Understanding the incentives operating in the industry is, therefore, essential as an entry point to influencing investment choices.

The Current System for Researching and Developing Pharmaceuticals

Developing a new pharmaceutical product is a long, heavily regulated, and expensive process, estimated to take 5 to 15 years, and to cost in excess of $2.5 billion (DiMasi, Grabowski, & Hansen, 2016). The drug discovery cycle typically goes through seven different phases: prediscovery, discovery, preclinical testing, clinical trials, regulatory agency review, scale-up to marketing, and postmarketing surveillance. This process involves laboratory experimentation, animal testing, three phases of clinical trials on humans, and several regulatory reviews before a drug can be made commercially available (**FIGURE 16-1**). In the United States, the regulatory function for pharmaceuticals is performed by the Food and Drug Administration (FDA).

The probability of identifying a promising molecule is reported to be small in the initial phase of R&D (0.1%): More than 5,000 molecules may have to be

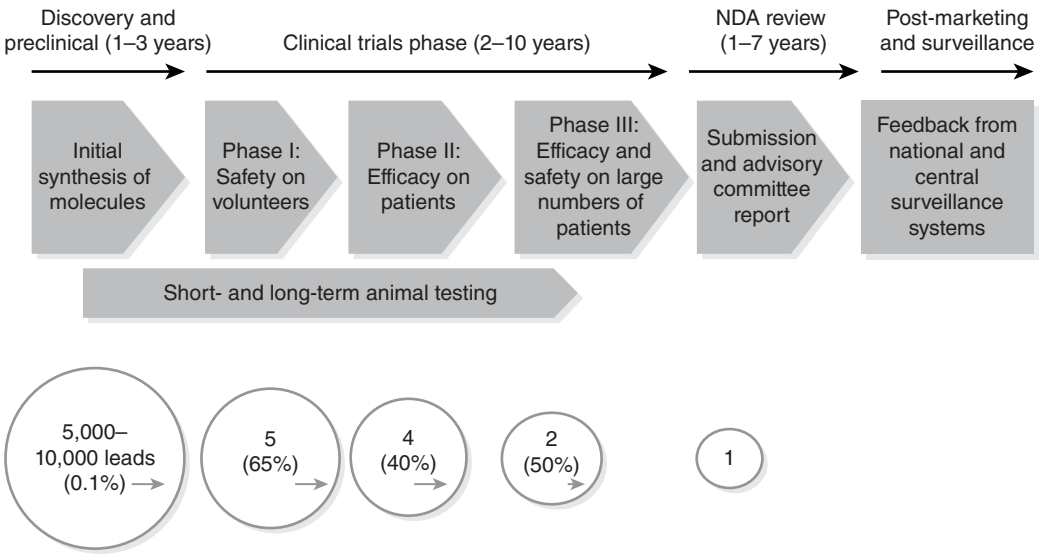


FIGURE 16-1 Overview of the key stages involved in the pharmaceutical R&D process and probabilities of compounds progressing to the next phase.

Modified from DiMasi et al., 2006; Petrova, 2014.

synthesized to identify 5 candidates that are suitable to progress to the next phase. Phase I clinical trials involving small numbers of volunteers are conducted to test the compounds for safety, tolerability, and stability of the active ingredient. Efficacy and dosage are explored in Phase II small trials on tens of patients. Safety and efficacy are then tested in larger trials involving thousands of patients from real-life settings in the final clinical phases (Petrova, 2014). The cumulative probability of success for the three clinical trial phases alone has been estimated to be in the range of 12% to 33%; considerable costs are associated with the preclinical industrial phase (approximately 25% of the total costs), and with the clinical trial phases (60% of the total costs) (DiMasi et al., 2016; Petrova, 2014).

To create incentives for companies to engage in this costly and risky pharmaceutical innovation, governments award inventors with patents, granting the inventors 20 years of exclusive rights to exploit the production of the patented medicine worldwide. Patents are defined as exclusive property rights on intangible creations of the human mind—hence the term "intellectual property rights" (IPR). While several forms of IPR protection are available, the product covered by a patent must be something that has never been previously disclosed anywhere in the world, and that would not be obvious to a person ordinarily skilled in the field involved. This is a crucial point for pharmaceutical entities, as their manufacturing process is often easy to replicate and can be copied with a fraction of the investment required for research and clinical testing (Saha & Bhattacharya, 2011). The appropriateness of 20 years as a patent term has been contested, with some arguing that a long protection period provides a negative incentive for real innovation (Leoni & Sandroni, 2016). In contrast, the industry has sought to extend these terms for medicines for which efforts are being made to conserve their efficacy by limiting their use (see the later discussion on antimicrobials) (Outterson, Samora, & Keller-Cuda, 2007).

Since the vast majority of medicine discovery activities are carried out in high-income countries and patents are awarded by specialized agencies based there (such as the U.S.-based FDA and the European Medicines Agency [EMA]), patents are upheld worldwide through the institution of Trade-Related Intellectual Property Rights (TRIPS), and enforced by the World Trade Organization (WTO) through its member nations. Negotiated between 1986 and 1994, the TRIPS agreement set the minimum standards for intellectual property protection that all WTO member countries—some of which did not provide patent protection for pharmaceuticals prior to the agreement—must respect. It also specified mechanisms for local enforcement and procedures for resolution of conflicts (WHO, 2005). Pharmaceutical patents and TRIPS were originally established as a remedy to the global markets' failure to protect intangible products of intellectual activity; it bans potential manufacturers from free-riding on inventions, thereby creating a financial "pull incentive" for innovation. The TRIPS agreement states that all pharmaceutical patents must be available for at least 20 years from the filing date for all products "invented" after January 1, 1995. To abide

by the agreement, WTO members must modify their intellectual property legislation to align them with the TRIPS standards, although countries at different levels of economic development were given different deadlines to comply with this mandate (e.g., the deadline for the least developed countries is 2033).

Although they are well intentioned, by imposing worldwide legal monopolies, patents prevent price competition and introduce a different market failure in the global market for pharmaceuticals, as the ensuing higher prices restrict consumption of pharmaceuticals (Levy & Rizansky, 2014). Health-sector advocates have also argued that the TRIPS standards should take protection of public health into account, so a provision to allow the patent requirement to be circumvented in the event of "public health emergencies" was included in the agreement (WHO, 2001a). To further clarify this provision, and after much pressure from the international health community, the WTO produced the Declaration on the TRIPS Agreement and Public Health, also known as the Doha Declaration of 2001, and a subsequent Decision on the Interpretation of Paragraph 6 of the TRIPS Agreement was reached in 2003. These clarifications affirmed the flexibilities available to member states seeking to protect public health (Kerry & Lee, 2007) (for further detail, see the *International Trade and Health* chapter).

Some of these measures, such as tiered pricing, aim to create the possibility of charging different sale prices in different countries according to consumers' differential ability to pay (Nicol & Owoeye, 2013). Such practice, however, creates the risk of "parallel imports," in which patented medicines are sold at a lower price in a low-and middle-income country (LMIC), only to be reimported illegally into high-income countries.

Other flexibilities, such as compulsory licenses, limits on data protection, patentability, and parallel imports, aim at allowing national governments to break patents and produce (or outsource production to a third party) generic versions of patented medicines in case of public health emergencies (Musungu & Oh, 2006). Such flexibilities have only rarely been used in LMICs, either owing to a lack of the legal and technical expertise required to take advantage of these complex options, or out of fear of trade retaliation from high-income countries, such as being included in annual watch lists like the Special 301 Report issued annually by the U.S. Trade Representative (Wirtz et al., 2017). Nevertheless, TRIPS flexibilities have proved useful in improving access to medicines in LMICs (see the "Affordability" section later in this chapter).

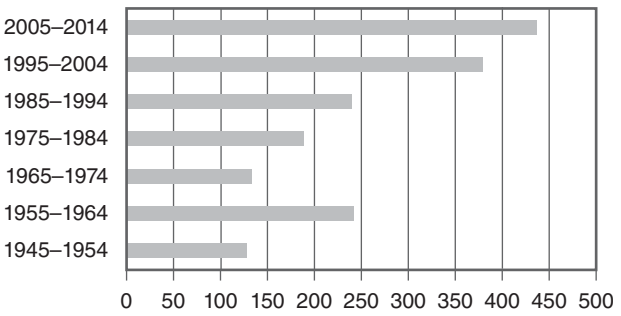


FIGURE 16-2 New molecular entity patents awarded by the FDA in the last seven decades.

Modified from U.S. Food and Drug administration, several years.

Availability Challenges with the Current Pharmaceutical R&D System

The pharmaceutical industry has been criticized for underperforming in terms of discovering new pharmaceutical products, and for being in a productivity crisis (Scannell & Bosley, 2016). Yet in a review looking back at the past 60 years of pharmaceutical discoveries, Munos (2009) showed that the FDA, which is perhaps the world's most important pharmaceutical regulatory agency, has awarded a steady number of patents for new molecular entities (NMEs) each year—45, on average. This number has been increasing in recent years (**FIGURE 16-2**).

Despite the undeniable advances of the last decades in pharmaceutical innovation, certain types of diseases appear to have been neglected by the R&D system. Light and Lexchin (2012) have identified a preponderance of FDA approvals for medicines with few significant therapeutic gains starting from the 1990s, caused by a change in the industry business model from seeking blockbusters to turning out small-variation medicines. This different level of investment in medicines markets is consistent with the distinction drawn by the 2001 Commission for Macroeconomics and Health (WHO, 2001a) between three categories of diseases:

- Type I diseases, which are prevalent in both high-income and LMICs. Their substantial commercial markets in high-income countries drive R&D activity, so there is no product shortfall. Affordability is often an issue, however. Examples include noncommunicable diseases such as diabetes, cardiovascular diseases (CVD), and tobacco-related illnesses, as well as infectious diseases such as pneumonia, hepatitis B, and *Haemophilus influenzae* type b (Hib).

TABLE 16-1 Diseases, Medicines, and R&D Issues

Type of Diseases	Geographical Location of Affected Population	Medicine Examples	Specific R&D Issues
Rare diseases	LMICs, HICs	Orphan drugs; hepatitis G; Fabry disease; hemophilia; cystic fibrosis	Limited sales due to small group of potential patients Need to charge high prices to recoup R&D investment
Tropical diseases	LMICs in the tropics	Dengue; leishmaniosis; leprosy	Limited ability to pay from potential consumers Risk of price controls imposed by authorities
Complex diseases	Mostly HICs, but also LMICs	Specialty medicines; specific cancer types; different type of hepatitis	Complex macromolecules requiring expensive technology Challenge to establish bioequivalence for biosimilar Extended data protection exclusivity granted for biologics
Antimicrobial-resistant infections	HICs and LMICs	Antibiotics for hospital superbugs; rising AMR levels in minimally regulated prescribing settings	Likely delays in wide commercialization due to prescription policies, leading to limited high-price sales during patent coverage

Abbreviations: AMR = antimicrobial resistance; HICs = high-income countries; LMICs = low- and middle-income countries.

- Type II diseases, which are "incident in both rich and poor countries, but with a substantial proportion of the cases in poor countries"—for example, acquired immunodeficiency syndrome (HIV/AIDS) and tuberculosis (TB). These diseases have modest markets in wealthy countries; thus, some R&D activity exists, although it is often targeted at high-income groups rather than LMICs (e.g., short-term courses of antimalarial agents for travelers and members of the military, or complex high-priced antiretroviral medicines).
- Type III diseases, which occur overwhelmingly or exclusively in LMICs and have no commercial market to drive R&D—for example, Chagas disease, dengue fever, Buruli ulcer, and African sleeping sickness (trypanosomiasis). There are often limited or no treatments for these diseases.

It is apparent that products aimed at certain diseases give the pharmaceutical companies access to more lucrative markets than others, and that this factor drives the industry's R&D investment decisions. A taxonomy of diseases and medicines for which the current R&D system does not seem to work adequately is offered in the following paragraphs (**TABLE 16-1**).

Aronson (2006) defines as "orphan drugs" those NMEs not associated with any recognized disease (such as for hepatitis G, a detectable hepatitis virus with as yet unknown implications), and those agents used to treat diseases affecting fewer than 5 to 10 individuals per 10,000 population. Despite the obvious lack of incentive for the industry to invest in such narrow market segments, the patents awarded by the FDA for so-called orphan-designation drugs have increased in recent years, and currently represent approximately 40% of the total number of patent awards (**FIGURE 16-3**).

Neglected tropical diseases (NTDs) is a term used to identify a diverse group of communicable diseases affecting individuals in low- and middle-income tropical and subtropical countries. Unlike the conditions targeted by orphan drugs, large numbers of people are affected by NTDs, but because of their limited disposable income, suppliers' ability to charge high prices for

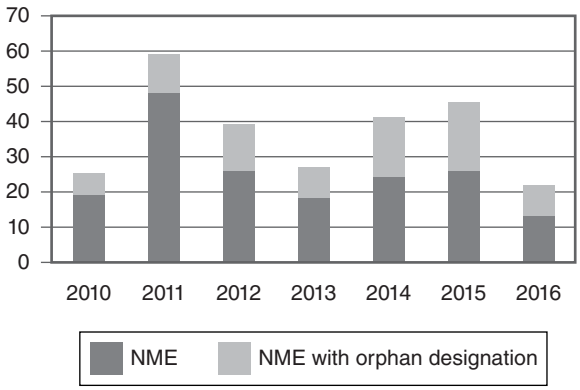


FIGURE 16-3 FDA-awarded patents for NMEs with and without orphan designation, 2010–2016.

related medicines is somewhat limited, and likely to be regulated for equity objectives. As a result, pharmaceutical companies claim that researching and developing medicines for NTDs is risky, as the expected returns would be unlikely to cover the costs in this area (Trouiller et al., 2002). The International Federation of Pharmaceutical Manufacturers Associations (IFPMA) reports that in 2014 multinational pharmaceutical companies were the third largest funders for R&D in NTDs, contributing 13% to the $3.3 trillion global expenditures in this area—up from 7% in 2007; by comparison, 62% of these activities were funded by public sources, and 20% by philanthropic organizations (IFPMA, 2017).

Specialty medicines for complex diseases are often, but not exclusively, complex macromolecules produced through recombinant DNA technology. They are typically used in new forms of treatment for cancer and hepatitis C. The costs involved in discovering and developing such entities are many times higher than those for regular small molecules, because of factors such as the technology involved in the process and the extended data protection exclusivity period granted for biologics. The lack of competitors in this market segment due to the challenge of establishing bioequivalence for biosimilars contributes to the high prices charged for such specialty products (Bigdeli, Peters, & Wagner, 2014).

For new antibiotics, the challenge of the current system of pharmaceutical innovation arises from the need to withhold the launch of new products to the market to delay the emergence of antimicrobial resistance (Power, 2006). The incentives influencing R&D for new antibiotics are addressed later in this chapter.

Alternative Push and Pull Mechanisms to Deliver Pharmaceutical Innovation

In response to the shortcomings of the privately funded R&D system highlighted in the previous subsection, new solutions have been proposed to create incentives for pharmaceutical innovation for specific diseases at affordable prices. Delinking R&D investment from costs and sales is typically the rationale for such strategies (Bigdeli et al., 2014), which have been broadly divided into "push," "pull," and mixed approaches (**TABLE 16-2**). "Push strategies" aim at providing an input at the beginning of the R&D process to help jump-start specific R&D activities, such as new facilities, laboratory capacity, or extra funds.

TABLE 16-2 Innovative R&D Models

Approach	Types	Examples
Push solutions	Product development partnerships (PDPs) between multiple public and private actors	■ GAVI Alliance to improve access to vaccines in LMICs ■ GlaxoSmithKline Diseases of the Developing World initiative ■ Medicines for Malaria Venture ■ International AIDS Vaccine Initiative ■ Global Alliance for TB Drug Development
	Public–private partnerships to generate novel scientific concepts through information and data sharing, and common infrastructure	■ Critical Path Institute ■ Innovative Medicines Initiative ■ Structural Genomics Initiative ■ Open Source Drug Discovery initiative
	Tax credits for R&D spending	■ U.S. tax breaks and differential income tax rates for pharmaceutical companies ■ Drugs for Neglected Diseases initiative
	Targeted research grants programs	

Pull solutions	Prize funds for innovation	■ Currently under study in the United States and by WHO
	Priority review vouchers to fast-track approval process for neglected diseases medicines	■ Priority voucher accorded to Novartis for Coartem antimalarial medicine ■ Priority voucher accorded to Jansen for Sirturo TB medicine ■ Priority review voucher agreed in 2014 in the United States for future Ebola medicines
	Advance market commitments	■ Gates Foundation, World Bank, UNICEF, and U.K. Department for International Development advance market commitment for vaccines ■ GAVI's Pneumococcal advance market Commitment
Mixed (regulatory) solutions	Regulatory measures discriminating positively towards neglected diseases medicines	■ Orphan drugs legislation in the United States and the European Union ■ Mandatory insurance coverage for specialty medicines in Thailand
	Patent pools	■ UNITAID's innovative financing mechanisms to stimulate the market for HIV tests ■ Medicines Patent Pool
	Market-shaping interventions to reduce transaction costs, increase market information, and balance supplier and buyer risk	■ DfID-UNITAID and CHAI's efforts to forecast increased demand for second-line antiretroviral (ARV) medications to reduce their price in 25 countries ■ Global Fund, President's Malaria Initiative, and UNITAID's decisions to make available planned public purchases of artemisinin-based combination therapy drugs to reduce their price volatility
	Strengthening LMICs' capacity to exploit TRIPS flexibilities	■ Incentives for local manufacturing of pharmaceuticals

Conversely, "pull strategies" aim at providing a "reward" at the end of the R&D process to entice participation, be it prize money, or privileged approval conditions for other medicines in a company's portfolio. Mixed approaches can create both inputs and rewards to R&D activity, typically through regulatory interventions.

Product development partnerships (PDPs) are not-for-profit organizational structures enabling entities from the public, private, academic, and philanthropic sectors to aggregate funding and capacity for the development of pharmaceuticals and public goods. Typically, PDPs target neglected diseases whose solutions lack commercial incentives and that disproportionately affect people in LMICs. Each partnership displays different features. In some cases, philanthropic institutions (such as the Bill & Melinda Gates Foundation) or national governments (such as France's support for UNITAID) provide the funds to set up specific laboratory capacity or an institute to mediate between different players, with academics and industry providing the research know-how. In these partnerships, the industry remains responsible for developing and selling the medicines, but commits to an agreed, lower profit margin (as in the case for the new malaria vaccine "RTS,S"—a partnership between GlaxoSmithKline and the Medicines for Malaria Venture).

In most cases, IPR on discoveries made through PDPs are owned by the private pharmaceutical partners that, in exchange for these rights, commit to

cap their profits to ensure wider access to the medicine. Despite some early success, such as Medicines for Malaria Venture's partnership with Novartis to develop a pediatric formulation of an anti-malarial agent, PATH's Meningitis Vaccine Project launch of a meningococcal vaccine for less than $0.50 per dose, and some fixed-dose combinations of malaria medications, no major breakthrough medicines have emerged so far from PDPs, partly because these experiences are quite recent (the oldest dating 15 years back) and the development of new medicines and vaccines is expected to take longer than this. IFPMA (2017) reports that in 2014 there were 17 active PDPs in the pharmaceutical sector.

Since 2012, the government of Brazil has engaged in PDPs with international pharmaceutical manufacturers. In those arrangements, production and technology of patented medicines are transferred to public laboratories for open local production, in exchange for the right to exclusive supply these medicines to the Brazilian national health service at negotiated prices for an agreed time-span—usually 5 years (Viana et al., 2016).

Precompetitive *public–private partnerships* aim to generate novel scientific concepts (e.g., research models) and infrastructures (e.g., databases) by fostering collaboration between multiple public and private entities based on mutual trust, pooling of complementary expertise and knowledge, and sharing of rewards. To avoid potential disputes, activities are limited to the precompetitive space; that is, they stop before the actual development of pharmaceutical products begins (Vrueh & Crommelin, 2017). Examples of such partnerships are the Critical Path Institute, whose mission is to foster the development of new evaluation tools and standards for drug therapy trials to accelerate regulatory qualification and medical product approval and adoption, and the Innovative Medicines Initiative, which aims to speed up the development of, and patient access to, innovative medicines, particularly for neglected diseases.

Advance market commitments (AMCs) are legally binding contracts between donors and R&D companies to subsidize the future purchase of a pharmaceutical (specifically, a vaccine) that is not yet available, if an appropriate vaccine is developed and if it is demanded by LMICs. The product profile and the price are agreed upon prior to the R&D being undertaken. Originally designed in 2005, AMCs have been specifically used to spur the development of vaccines for neglected tropical diseases. In 2010, the launch of an inexpensive vaccine for pneumococcal disease in more than 40 LMICs was made possible by an AMC organized by the GAVI Alliance (Scudellari, 2011).

Patent pools are mechanisms through which patent holders voluntarily exchange the property rights of their inventions for payment of royalties by manufacturers (Bermudez & 't Hoen, 2010). The Medicine Patent Pool is currently the only pool backed by the United Nations; it is supported by specific (innovative financing) taxes channeled through UNITAID. To date, the pool has signed agreements with nine private-sector patent holders for a total of 16 pharmaceutical products: 12 ARV medicines, 2 hepatitis C treatments, 1 TB treatment, and 1 HIV technology platform. Seventeen manufacturers have received sublicenses from the medicine patent pool to produce generic versions of those compounds (De Luca & Calzavara, 2016). Unfortunately, despite the existence of TRIPS flexibilities and exceptions to current patent laws, many LMICs lack the technical regulatory capacity or manufacturing muscle to take advantage of such provisions.

Policies aimed at strengthening local capacity to take full advantage of the current R&D system have been identified as a possible measure to increase access to pharmaceuticals worldwide (WHO, 2016). While the vast majority of pharmaceutical R&D activities are conducted in high-income countries, there is evidence that local production of medicines offers advantages for national supply systems (Mackintosh, Banda, Wamae, & Tibandebage, 2016). For example, having sizable national pharmaceutical industries allowed the governments of Brazil and Thailand to effectively negotiate down the price of patented ARV medications by credibly threatening the use of compulsory licenses (Flynn, 2008). Recent work from WHO (2016) suggests that strengthening the development of national intellectual property systems and local production capacity in LMICs could improve access to medicines worldwide.

Availability: Country-Level Distribution Systems and Pharmaceutical Management

Access to pharmaceuticals depends on strong pharmaceutical supply systems to ensure the consistent availability of affordable and high-quality diagnostic and treatment commodities at the places where the target population seeks care. Furthermore, the supply chain is responsible for transferring information on supply and demand back to the central level where planners—along with policy and decision makers—handle financial flows, so that the system is adequately resourced and replenished.

Pharmaceutical Supply Chain Systems: Introduction and Scope

In recent years a series of global initiatives has greatly increased the resources flowing into LMICs for the procurement of medicines. The value of donor-financed health commodities now exceeds $10 billion, and it continues to rise. While these resources are greatly welcomed as a means to improve the availability of essential commodities, they also have the potential to place considerable pressure on the already weak supply chain management systems in LMICs.

The pharmaceutical management framework (**FIGURE 16-4**) provides the foundation for improving access to medicines. It illustrates four basic pharmaceutical management functions: selection, procurement, distribution, and use. These are all inter-related, with each major function building on the previous one and leading logically to the next. A breakdown in one part of the framework leads to failure of the whole pharmaceutical management process and compromises access (Management Sciences for Health, 2012).

- *Selection* involves reviewing the prevalent health problems, identifying treatments of choice, choosing individual medicines and dosage forms, and deciding which medicines will be available at each level of health care.
- *Procurement* includes quantifying medicine requirements, selecting procurement methods, managing tenders, establishing contract terms, assuring pharmaceutical quality, and ensuring adherence to contract terms.
- *Distribution* includes clearing customs, stock control, stores management, and delivery to medicines depots and health facilities.

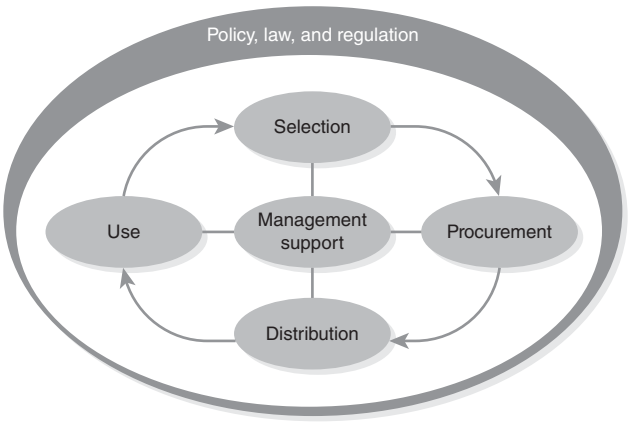


FIGURE 16-4 The pharmaceutical management framework.

Courtesy of Management Sciences for Health. (2012). *MDS-3: Managing access to medicines and technologies.* Arlington, VA: Management Sciences for Heatlh.

- *Use* includes diagnosing, prescribing, dispensing, and proper consumption by the patient.

At the center of the pharmaceutical management framework is a core of *management support* systems: organization, financing and sustainability, information management, and human resources management. Finally, the entire framework relies on *policy, law, and regulation*, which, when supported by good governance, establish and support the public commitment to quality essential medicine supply. This section focuses on the supply chain components of the framework: selection, procurement, and distribution. Medicine use is covered in a later section.

Selection

Essential medicines are those that satisfy the priority healthcare needs of the majority of the population. They are intended to be available within the context of functioning health systems at all times in adequate amounts, in the appropriate dosage forms, with assured quality and adequate information, and at a price the individual and the community can afford. Medicines should be selected according to the following criteria:

- Relevance to the pattern of diseases
- Proven efficacy and safety
- Evidence of performance in a variety of settings
- Quality
- Cost-effectiveness
- Pharmacokinetics
- Acceptability

In 1977, WHO defined the first *Model List of Essential Drugs*, a limited list of medicines whose use, it was hoped, would result in lower costs, better supply, and more rational use. The list has since been updated every two years, and many countries have used it as the basis for developing their own National Essential Medicines List. The 20th list was published in June 2017 (WHO, 2017g). In 2007, a separate *Model List of Essential Medicines for Children* was introduced. The sixth edition appeared in March 2017 and was amended in August 2017 (WHO, 2017f).

Procurement

Effective procurement ensures the availability of the right medicines of appropriate quality in the right quantities, at reasonable prices with timely delivery to avoid stock-outs. Effective procurement is a collaborative process between the procurement department of the national ministry of health and technical committees or

disease control programs (e.g., HIV/AIDS or malaria). These committees often make the final decision on which medicines to buy and in what quantities. In most cases in the public sector, a form of group purchasing is used in which a central procurement department, such as a central medical store, negotiates contracts for its members. The procurement process involves the following steps:

1. Review the medicine selections.
2. Determine the quantities needed, using the most appropriate method for forecasting based on the quality of data available (consumption and/or morbidity).
3. Reconcile needs and funds.
4. Develop supply schedules.
5. Choose the procurement method.
6. Locate and select suppliers.
7. Specify the contract terms.
8. Monitor the order status.
9. Receive and check the medicine.
10. Make payment.

Four main methods of procurement are used: open tender, restricted tender with performance monitoring, negotiated procurement, and direct procurement. Some countries and donors may opt for some form of direct procurement from nonprofit suppliers such as UNICEF or the International Dispensary Association Foundation. The Global Fund and most major donors require that tenders for the procurement of HIV/AIDS, malaria, and TB commodities be restricted to WHO-prequalified products wherever possible, to facilitate access to medicines that meet high standards of quality, safety, and efficacy (International Dispensary Association, 2017; United Nations International Children's Emergency Fund, 2017).

Challenges can arise at any point in the procurement process. For example, they may be the result of factors arising at the level of regulation, lack of financial and human resources, or broader management and organizational shortcomings such as fragmentation and corruption (**EXHIBIT 16-1**).

A number of tools are available to facilitate pharmaceutical procurement. The Global Fund secretariat developed the *Price and Quality Reporting System*, a publicly accessible online database, as a means to gather information about product prices, product quality, and supplier performance. This system represents a source of information that purchasers can use during negotiation of price and delivery conditions (Global Fund, 2017a). Management Sciences for Health has published an *International Drug Price Indicator Guide* since 1986 and updates it annually; it reports procurement prices from multiple international sources (Management Sciences for Health, 2017). In 2001, the Global Drug Facility was established by the Stop TB Partnership to enable health ministries in LMICs to procure quality medicines at competitive prices (Stop TB Partnership, 2017). Similarly, wambo.org is an online platform developed by the Global Fund to provide accessible price information, increased transparency, and improved reliability in the supply of medicines, health products, and nonhealth commodities necessary for HIV/AIDS, TB, and malaria programs. This platform is available to Global Fund recipients that use Global Fund grant funds to procure medicines through the Pooled Procurement Mechanism.

EXHIBIT 16-1 Common Procurement Challenges

- Absence of a comprehensive procurement policy
- Inadequate rules, regulations, and structures
- Public-sector staff with little experience of working with the private sector
- Government or donor funding that is insufficient or disbursed at irregular intervals
- Donor agencies with conflicting procurement regulations
- Fragmented medicine procurement at the provincial or district level
- Lack of unbiased market information
- Currency fluctuations
- Limited number of prequalified suppliers
- Long procurement lead times
- Insufficient product specifications
- Insufficient capacity and data for forecasting
- Poor procurement and supply planning
- Corruption and lack of transparency

Competitive tenders are recommended for most public-sector procurement. A formal tender process includes medicine selection, quantification, preparation of tender documents and contracts, notification and invitation to bid, receipt of bids, formal opening and collation of bids and supplier selection, award of contracts, performance monitoring of suppliers and clients, and enforcement of contract terms when necessary. Supplier selection should be based on formal written criteria.

Quantification and Forecasting

Quantification is the process used to estimate how much of a product is required for the purposes of procurement. It involves estimating the quantities needed of a specific item, the funding required for purchasing the item, and when the products should be delivered to ensure an uninterrupted supply for the program. The quantification process has two parts:

- *Forecasting:* Estimating the quantities of product required to meet customer demand that will actually be used during a particular time frame.
- *Supply planning:* Detailing the quantities required to fill the supply pipeline, costs, orders, and arrival dates of shipments.

Three main methods are used for forecasting medicine requirements. The *consumption method* uses historical data on medicine consumption. If these data are available from a reliable system of reporting, this technique provides the most accurate prediction of future needs. New programs or new treatments, however, will not have historical data on consumption available. Thus, in such cases, purchasers must rely on the *morbidity method*, which estimates future demand based on estimates of disease burden using data on incidence, cases (from attendance at health facilities) or episodes, and standard treatment guidelines. If there are no data on consumption or morbidity, it may be possible to use *proxy consumption*—that is, extrapolation of data from proxy facilities, regions, or countries—as a means of estimating need. **TABLE 16-3** summarizes the three main methods for quantification and forecasting. A fourth method, service-level projection of budget requirements, can be used for estimating budget needs.

Supply planning works alongside forecasting to ensure an optimal procurement and delivery schedule. Once the total requirements have been determined and the product specifications have been agreed, the estimated requirements should be translated into actual orders, including the timing of the orders, by taking the existing pipeline and seasonality into account (Management Sciences for Health, 2012).

Distribution

Effective and efficient distribution of medicines relies on good management and system design.

TABLE 16-3 Comparison of the Various Quantification Methods

Method	Uses	Essential Data	Limitations
Consumption	First choice for procurement quantifications given reliable data. Most reliable predictor of future consumption.	Reliable inventory records. Records of supplier lead time. Projected pharmaceutical costs.	Must have accurate consumption data. Can perpetuate irrational use.
Morbidity	Estimating need in new and scaling-up programs or disaster assistance. Comparing use with theoretical needs. Developing and justifying budgets.	Population and patient attendances. Actual or projected incidence of health problems. Standard treatments (ideal, actual). Records of supplier lead time. Projected pharmaceutical costs.	Morbidity data are not available for all diseases. Standard treatments may not really be used. Accurate attendance is difficult to predict.
Proxy consumption	Procurement quantification when other methods are unreliable. Comparing use with other supply systems.	Comparison area or system with good per-capita data on consumption, patient attendance, service level, and morbidity. Number of local health facilities by category. Estimates of local user population by age.	Questionable comparability of patient populations, morbidity, and treatment practices.
Service-level projection of budget requirements	Estimating budget needs.	Use by service levels and facility type. Average medicine cost per attendance.	Variable facility use, attendance, treatment patterns, and supply system efficiency.

Reproduced from Management Sciences for Health. (2012). *MDS-3: Managing access to medicines and technologies.* Arlington, VA: Management Sciences for Heatlh. http://apps.who.int/medicinedocs/documents/s19577en/s19577en.pdf

An efficient distribution system should have the following characteristics:

- Maintain an uninterrupted supply of medicines.
- Provide appropriate storage.
- Keep medicines in good condition throughout the distribution process.
- Minimize losses due to damage and expiry.
- Maintain accurate inventory management records.
- Use available transportation efficiently.
- Reduce pilferage.
- Provide accurate information for quantification.

A distribution system consists of four main elements (Management Sciences for Health, 2012):

- *System design:* Geographic or population coverage, number of levels in the system, push versus pull system, degree of decentralization.
- *Information system:* Inventory control, records, consumption reports, information flow.
- *Storage:* Selection of storage sites, building design, handling systems, order picking.
- *Delivery:* Collection versus delivery, choice of transport, vehicle procurement and maintenance, delivery schedules and routes.

As with procurement, distribution systems face a number of challenges that can arise at any point in the process (**EXHIBIT 16-2**). Some are a consequence of the system itself; for example, supply systems may be vertically multilayered and highly fragmented. The information management systems, which are needed to monitor stocks and enable proper planning and forecasting, may be poor or absent. There may also be a lack of adequate resources to provide the infrastructure and facilities necessary for effective distribution. For example, there may be insufficient funding to develop appropriate storage, distribution, and inventory management systems; and there may be a scarcity of warehousing and transport infrastructure, leading to limited geographic reach of the distribution network.

Management and organizational challenges at the local level include poor planning, poor security, and "leakage" of products. Limited provision of cold-chain facilities may lead to inadequate supplies of heat-sensitive medicines such as vaccines and sera. Problems with distribution chains may lead to medicines having a short shelf life when they finally reach their destination, to their reaching their expiry dates and having to be destroyed, and to stock-outs.

Stock-outs have serious implications for the implementation of programs, as they undermine the ability of health workers to follow treatment protocols. When a first-line medicine is not available, patients are either sent away to buy the medicine in the private sector, which leads to a progressive lack of confidence in the public health system, or treatment is adapted based on the availability of other medicines with a similar effect. Stock-outs can be prevented through efficient procurement and financial planning, accurate forecasting, and good inventory management.

Several initiatives have been undertaken to help prevent stock-outs in health facilities. The Stop Stock-outs Project is a consortium of civil society organizations in South Africa that monitors availability of essential medicines and vaccines in public health facilities. Its annual *National Stop Stockouts Surveys* are used to raise public awareness of the consequences of medicine stock-outs, and to provide information about medicine availability to the member countries' ministries of health to aid in their planning (Stop Stockouts Campaign, 2015). A number of countries use mobile phones and broadband access to track commodities and improve access to medical resources and health care to the rural poor (Zurovac, Talisuna, & Snow, 2012).

Storage and distribution costs represent a significant component of the health budget, and adequate resources must be allocated to these functions to ensure that the medicines get to their point of use. In fact, transport costs can be more than the value of the medicines themselves, especially when pharmaceuticals must be moved to remote, sparsely populated areas. At least five alternative ways have traditionally existed for supplying medicines and medical supplies to governmental and nongovernmental health services:

- *Central medical stores (CMS):* In the traditional public-sector pharmaceutical supply system, medicines are procured and distributed by a

EXHIBIT 16-2 Commonly Encountered Distribution Challenges

- Vertical and fragmented supply systems
- Insufficient funding for storage, distribution, and inventory management systems
- Poor planning
- Poor security and leakage
- Short shelf life of medicine
- Limited geographic reach
- Insufficient warehousing and transport
- Poor information management systems to monitor stocks for planning and forecasting
- Stock-outs and expiry
- Poor cold chain
- Vertical and parallel distribution chains

centralized government unit. The state is the owner and funder (using national or donor resources) for the entire supply system. This approach has been employed in a number of countries in Africa, Asia, Europe, and Latin America. In Ghana, the CMS distributes products to 10 regional medical stores, which in turn distribute those supplies to roughly 2,200 service delivery points. In Zambia, the CMS distributes products to 72 district medical stores, which then distribute the supplies to roughly 1,500 service delivery points.

- *Autonomous supply agency:* In this alternative to the CMS system, distribution is managed by an autonomous or semi-autonomous pharmaceutical supply agency. This strategy has been used in Tanzania and Benin.
- *Direct delivery system:* In this decentralized, non-CMS approach, medicines are delivered directly by suppliers to districts and major facilities. The government pharmaceutical procurement office selects the supplier and establishes the price for each item, but the government does not store and distribute medicines. This approach has been used in the Caribbean.
- *Primary distributor (or prime vendor) system:* In this non-CMS system, the government pharmaceutical procurement office establishes a contract with one or more primary distributors, as well as separate contracts with pharmaceutical suppliers. The contracted primary distributor receives medicines from the suppliers and stores, and distributes them to districts and major facilities. Thus, private channels are used to supply publicly funded medicines to government-operated health facilities. This strategy is used in the United States.
- *Primarily or fully private supply:* In some countries, medicines for public-sector patients are provided by private pharmacies in or near government health facilities. With this approach, measures are required to ensure financial support that facilitates access for poor and other vulnerable populations—for example, through insurance reimbursement. Private-sector supply is used in Canada and Australia.

These systems vary considerably with respect to the roles of government and the private sector and in terms of incentives for efficiency. Mixed systems in which different categories of pharmaceuticals are supplied through different mechanisms are increasingly needed to respond to expanding demands for products and services (Management Sciences for Health, 2012). The systems used are not static, but rather evolve over time; they are also subject to revision by governments in response to changing political and economic priorities (**EXHIBIT 16-3**).

EXHIBIT 16-3 A Semi-Autonomous Medical Supplies System in Sudan

In Sudan, a Central Medical Stores unit had originally been established in 1935 as a department of the Sudan Medical Services. In 1954, it became a department within the Ministry of Health and was renamed Central Medical Supplies. In 1991, its status was again changed to enable this system to exercise maximum autonomy within the framework of the government of Sudan. A Central Medical Supplies Public Corporation (CMS) was established as a semi-autonomous body to facilitate the selection, procurement, storage, and distribution of medical supplies for the public sector. Central Medical Stores then became the national center for procurement, storage, and distribution of medicines.

Prior to 1991, all medicines and medical supplies procured and warehoused by CMS were given free of charge to all public health facilities in Sudan. However, since 1992 and since it has become a semi-autonomous organization, CMS has operated on a cost-recovery basis. This change was introduced to bring it into line with the cost-recovery policy implemented by the government at all health facilities in the Sudanese public health sector.

Unfortunately, this policy resulted in reduced access to medicines for those patients who could not afford to pay. In 1996, the government announced a project to offer free treatment at hospital emergency units. This policy applied only to those needing emergency treatment in a hospital casualty department. All patients were entitled to free services, including medicines during the first 24 hours of admission. CMS receives a special budget from the Ministry of Finance and National Economy to cover the costs of this care. Medicines are distributed monthly, once the relevant sum has been deposited in the CMS account. The budget allocated for this amounted to SDG 60 million (approximately $20 million) in 2015, accounting for more than 35% of CMS sales.

CMS now provides 73% of its sales to the public sector, 26% to private pharmacies, and 1% to others. In Sudan, there is now also public procurement of medicines through the Revolving Drug Fund, through health insurance, and through police and military medical facilities. In addition to the increased access to these services fostered by the CMS, a number of initiatives are designed to increase demand for HIV, TB, and malaria services.

Alfadl et al., 2016.

EXHIBIT 16-4 The Private-Sector Distribution Chain for Antimalarial Medicines

In many LMICs, the retail sector plays an important role in the treatment of malaria. Retailers are perceived as being more accessible and responsive than other providers, and it is often claimed that they are better able to respond to public-sector failures. A wide range of retail outlets exist, including pharmacies, drug shops, general stores, market stalls, and itinerant hawkers, and the availability of antimalarial medicines is generally high. Yet important concerns have arisen regarding the performance of retailers in providing high-quality and affordable malaria treatment. Highly effective, but more expensive medicines are rarely available, especially in more remote and less formal outlets, and antimalarial medicines are often substandard or even fake.

A number of studies have explored the retail-sector distribution chain for antimalarial medicines, with a view toward understanding how retailers decide which medicines to stock and what prices to charge. Notably, variation in the number of levels within the chain has been found across countries. Within countries, there are generally more levels in the distribution chains serving more remote outlets and in those employing less qualified staff. As would be expected, fewer suppliers operate at the top of the chain than at the bottom, and some operate at more than one level, creating considerable overlap across chain levels. Different types of wholesalers usually supply different types of retail outlets. In Tanzania and Kenya, wholesalers that supply medicines alongside other commodities serve general shops, whereas wholesalers that specialize in handling medicines serve pharmacies and drug shops. Overall, wholesale markets tend to be relatively concentrated, especially at the top of the chain, where a relatively small number of importers account for most of the antimalarial volumes sold.

Wholesale price markups play a key role in determining affordability at the retail level. Markups vary significantly across chain levels, ranging from 2% to 67% at the level supplying retailers directly, and from 8% to 99% at those supplying higher levels. Retail markups tend to be higher, ranging from 3% to 566% in pharmacies, 29% to 669% in drug shops, and 100% to 233% in general shops.

Data from Patouillard, Hanson, & Goodman, 2010.

Private-Sector Supply of Pharmaceuticals

The for-profit private sector includes manufacturers at the international and local levels, importers and wholesalers, distributors, and private prescribers, pharmacies, retail outlets, and itinerant medicine vendors. It can play four roles in support of one or more channels within the health system (Ballou-Aares et al., 2008):

- Selling medicines, supplies, and equipment (e.g., manufacturers, wholesalers).
- Selling supply chain services, such as procurement, transportation, warehousing, and information and financial services to one or more channels. Services may include supply chain design, needs quantification, or logistics management and information system design.
- Providing a distribution channel (e.g., via a pharmacy, private health clinic, or franchise network).
- Assisting in implementation of supply-chain best practices, such as scheduled delivery networks, integrated supply chains, pay-for-performance systems, or cash-to-cash cycle time management approaches (e.g., consultants, trainers, educators, change agents).

Private pharmaceutical outlets are often more numerous than their government counterparts, and play an important role in the provision of health services in many countries, particularly in rural and underserved communities (Abuya et al., 2007). Many countries also have thriving illegal markets that offer pharmaceuticals at retail and sometimes wholesale levels. Understanding the systems that supply medicines to these providers is critical for designing effective interventions to improve quality and affordability in this sector (**EXHIBIT 16-4**).

Several strategies have been developed to engage the private sector, particularly retail shop owners, in an attempt to increase access to essential medicines. These efforts include training (Marsh et al., 2004; Patouillard, Hanson, & Goodman, 2010) franchising (Chiguzo, Mugo, Wacira, Mwenda, & Njuguna, 2008), and accreditation (Management Sciences for Health, 2010), all of which have had some success. **EXHIBIT 16-5** describes how Tanzania has used accredited drug dispensing outlets to increase access to artemisinin-based combination therapies for malaria.

Mixed Systems and Interconnectedness Between the Public and Private Sectors

Most countries have mixed public and private financing and delivery of care, such that the public sector may supply public health facilities, while the private sector supplies the public sector, private hospitals, clinics, and retail outlets. In many instances, there is a great deal of interaction between the two sectors, with both being

EXHIBIT 16-5 Increasing Access to Malaria Treatment in Tanzania Through Accredited Drug Dispensing Outlets

When Tanzania shifted its recommended first-line treatment to an artemisinin-based combination therapy (ACT), the National Malaria Control Program (NMCP) recognized that a public sector–focused program alone would not benefit the majority of Tanzanians who treat malaria at home. Nevertheless, providing malaria therapy through the private sector in this country has its challenges: Registered pharmacies are scarce and charge unaffordably high prices for ACT, while retail medicine sellers are untrained and largely unregulated.

With funding from Management Sciences for Health's Rational Pharmaceutical Management Plus Program, the Tanzania Food and Drug Authority and NMCP designed a plan to make subsidized ACTs available through accredited drug dispensing outlets (ADDOs) in the Morogoro and Ruvuma regions. The ADDO program is an innovative public–private initiative that uses accreditation, inspection, incentives, and training to increase access to essential medicines. Its design included allowing prescription ACTs to be sold in ADDOs, training ADDO dispensers on the new treatment, and collaborating with stakeholders to determine the price at which ADDOs would sell subsidized ACTs.

The pilot program covered approximately 2.9 million people, or approximately 8.4% of the population of Tanzania. In the Morogoro and Ruvuma regions, 1,363 ADDO dispensers from 650 ADDOs have been accredited and trained and are now identifying and treating uncomplicated malaria with ACTs. In comparison, only 600 pharmacists exist in the public and private sectors in the entire country.

One year after the pilot program was initiated, a review of records from 448 ADDOs indicated that the percentage of ADDOs that dispensed at least one course of ACT rose from 26.2% during July–September 2007 to 72.6% during April–June 2008, and that ADDOs dispensed more than 300,000 treatments over the same period. For the five districts in Ruvuma that reported data for July–September 2008, this percentage increased to 81.5%. These results illustrate the enormous potential of ADDOs to improve community access to recommended malaria medicines through the private sector.

Data from Rutta, 2013; Management Sciences for Health, 2010.

involved in procurement, importation, financing, service delivery networks, and market and product information. In many countries, illicit flows may also occur within and between the sectors due to leakage from public sector facilities, diversions from the port, and donations by nongovernmental organizations (NGOs).

FIGURE 16-5 illustrates the potential flow of medicines in the public and private sectors, highlighting interactions among the various sectors and stakeholders. Understanding these systems, and the connections between the public and private sectors, is critical to intervening to improve access and quality.

The Effect of Global Health Initiatives on the Pharmaceutical Distribution Chain

Since 2000, the emergence of several major disease-specific global health initiatives (GHIs) has contributed to an unprecedented increase in the resources available for programs targeting some diseases. Establishment of GHIs has led to large increases in the supply of and demand for medicines, vaccines, insecticide-treated bed nets, and diagnostic and laboratory materials, and has been associated with improvements in the quality, availability, and affordability of many of these commodities, particularly in relation to vaccines and antiretroviral medicines for HIV/AIDS (Chauveau, Meiners, Luchini, & Moatti, 2008; Hagmann, 2001; Samb et al., 2009; Sharma, 2003).

In some instances, national governments and GHIs have worked together to strengthen national procurement and distribution networks (Stillman & Bennett, 2005). Nevertheless, although alignment with country procurement systems is included as one of the indicators of aid effectiveness in both the 2005 *Paris Declaration on Aid Effectiveness* and the 2008 *Accra Agenda for Action*, such synergy has not always been realized (Sjöstedt, 2013). In other instances, GHIs have duplicated or displaced country supply chains, thereby creating parallel systems (Laing & McGoldrick, 2000).

Last Mile Distribution

Every supply chain has a literal last mile, the crucial final leg to reach the customer. In public health there is also the figurative "last mile," referring to reaching the most isolated segments of the population. Most of those persons who lack access to medicines and vaccines are either the most remote, living far from paved roads and affordable transport, or the most economically impoverished.

Several initiatives have been developed to address this last mile. Project Last Mile (PLM) is a public–private partnership that transfers Coca-Cola's supply-chain management and marketing expertise to

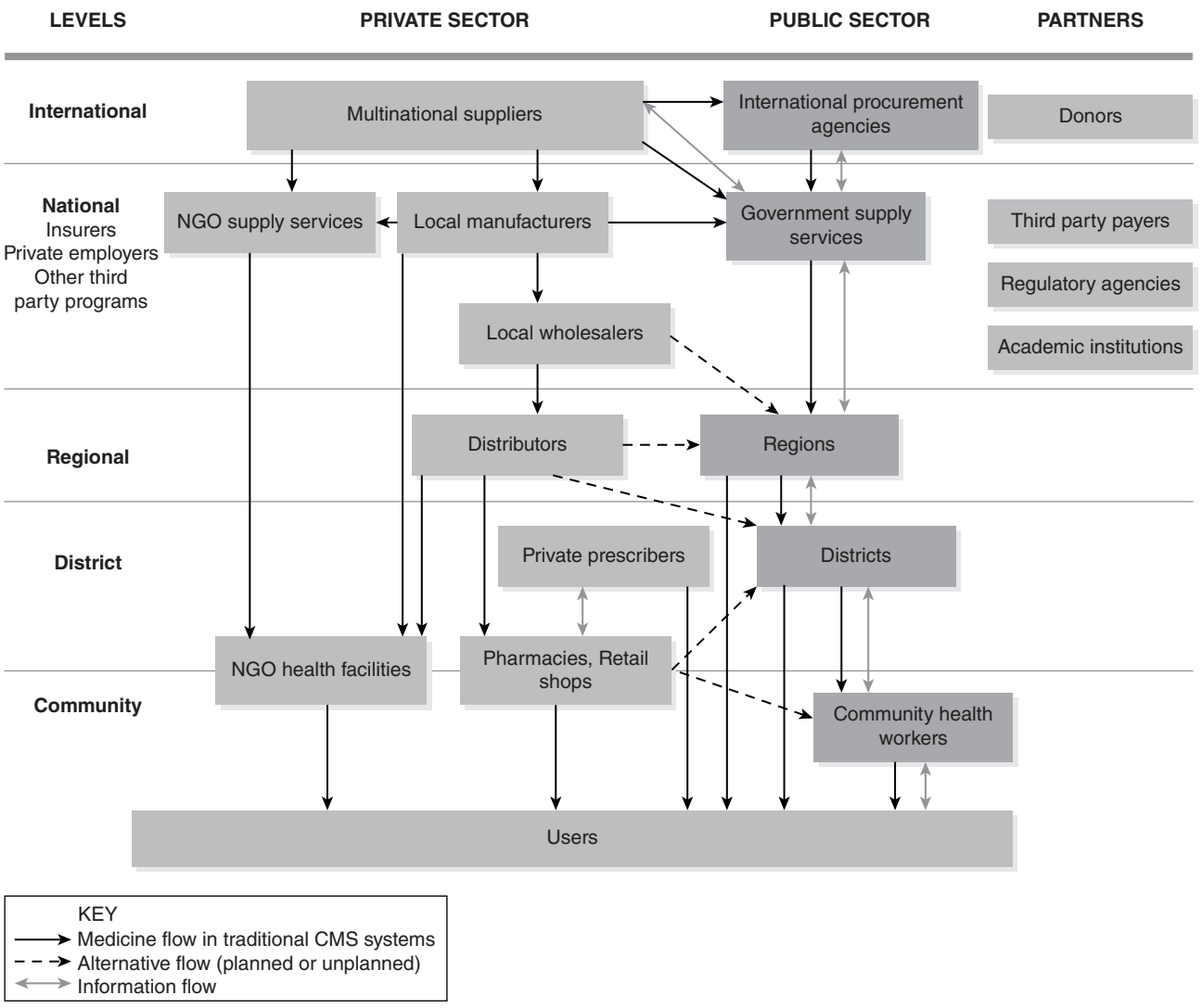


FIGURE 16-5 Flow of medicines in the public and private sectors.

Modified from Management Sciences for Health. (2012). *MDS-3: Managing access to medicines and technologies*. Arlington, VA: Management Sciences for Health.

help governments and health systems in seven countries deliver life-saving medicines and supplies to hard-to-reach communities across Africa. The initiative began in 2010 with a mission to transform the delivery of medical supplies in Tanzania, and expanded in 2014 with more than $21 million in additional investment and an ambition to help support 10 African countries by 2020. PLM teams have since worked with Ministries of Health in Ghana, Mozambique, Nigeria, and South Africa to improve the availability of essential medicines (Village Reach, 2012).

Affordability

Affordability is a key factor affecting access to pharmaceuticals. It relates the price that a patient must pay for a product to the amount that the person is able to pay. While medicines are sometimes provided for free at public health facilities, patients frequently seek care outside the public sector—for example, in private drug shops, clinics, or pharmacies. Even where medicines are officially free in the public sector, stock-outs and informal charges may mean that patients still incur payments to obtain them. Medicines are sometimes included in insurance benefit packages, but they may nonetheless require a copayment. Thus, in many cases, patients are required to make out-of-pocket payments to purchase medicines when they are ill.

Affordability of medicines is commonly measured in three ways (Niëns et al., 2012; WHO & Health Action International [HAI], 2008):

- The "catastrophic payment" method is based on the ratio of medicine expenditure to total household resources.
- The impoverishment method examines a household's residual income after paying for a good.

TABLE 16-4 Affordability of One Salbutamol Inhaler (100 mcg/dose) Purchased from the Private Sector in Selected Countries, 2013

Country (Date of Survey)	Originator Brand	Lowest-Priced Generic Brand
Burundi (August 2013)	4.2 days	1.4 days
Egypt (September 2013)	0.3 day	0.3 day
Sudan (February 2013)	2.5 days	1.4 days
Tajikistan (June 2013)	3.3 days	3.1 days

Data from Health Action International, 2017.

- The WHO/HAI method expresses affordability as the number of days' wages (e.g., of the lowest-paid, unskilled government worker in a given setting) required to purchase a treatment course.

To illustrate, the affordability of salbutamol inhaler, a medicine commonly used to treat asthma, purchased in the private sector is outlined in **TABLE 16-4**. As this example demonstrates, the significant variation in affordability for the same treatment is pronounced not only across countries, but also within countries between originator and generic brands (HAI, 2017).

This section examines some of the factors that have impacted the affordability of pharmaceuticals in LMICs, and explores the strategies that have been devised both to reduce the prices of pharmaceuticals and to increase financing for them.

Pharmaceutical Prices and Affordability

In LMICs, governments fund less than 35% of total pharmaceutical spending, compared with more than 60% of such spending in high-income countries (Lu et al., 2011). While the affordability of pharmaceuticals is a key issue in all countries, it is of particular importance in LMICs, where less-developed health systems and lower wages generally result in a greater share of personal income being spent on health, thereby exposing households to the risk of catastrophic health expenditures (Lu et al., 2011). This situation is compounded by the chronically low public-sector availability of pharmaceuticals, which forces many patients in these countries to purchase a high proportion of their pharmaceuticals from the private sector, often at greater cost (Cameron, Ewen, Auton, & Abegunde, 2011).

Price is the most immediate factor that affects the affordability of pharmaceuticals. In the public sector, the efficiency of the procurement process is an important determinant of both the affordability and the availability of public-sector pharmaceuticals. A 2009 study that examined the government procurement prices of 15 commonly purchased medicines in 36 countries revealed that, although the median government procurement prices for those medicines were only 1.11 times the international reference prices from open international procurements for generic products (indicating a relatively good level of procurement efficiency, on average), the range of procurement prices varied widely, from 0.1 to more than 5 times the international reference prices, indicating that the procurement process often did not secure the best value (Cameron, Ewen, Ross-Degnan, Ball, & Laing, 2009).

To understand why medicine prices vary or may become unreasonably high, their many components must be elucidated. Beyond the upstream influences operating at the manufacturer level described earlier, the "downstream" elements differ by country, by sector of the health system, and by type of pharmaceutical, but generally include the manufacturer's selling price; insurance, freight, port charges, inspection charges, and duties for imported products; the markups added by importers, wholesalers, and retail distributors; various taxes; and fees associated with dispensing.

Strategies to Improve Affordability by Reducing Pharmaceutical Prices

Several international panels have recommended a range of policies in an attempt to deliver more affordable pharmaceutical prices. Note, however, that the evidence for these recommendations comes largely from high-income settings, as research in this area has until recently been lacking in lower-income countries (Nguyen, Knight, Roughead, Brooks, & Mant, 2015). **EXHIBIT 16-6** synthesizes a number of these recommendations from WHO's Guideline on Country Pharmaceutical Pricing Policies and *The Lancet*'s Commission on Essential Medicines Policies (WHO, 2015c; Wirtz et al., 2017). Along with these recommendations, the panels agree on a number of key principles for consideration when choosing to act on high pharmaceutical prices. Generally, countries should select several complementary policy options that are appropriate to their particular health system, pharmaceutical market composition, national priorities, available resources, and human rights considerations.

EXHIBIT 16-6 Recommended Policies to Reduce Prices of Essential Medicines in LMICs

Procurement Interventions

- Pooled procurement using limited competitive bidding/tendering
- Pooled procurement or use of monopsony power (e.g., where large public or social health insurers operate), with price negotiation based on volumes procured or on inclusion in a reimbursement list

Pro-Generic Policies

- Use of international nonproprietary names when prescribing, rather than brand names
- Mandatory generic substitution or enablement of generic substitution by pharmacists and other dispensers
- Measures to enable early and prompt entry of generics

Pricing Interventions

- Reduction or removal of duties, tariffs, import taxes, or sales taxes, particularly for essential medicines
- Internal/within-country reference pricing by chemical entity, pharmacologic class, or indication
- External reference pricing for procurement and price benchmarking
- Regulation of wholesaler and retailer markups, with preference for regressive markups (i.e., lower markup for higher-priced products), rather than fixed-percentage markups
- Regulation of professional fees
- Regulation of annual factory-gate price increases
- Patent-related interventions such as encouragement of voluntary licensing and patent pools

Trade-Related Aspects of Intellectual Property Rights Flexibilities

- For LMICs, postpone the granting or enforcement of medicine patents or other market exclusivity rights that would prevent the registration of a generic product
- Parallel importation
- Compulsory licensing
- Government use licensing
- Application of strict patentability criteria

Modified from WHO, 2015c, Wirtz et al., 2017b.

Two of the price-reduction mechanisms in Exhibit 16-6 have more recently been debated: The reduction of taxes and tariffs and the use of the public health safeguards incorporated into the WTO's Agreement on Trade-Related Intellectual Property Rights (see the *International Trade and Health* chapter for greater detail on TRIPS).

Tariffs and taxation are important tools wielded by governments to ensure economic growth and national development. For medicines, it has been argued that measures such as import duties foster the growth of the local pharmaceutical industry by protecting it from international competition. However, tariffs and taxes levied on medicines are widely regarded as regressive, disproportionately affecting the poor, who are more likely to experience ill health and pay a larger share of their income for treatment than their richer counterparts. Recent studies have concluded that, although tariffs do contribute to the final price the patient pays for medicines in many countries, their overall share of the final price has been falling over time. Domestic taxes such as value-added tax (VAT) or sales tax are now more likely to be a key driver of medicine price, with these tax rates in LMICs ranging from 2.9% to 34% (Creese, 2011). Although government revenues raised by taxing medicines have been shown to be small (Olcay & Laing, 2005), advocates for removing or reducing these taxes have recognized the challenge of lobbying effectively for a loss of government revenue. "Healthy taxation," which sanctions the taxing of risky and unhealthy behavior and products (e.g., cigarettes) rather than targeting health-promoting ones (e.g., essential medicines), has become a potent argument to reform tax regimes to improve medicine affordability (Creese, 2011).

Although the TRIPS flexibilities designed to protect public health have not been widely used, they have been most successfully applied to supply generic antiretroviral medicines in a number of LMICs; less frequently, they have been used to secure access to affordable treatments for cancer, arthritis, and avian flu ('t Hoen, 2016). Even the simple threat of issuing

compulsory licenses by several LMIC governments has proved useful in the past to negotiate lower prices for patented medicines for HIV and cancer (Coriat, 2008; 't Hoen, 2016). Nevertheless, these flexibilities are under continual threat from the inclusion of so-called TRIPS-plus measures in bilateral and regional trade agreements between powerful trading nations and LMICs. These agreements include intellectual property rights obligations that go well beyond the TRIPS minimum standards and supersede them in these jurisdictions (Kerry & Lee, 2007).

External Financing of Pharmaceuticals

Chronic underfunding of their broader healthcare infrastructure has led some countries to rely on external funding to support access to medicines. For example, by 2016, 45% of the $30 billion disbursed by the Global Fund went directly to ministries of health, much of which was spent on pharmaceuticals and strengthening the distribution chain; another 30% of Global Fund disbursements went to civil society organizations, many of which run pharmaceutical-related programs that complement those found in the public sector (Global Fund, 2017b). Other prominent sources of external funding for pharmaceuticals include the GAVI Alliance, which has been co-financing country procurement of vaccines for immunization programs since 2007 (GAVI, 2016), and the Global Drug Facility, which provides financial and technical support to countries to procure TB treatments and other commodities (Arinaminpathy, Cordier-Lassalle, Lunte, & Dye, 2015; Stop TB Partnership & WHO, 2014).

Safe and Effective Use of Medicines

Even when the right medicines are in the right place and are affordable, many challenges remain in ensuring safe and effective therapy. Are the medicines of appropriate quality? Are they being used appropriately? What must be done to ensure their continued safety? These are the questions to which we now turn.

Rational Use of Medicines

The concept of rational medicine use has been adopted by many countries. According to WHO (2010d), the rational use of medicines requires that "patients receive medications appropriate to their clinical needs, in doses that meet their own individual requirements, for an adequate period of time, and at the lowest cost to them and their community." This definition is expressed in medical and financial terms, but patients have their own rationale for taking medicines. Thus, what may be seen as irrational from a medical perspective—for example, interrupting a treatment course of antimalarials when symptoms recede—may be perceived as entirely rational from the consumer's point of view.

Nonrational use of medicines takes many forms and can occur at the prescribing, dispensing, and patient administration stages. Studies in both high-income countries and LMICs have described numerous examples of nonrational medicine use (**EXHIBIT 16-7**).

The nonrational use of medicine carries enormous costs and impacts. In addition to inefficient use of limited resources, such use may lead to adverse clinical consequences and unnecessary suffering in patients. Wasting resources in this way may lead to increased costs later and the need to use more expensive medicines. In addition, nonrational use can lead to increased morbidity and mortality, along with increased risks of adverse drug reactions and the emergence of antibacterial resistance (discussed later in the chapter).

Medicine Use Behavior

The use of medicines has to be viewed in both biomedical and social contexts. A number of actors (principally prescribers, dispensers, and patients) are

EXHIBIT 16-7 Examples of Nonrational Medicine Use

- Polypharmacy (multiple prescriptions or over-prescription)
- Use of medicines that are not related to diagnosis
- Unnecessarily costly medicines
- Inappropriate prescription and use of antibiotics
- Indiscriminate use of injections
- Irrational self-medication with underdosing and overdosing
- Incorrect administration, dosages, timing, formulation, or duration
- The use of medicines when no medicine therapy is indicated
- Failure to prescribe available, safe, and effective medicines
- Poor adherence to tuberculosis treatment
- Underuse of effective medicines for hypertension and depression
- Hospital medicine use problems such as antibiotic misuse for surgical prophylaxis

Wiedenmayer, 2004.

engaged in this process, which involves a number of steps:

1. Diagnosis: Identification of what is wrong.
2. Therapy: Development of a therapeutic objective and plan.
3. Prescribing: Information, instructions, warnings, and prescription writing.
4. Dispensing: Supply, advice, counseling, instructions, and warnings.
5. Adherence by the patient: Understanding of the therapy, patient responsibility, and valuing of treatment.

The factors that determine patterns of individual medicine use are complex. The medicine use system in which they occur is part of a larger healthcare system, which is itself shaped by the social, cultural, economic, and political contexts of the country concerned. Furthermore, both prescriber and patient bring a host of beliefs and motivations to their interaction. For example, the use of medicines may be discouraged as a result of social stigma associated with enrollment in an HIV medicine program, as this signals the patient's HIV/AIDS status (Connelly & Rosen, 2006). Failure to understand the local cultural context is a common cause of inappropriate medicine use (Wiedenmayer, 2004). For example, in some countries, medicines that are red in color are thought to be good for the blood. Injections and suppositories may be more acceptable in some countries than others based on local beliefs. Moreover, there is a widespread belief that antibiotics are effective against viruses, as well as lack of trust in the safety of vaccines (Vaccine Confidence Project, 2015).

Strategies to Improve Medicine Use

Improving the use of medicines by health workers and the general public is crucial to reducing morbidity and mortality from both communicable and noncommunicable diseases, preventing and minimizing drug resistance, and controlling expenditures. Therapeutically sound and cost-effective use of medicines by health professionals and consumers needs to be achieved at all levels of the health system, in both the public and private sectors. A sound rational-use program has three elements that are formulated to reflect the main responsibilities of a national essential medicine program (**EXHIBIT 16-8**).

WHO and the International Network for Rational Use of Drugs (INRUD) have developed extensive resources and literature on the promotion of rational medicine use. INRUD's *Drug Use Bibliography* is an annotated list of published and unpublished articles, books, reports, and other documents related to medicine use, with a special focus on LMICs (INRUD, 2017). A framework linking strategies to improve use and other elements of the process has been developed (**FIGURE 16-6**).

EXHIBIT 16-8 Elements of Rational Medicine Use Programs

- *Rational use of medicine strategy and monitoring:* Advocating rational medicine use, identifying and promoting successful strategies, and securing responsible medicine promotion.
- *Rational use of medicines by health professionals:* Working with countries to develop and update treatment guidelines, national essential medicine lists, and formularies, and supporting training programs on rational use of medicines.
- *Rational use of medicines by consumers:* Supporting creation of effective systems of medicine information, and empowering consumers to take responsible decisions about their treatment.

Reprinted from Essential Medicines and Health Products. (2010). Activities, world health organization, rational use of medicines. http://www.who.int/medicines/areas/rational_use/rud_activities/en/

Effectiveness of Interventions to Improve Medicine Use

In general, medicine use can best be improved through interventions targeted at specific problems with an identifiable audience, or by a system change such as preventing the unnecessary prescribing of expensive branded medicines through use of nationally accepted clinical practice guidelines. The most effective medicine use intervention is one that is participatory, interactive, problem based, and focused. It is usually more efficient and effective to combine a number of strategies into a multifaceted intervention to improve medicine use.

Some strategies have proved more successful than others. A review by Laing et al. (2001) identified 10 key factors that have been shown to make a difference (**EXHIBIT 16-9**).

Quality of Medicines

A prerequisite for the rational use of medicines is that they should be of good quality, both at the time of manufacture and at the time of consumption by the patient. Unfortunately, a wide variety of potential hazards may emerge between the manufacture and end use of these products. Substandard and counterfeit medicines are major problems on a worldwide basis, as are smuggling and the illegal importation of

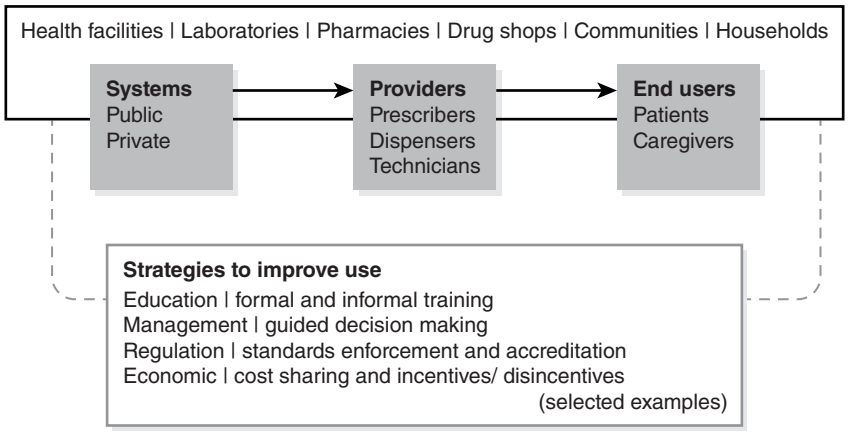


FIGURE 16-6 Framework for improving medicine use.

From Management Sciences for Health. (2012). *MDS-3: Managing access to medicines and technologies.* Arlington, VA: Management Sciences for Health. http://apps.who.int/medicinedocs/documents/s19577en/s19577en.pdf

EXHIBIT 16-9 Recommendations to Improve Use of Medicines in LMICs

1. Establish procedures for developing, disseminating, utilizing, and revising national (or hospital-specific) standard treatment deadlines.
2. Establish procedures for developing and revising an essential medicine list (or hospital formulary) based on treatments of choice.
3. Request hospitals to establish representative Pharmacy and Therapeutics Committees with defined responsibilities for monitoring and promoting quality use of medicines.
4. Implement problem-based training in pharmacotherapy in undergraduate medical and paramedical education based on national standard treatment guidelines.
5. Encourage development of targeted, problem-based, in-service educational programs by professional societies, universities, and the ministry of health, and require regular continuing education for licensure of health professionals.
6. Stimulate an interactive group process among health providers or consumers to review and apply information about appropriate use of medicines.
7. Train pharmacists and medicine sellers to be active members of the healthcare team and to offer useful advice to consumers about health and medicines.
8. Encourage active involvement by consumer organizations in public education about medicines, and devote government resources to support these efforts.
9. Develop a strategic approach to improve prescribing in the private sector through appropriate regulation and long-term collaborations with professional associations.
10. Establish systems to monitor key pharmaceutical indicators routinely so as to track the effects of health-sector reform and regulatory changes.

Data from Laing et al., 2001.

medicines. These activities present major challenges to national authorities, and seriously undermine initiatives aimed at rational medicine use.

Defining Quality of Medicines

Two concepts that need to be distinguished are quality assurance (QA) and quality control (QC). QA is the whole process of assuring quality from manufacturer to end user. By contrast, QC involves testing key characteristics of the final product, including the identity and potency of the active ingredient, the content range (with the acceptable range usually set at 95% to 110%), purity (involving checks to exclude contaminating substances or microorganisms), uniformity (in terms of color, shape, and size), bioavailability (the rate and extent of absorption of a medicine into the body), and stability (adequate shelf life and expiry). Thus, QA in pharmaceutical supply is not the same as QC in manufacturing.

Detailed descriptions of medicine characteristics and of analytical techniques to verify them are laid down in national pharmacopoeias such as the *United States Pharmacopoeia* (USP) and the *British*

Pharmacopoeia (BP). WHO publishes an *International Pharmacopoeia* (Ph.Int.), which from 1975 has focused on WHO's Model List of Essential Medicines (WHO, 2017c). More recently, it has emphasized medicines of major public health importance and those recommended by specific WHO disease programs, such as medicines to treat malaria, TB, and HIV/AIDS, as well as medicines for children. Priority is also given to medicines evaluated by WHO's Medicines Prequalification Program.

Two other concepts that need to be distinguished are substandard and counterfeit medicines. Counterfeit medicines are addressed later in this section. WHO (2010e) defines substandard medicines (also called out-of-specification [OOS] products) as "genuine medicines produced by manufacturers, authorized by the National Medicines Regulatory Authority (NMRA), which do not meet quality specifications set for them by national standards." Normally, medicines produced by manufacturers have to comply with a set of quality specifications, and NMRAs review products against these criteria before they are authorized for marketing. The quality of medicines reaching the patient can be affected by the manufacturing process, packaging, transportation and storage conditions, handling, and other factors.

These influences can be cumulative. Loss of activity due to the instability of the drug itself is unusual; instead, poor initial quality of the medicine is the more serious problem. Substandard medicines include those with a loss of potency due to poor bioavailability, expiry, or poor storage conditions; those with too-low or too-high concentration of active ingredients due to manufacturing or compounding error and counterfeiting; those that have degraded into toxic substances; those producing adverse reactions; and injectables, creams, syrups, and eye drops that have been contaminated with bacteria or fungi. Substandard medicines remain a serious problem in resource-poor settings (Caudron et al., 2008).

Quality Assurance Programs

The three main activities involved in assuring the quality of medicines are registration of medicines with the NMRA, inspection of manufacturing plants to ensure compliance with Good Pharmaceutical Manufacturing Practice (GMP), and testing of medicines based on product types and characteristics. GMP covers all aspects of production, with its guidelines focusing on personnel, facilities, equipment, sanitation, raw materials, manufacturing processes, labeling and packaging, quality control systems, self-inspection, distribution, documents and records, and complaints and adverse medicine reaction systems.

Quality assurance programs must include training and supervision of staff involved at all stages of the process. Effective information systems are essential for following up and documenting quality problems. Although the foundation of QA consists of regulations and standards, it is the people who enforce the regulations or work to comply with the standards who make the difference between QA and the lack of it (WHO, 2007). WHO's prequalification program is another mechanism for ensuring pharmaceutical quality, including the quality of new products (WHO, 2017e).

Role of Regulatory Agencies

An important determinant of both QA and medicine quality is effective medicine regulation. When enforcement is weak or lacking, pharmaceutical manufacturers, importers, and distributors may ignore regulatory requirements; in turn, the quality, safety, and efficacy of both imported and locally manufactured medicines may be compromised. In most countries, national legislation is often not equipped to deal with the extremely serious health consequences of counterfeit medical products, and penalties for counterfeiters are too light to act as real deterrents (International Medical Products Anti-Counterfeiting Taskforce [IMPACT], 2011).

Both NMRAs and WHO play an important part in ensuring the quality of medicines. Regulatory decisions on medicines have been compiled by the United Nations in its *Consolidated List of Products Whose Consumption and/or Sale Have Been Banned, Withdrawn, Severely Restricted or Not Approved by Governments* (United Nations, 2005; WHO, 2010c)

Counterfeit Medicines

A separate and distinct issue from substandard medicines is the proliferation of counterfeit medicines. The problem of counterfeit medicines was first addressed at the international level in 1985 at the Conference of Experts on the Rational Use of Drugs in Nairobi. A number of initiatives on this front have since been undertaken. In 2006, WHO launched the International Medical Products Anti-Counterfeiting Taskforce as part of the Declaration of Rome. Today, this body remains the main conduit for WHO's work against counterfeiting operations.

WHO has defined a counterfeit medicine as "one which is deliberately and fraudulently mislabeled with respect to identity and/or source. Counterfeiting can apply to both branded and generic products, and counterfeit products may include products with the correct

ingredients or with the wrong ingredients, without active ingredients, with insufficient active ingredients or with fake packaging" (IMPACT, 2006). This definition has not been universally accepted, however, which has made it difficult to exchange information between countries and to measure the size of the problem.

The cost of counterfeit medicines is not measured purely in financial terms (Burki, 2010), because these products constitute a serious health risk for the world's population. Trade in counterfeit medicines is a serious crime. In recent years, increasing international trade of pharmaceuticals and sales via the Internet have facilitated the entry of counterfeit products into the distribution chain. In areas in some LMICs, more than 30% of medicines sold are counterfeit. Counterfeiting is increasing, and it is estimated that more than 70% of this practice occurs in LMICs (IMPACT, 2006).

Medicines are attractive to counterfeiters for a number of reasons. They are high-value items for which demand exceeds supply, and their ingredient costs can be low. Moreover, production can often be undertaken using simple equipment, and overhead costs are minimal. Finally, patients are rarely in a position to judge the quality of a medicine. The counterfeiting of medicines tends to thrive where medicines distribution systems are weak or uncoordinated.

The majority of counterfeit cases involve tablets and capsules. Antibiotics, hormones, analgesics, antihistamines, and steroids account for almost 60% of reported cases of counterfeit medicines. In approximately one-third of cases, products contain no active ingredient; in 20% of cases, they contain an incorrect quantity of that ingredient. In a further 20% of cases, the medicines contain the wrong ingredient, and in almost 10% of cases, they contain high levels of impurities and contaminants (Deisingh, 2005). Counterfeit medicines pose a serious health risk to patients and are a major cause of waste of both public and private funds.

A wide range of methods are now available for the detection of counterfeit medicines (WHO, 2010a). Although some require the use of expensive equipment, simple chemical approaches have also been developed. An initiative of pharmaceutical companies in Germany, the Global Pharma Health Fund (GPHF) has developed an inexpensive field test kit or "Minilab" with rapid, simple-to-use tests for medicine quality verification and counterfeit medicine detection. More than 800 Minilabs have since been supplied across 95 countries; Cambodia, Laos, Vietnam, Madagascar, Nigeria, and Tanzania have also adopted this technology for postmarketing antimalarial drug quality monitoring (GPHF, 2017).

EXHIBIT 16-10 WHO-Recommended Steps to Curb Medicines Counterfeiting

- National laws should regulate the manufacture, trade, distribution, and sale of medicines effectively, with severe penalties being imposed for manufacturing, supplying, or selling counterfeit medicines.
- The NMRAs responsible for the registration and inspection of locally manufactured and imported medicines should be strengthened.
- NMRAs should develop standard operating procedures and guidelines for the inspection of suspected counterfeit medicines, and should initiate widespread screening tests for the detection of such products.
- Adequate training and powers of enforcement against counterfeit medicines operations should be given to personnel by NMRAs, the judiciary, customs personnel, and police.
- Partnerships should be established between health professionals, importers, industry, and local authorities to combat counterfeit medicines.
- Countries should systematically use WHO's Certificate Scheme on the Quality of Pharmaceutical Products Moving in International Commerce. Countries in the same region should work toward harmonization of their marketing authorization procedures.
- Countries should exchange experience and expertise in areas related to quality control, medicine detection, and enforcement.

Modified from World Health Organization (WHO). (2003). Counterfeit drugs: Action sheet. http://www.wpro.who.int/mediacentre/factsheets/fs_200311_Counterfeit_drugs/en/

WHO has published guidelines for the development of measures to combat counterfeit medicines. **EXHIBIT 16-10** lists the steps that should be taken to curb medicines counterfeiting.

Many countries are stepping up their offensives against the counterfeiting problem, but the results have so far been mixed. In Nigeria, intensive action has led to the proportion of fake or adulterated medicines in circulation being reduced from 70% to 16%; in contrast, in India, the share of spurious medicines has grown from 10% to 20% of the market (IMPACT, 2006). Global action will be necessary if this problem is to be contained, and suggestions have been made as to how this goal might be achieved (Attaran et al., 2012). In Europe, a concerted effort is being made to tackle the problem by means of a Falsified Medicines Directive (**EXHIBIT 16-11**).

EXHIBIT 16-11 Europe and the Falsified Medicines Directive

For several years the European Union has been concerned about the threat of counterfeit or "falsified" medicines to public health and safety in its member states. In 2011, the European Commission (2017) started work to amend Directive 2001/83/ED to address these concerns under the Falsified Medicines Directive (FMD). The FMD is designed to protect patient safety by minimizing the chances of counterfeit medicines entering the established medicines supply chain across Europe. It will enable manufacturers, wholesalers, distributors, and all parties who supply medicines to patients to verify the authenticity of a medicinal product, identify individual packs, and check whether the outer packaging of medicines has been tampered with. Manufacturers will be obliged to apply a range of safety features to every pack, including a tamper-proof security seal and a two-dimensional barcode, which will be authenticated before dispensing. Data will be managed by the European Medicines Verification Organization, supported by the Medicines Verification Organizations of EU member states. Manufacturers will check packs into the database, and only those that can be authenticated by scanning the barcode will be dispensed.

Data from Robinson, J., 2016; European Commission, 2017.

Medicine Safety Issues

The availability of medicines of appropriate quality presents its own problems, as the use of any medicine brings with it the possibility of unintended consequences. Any injury occurring at the time a medicine is used, whether or not it is identified as a cause of the injury, is described as an adverse drug event (ADE) (International Council for Harmonization of Technical Requirements, 1996). An adverse drug reaction (ADR) is a special type of ADE in which a causative relationship can be shown. Both ADRs and ADEs are different from accidental or deliberate excessive dosage and administration errors. They also differ from "side effects," as this term also applies to effects that may be beneficial (Nebeker, Barach, & Samore, 2004). A range of approaches are used to report and analyze ADRs, including pharmacovigilance and postmarketing surveillance.

Adverse Drug Reactions

An ADR has been defined as "any response to a medicine which is noxious, unintended, and occurs at doses normally used for prophylaxis, diagnosis, or therapy" (Kanjanarat et al., 2003). ADRs are unwanted effects of a medicine, including idiosyncratic ones, which occur during its proper use. They may occur following a single dose or after prolonged administration of a medicine, or result from the combination of two or more medicines; they can be either predictable or unpredictable. ADRs may be easily confused with outcomes of disease processes and are sometimes difficult to distinguish from other causes. WHO has proposed a stepwise approach to assessing possible ADRs and has published a guide to detecting and reporting adverse drug reactions (WHO, 2017a).

Patients may respond differently to the same treatment, so their risk of ADRs likewise varies. Genetic makeup and concurrent disease, for example, can make a patient more prone to ADRs. The very old and the very young are more susceptible to such reactions. Drug–drug interactions are some of the most common causes of adverse effects. Interactions can also occur with alcohol and traditional medicines.

Pharmacovigilance

Pharmacovigilance is defined as the science and activities relating to the detection, assessment, understanding, and prevention of adverse effects or any other medicine-related problem (WHO, 2017d). In recent years, pharmacovigilance has extended its remit to include herbals, traditional and complementary medicines, blood products, biologicals, medical devices, and vaccines. The major source of new information about ADRs is spontaneous reporting.

Pharmacovigilance also extends to the regulatory measures required to prevent future ADRs and improve the benefit–risk ratio of pharmaceuticals. It supports effective decision making by NMRAs and facilitates the exchange of medicine safety information between countries and the monitoring of global trends in ADRs.

Once limited to the high-income world, pharmacovigilance has developed rapidly. Reliable systems of pharmacovigilance are now recognized as a prerequisite for the rational, safe, and cost-effective use of medicines in all countries (Mahmood, Tahir, & Haq, 2011). A regional pharmacovigilance network and training center for West Africa has been developed (Strengthening Pharmaceutical Systems Program, 2011), and similar collaborations are under development elsewhere. Progress is patchy, however: In Pakistan, pharmacovigilance is practically nonexistent despite widespread recognition of the vital contribution it can make to best patient care (Hussain & Jamshed, 2016).

Postmarketing Surveillance

Premarketing trials of medicines are undertaken on relatively small numbers of patients and lack the

power to detect important but less common ADRs, to detect ADRs that occur in a delayed fashion after the original use of the medicine, or to detect consequences associated with long-term medicine administration. Premarketing trials often do not include special populations such as pregnant women, the elderly, or children, all of whom may be at risk from unique ADRs or from an increased frequency of ADRs compared with the general population.

For these reasons, postmarketing surveillance is an important tool to detect less common, but sometimes serious ADRs. The spontaneous reporting of ADRs is the cornerstone of postmarketing surveillance. Such reports can be made to the NMRA or to the pharmaceutical company concerned. Companies may also have to submit international reports, giving details of ADRs reported in other countries, to the national authority. In certain countries, only doctors are allowed to report suspected ADRs; in others, reports from a wider group of health professionals—sometimes including pharmacists and nurses—are accepted. The frequency of reporting varies considerably between countries, but it is estimated that only 1% to 10% of all cases are actually reported to the authorities.

Postmarketing surveillance also includes the inspection of premises, documents, products, and the practice of personnel in charge to ensure compliance with laws and regulations (Chaiyakunapruk & Jones, 2016).

Many countries have established monitoring systems for early detection and prevention of medicine-related problems. Postmarketing surveillance is mainly coordinated by national pharmacovigilance centers, which collect and analyze case reports of ADRs, distinguish signals from background noise, make regulatory decisions based on strengthened signals, and alert prescribers, manufacturers, and the public to new risks of ADRs. The number of national centers participating in WHO's International Drug Monitoring Program increased from 10 in 1968 to 120 in 2014, greatly strengthening its ability to detect ADRs (WHO, 2017h).

Antimicrobials and Antimicrobial Resistance

Antimicrobial medicines, which kill or inhibit the growth of microbial organisms, are fundamental to modern medicine. They have made control and treatment of previously incurable infectious diseases possible and enabled myriad medical procedures to be performed. Antimicrobial resistance (AMR), which refers to the ability of microbes to grow in presence of antimicrobials, presents a serious and imminent threat to global public health by undermining the effectiveness of existing antimicrobials in treating infections (see also the *Infectious Diseases* and *Public Health Infrastructure* chapters).

In a narrow sense, the term "antimicrobial resistance" is used loosely to refer to antibiotic resistance. Although, scientifically speaking, both "antimicrobials" and "antimicrobial resistance" are umbrella terms that apply to all types of microbes, including bacteria, viruses, parasites, and fungi, many problems and challenges associated with antimicrobial resistance mainly concern the resistance of bacterial species. In those cases, policy makers, scientists, and the media alike sometimes use the term "antimicrobial" interchangeably with "antibiotic" or "antibacterial"; the latter actually concerns medicines that act only against bacteria (Mendelson, 2017).

AMR is driven by human activities. While AMR is a natural phenomenon that manifests the principles of Darwinian evolution, human activities can accelerate its generation and amplify its spread in a highly dynamic and interconnected ecological network (Holmes et al., 2016). Bacteria naturally colonize all types of physical environments of the socio-ecosystem, including residential, agricultural, aquacultural (fish farming), and industrial settings. They also reside in the living organisms that live within these settings, such as humans, farm animals, crops, and fish. Both extensive use of antibiotics in medicine, livestock production, and fish farming, and industrial antibiotic pollution, which directly or indirectly releases antibiotics into these environments and the living organisms within them, increase the exposure of bacteria to antibiotics, and consequently the selection pressure for antibiotic resistance. The resultant resistant genetic elements and bacteria are liable to spread across geographies and systems.

Consequences of Antibiotic Resistance

Antibiotic resistance across the globe has reached an alarming level. High rates of resistance to some of the most advanced antibiotics (such as third-generation cephalosporins and carbapenems) and to some of the most widely used antibiotics (such as fluoroquinolones) have been observed in bacteria that cause common healthcare-associated and community-acquired infections in all WHO regions (WHO, 2014). In healthcare settings, a mix of gram-positive and gram-negative pathogens identified by the acronym ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas*

aeruginosa, and *Enterobacter*) are the major causes of nosocomial infections (Boucher et al., 2009). At the community level, the leading causes of resistant infections are pneumonia- and meningitis-causing *Streptococcus pneumoniae*; the enteric pathogens *Vibrio cholerae*, *Salmonella*, *Shigella*, *Escherichia coli*, and *Campylobacter*; and the urinary tract pathogens *Escherichia coli and Neisseria gonorrhoea* (Laxminarayan et al., 2013).

The health and economic burdens imposed by such resistance are immense. Antibiotic resistance leads to increased morbidity and mortality, as well as longer treatment durations and higher costs (Friedman, Temkin, & Carmeli, 2016). In the United States and Europe, resistant infections cause an estimated 23,000 and 25,000 deaths, respectively, each year; in LMICs, neonatal sepsis attributable to resistant infections alone causes an estimated 214,000 deaths each year (Laxminarayan, Sridhar, Blaser, Wang, & Woolhouse, 2016). The economic burden of antibiotic resistance, although under-studied, is considered substantial. The overall economic burden of antibiotic resistance is roughly estimated to be $1.8 billion in Europe (European Centre for Disease Control, 2009) and $55 billion in the United States (Centers for Disease Control and Prevention [CDC], 2011). However, the lack of broader consideration for the fundamental role of antibiotics in medicine means that even existing upper-limit estimates are likely to be underestimates (Smith & Coast, 2013). In the worst-case scenario, the foundation of modern medicines could be shaken by an "antibiotic apocalypse," in which effective antibiotics run out in the face of rising resistance.

Upstream Availability of Antimicrobials

The scale of the challenge of AMR, as illustrated by the global disease burden of infectious diseases and its potential economic costs, contrasts sharply with the dwindling pipeline for novel classes of antimicrobials, especially antibiotics. Despite being one of the most widely used and profitable class of medicines—accounting for 5% of the global pharmaceutical market and generating global sales of $42 billion in 2009 (So & Shah, 2014)—there has been significant stagnation in the R&D of antibiotics. After the discovery of penicillin by Alexander Fleming in 1928, most other major classes of antibiotics, such as cephalosporins, tetracyclines, macrolides, and quinolones, were discovered between the end of the 1940s and the early 1960s (Monnet, 2005). Between 1960 and 2000, no new classes of antibiotics with novel molecular structures were approved (Walsh & Wencemicz, 2016). Although interest in antibiotic discovery has revived in recent years with the approval of several "first in class" antibiotics, the current pipeline does not offer potential treatment options for priority resistant bacteria, especially for multidrug- and extensively drug-resistant TB and gram-negative pathogens such as *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and Enterobacteriaceae (WHO, 2017b).

This lack of investment is the result of a market failure created by the misalignment between the incentives created by the current pharmaceutical business model and basic public health considerations. First, the target diseases of antimicrobials—infectious diseases—are categorized as Type II and III diseases, which mostly affect low-income countries; compared with high-income countries addressed by medicines for Type I diseases, the target market of antimicrobials is considered far less lucrative.

Second, in a system where maximum profit is achieved by products that can be launched early, maintain effectiveness during their shelf life, and be sold as widely as possible, the unique properties of antibiotics make them an especially unattractive business proposition. To start with, efforts to preserve the effectiveness of these novel products significantly limits their window of sales. Novel antibiotic products are advanced treatments whose launch is likely to be delayed until the effectiveness of existing products is near exhaustion. Moreover, the highly regulated and restricted consumption of novel antibiotics constrains the volume of sales. The duration of antibiotic treatment in general is already short, and use of novel products is constrained due to considerations for conservation. This creates a paradox that disincentivizes the pharmaceutical industry: The more advanced a novel product, the less often it can be used and the lower its revenues.

With new antibiotics urgently needed, an increasing number of innovative financing mechanisms and initiatives are being introduced to address this market failure and stimulate the R&D pipeline. These mechanisms aim to resolve the tension between antibiotic conservation and the current volume-based business model by delinking financial rewards from volume of sales. Delinkage rewards companies for successful innovation on some basis other than sales volume, such as value- or milestone-based payments (Rex & Outterson, 2016). With such a scheme, previously underestimated public health and societal benefits of antibiotics are recognized and realigned with economic values. These include option value (in allowing preparedness for future epidemics), enablement value (in allowing other medical procedures to be undertaken), and diversity value (in reducing selection pressure on existing medicines) (Laxminarayan, Matsoso, et al., 2016; Laxminarayan & Weitzman, 2002).

Numerous push, pull, and mixed incentives for delinkage have been proposed since the mid-2000s (Mossialos et al., 2010), and some of these mechanisms have been implemented by initiatives in Europe and North America (Renwick, Simpkin, & Mossialos, 2016). Prominent examples of international public–private partnership programs are the European Commission's Innovative Medicines Initiative, WHO's Global Antibiotic Research and Development Partnership, and Combating Antibiotic Resistant Bacteria Biopharmaceutical Accelerator (CARB-X), a partnership of the U.S. Department of Health and Human Services, the Wellcome Trust, and Boston University School of Law (Renwick, Brogan, & Mossialos, 2016).

Access to, and Safe and Effective Use of, Antimicrobials

In tackling AMR, stimulating innovation of antimicrobials would be futile without ensuring adequate access to quality antimicrobials and their safe and effective use. Without access, innovation would be unjust; without efforts to conserve their effectiveness, investments in innovation and access would be wasteful (Hoffman & Outterson, 2015).

Access to novel antimicrobials is vital to treating increasingly resistant infections. Under the existing business model, the pricing of a novel medicine that has to be held back from the market and restricted in sales would likely be prohibitively high. Delinkage, with its rationale of separating costs from volume-based sales by rewarding innovation itself, could also facilitate access by enabling scale-up of affordable products through licensing arrangements such as patent buy-out or other upfront agreements conditioned to ensure affordable access (Laxminarayan et al., 2013; So & Shah, 2014). However, delinkage is an indirect mechanism for addressing access, and practical challenges arise in incorporating access-enabling features in the design of such mechanisms (Outterson et al., 2016).

Access to "old" but still useful antimicrobials is vital to safeguarding global health. Adequate access to antimicrobials has been a decisive factor in enabling modern medical procedures, significantly lowering the mortality associated with infectious diseases and leading to progress in maternal, neonatal, and child health worldwide (Laxminarayan et al., 2016; Mendelson et al., 2016). However, the lack of access to antibiotics remains a critical and worsening problem that affects both high-income countries and LMICs (Pulcini et al., 2017). The most influential upstream barrier to access concerns the lack of market availability, as pharmaceutical companies are deterred from entering the old antibiotics market—a generics market—due to low profit margins (Pulcini et al., 2017). The main downstream barriers are financial (affecting affordability), structural (affecting supply and distribution), and human resources (affecting capacity and management) barriers.

While there are currently no specific mechanisms to support access to antibiotics, experience can be drawn from novel organizational, financing, logistic, and procurement models developed for infectious diseases including malaria, HIV, and TB, exemplified by the Global Fund's Affordable Medicines Facility–Malaria and the U.S. President's Emergency Plan for AIDS Relief (Årdal et al., 2016; Mendelson et al., 2016). Implementation of WHO's Essential Medicines Policy, which has recently been overhauled in response to the escalating AMR crisis (Gulland, 2017), can also help LMICs improve access to quality assured antibiotics (Holloway & Henry, 2014). The Access to Medicine Foundation (2017) has developed a framework for benchmarking actions of pharmaceutical companies in improving global access to antimicrobials.

As much as ensuring adequate access is a critical issue, conserving the effectiveness of antibiotics by promoting rational use is a top priority at a time when consumption of antibiotics in humans is increasing worldwide (Van Boeckel et al., 2014). WHO's Essential Medicines Policy is a good example of how improvement in access can be combined with explicit policies for promoting rational use (Holloway, Rosella, & Henry, 2016).

One key strategy to improve rational antibiotic use is to change behaviors through addressing determinants of antibiotic use. The behaviors of healthcare providers (doctors and pharmacists) and users (patients) are influenced by multiple layers of factors that shape their respective incentives for antibiotic use (Belongia & Schwartz, 1998; Haak & Radyowijati, 2010; Hulscher, Grol, & Van Der Meer, 2010; Radyowijati & Haak, 2002, 2003; Teixeira Rodrigues, Roque, Falcão, Figueiras, & Herdeiro, 2013; Tonkin-Crine, Yardley, & Little, 2011). These factors are summarized in **TABLE 16-5**.

Although some of the structural and cultural determinants of antibiotic use can be addressed only indirectly by large-scale healthcare reforms or health system strengthening, determinants of antibiotic use in both hospital and community settings can be targeted by specific policies and interventions (Dar et al., 2015). These interventions, which are referred to as antibiotic stewardship, represent an organizational or healthcare system-wide approach to promoting and monitoring judicious use of antimicrobials to preserve their future effectiveness (NICE, 2015).

TABLE 16-5 Determinants of Antibiotic Use

Level of Determinants	Factors That Mainly Influence Providers	Factors That Influence Both Users and Providers
System	Governance, organization, financing, and delivery of healthcare in the health system; major policies and regulations in the healthcare and pharmaceutical sectors; role and influence of the pharmaceutical sector	Cultures, beliefs, values, and medical universes in different societies; availability, accessibility, and quality of healthcare services and medicines
Organization	Governance, organization, financing, and delivery of healthcare in hospitals; organizational policies and regulations specific to hospitals; interactions with the pharmaceutical sector	Availability, accessibility, and quality of healthcare services and medicines local to patients
Provider–user interpersonal		Interactions and relationships between doctors and patients
Individual		Personal background, experience, and socioeconomic characteristics; interactions with other actors in the health system; level of medical knowledge; attitudes and expectations toward health care and medicine use

In healthcare settings, antimicrobial stewardship can be introduced at national, hospital, and individual levels in the forms of national guidelines and programs, hospital antibiotic and therapeutic policies and guidelines, and training and educational activities. Effective implementation (Hulscher et al., 2010) of antimicrobial stewardship in inpatient (Davey, Wilcox, Irving, & Thwaites, 2015) and outpatient settings (Arnold & Straus, 2005) requires a multidisciplinary team comprising doctors, pharmacists, microbiologists, epidemiologists, and infectious disease specialists (Davey et al., 2015; Lee, Cho, Jeong, & Lee, 2013). In community settings, where both prescribed and nonprescribed use of antimicrobials take place and noncompliance and self-medication are prevalent (Kardas, Devine, Golembesky, & Roberts, 2005; Morgan, Okeke, Laxminarayan, Perencevich, & Weisenberg, 2011), antimicrobial stewardship activities target both providers (in this case, retail pharmacies) and patients.

On the supply side, more stringent regulation of over-the-counter sale of antibiotics can be introduced to restrict access to nonprescribed antibiotics (Santa-Ana-Tellez, Mantel-Teeuwisse, Dreser, Leufkens, & Wirtz, 2013). On the demand side, national-level public campaigns and lower-level educational interventions aimed at the public can improve patients' knowledge of antibiotic use and their use behaviors, especially when those campaigns are introduced alongside stewardship interventions aimed at providers (Haynes & Mcleod, 2015).

Other strategies to improve rational antibiotic use include measures to reduce reliance on antibiotics by reinforcing prevention and infection control measures (including consistent hand washing, and developing and ensuring adequate access to vaccines), developing alternative treatments to antibiotics, and ensuring the availability of convenient diagnostic technology to assist decision making.

Antimicrobial Resistance as a Multisectoral Issue

In addition to the challenges surrounding antibiotic use in humans, tackling AMR involves addressing human drivers of resistance in other sectors. The emergence and spread of carbapenem resistance mediated by the NDM-1 enzyme in drinking and sewage water in India, and colistin resistance encoded by the *mcr-1* gene in pig farms in China, are examples of how human activities in different sectors can significantly contribute to the rise of resistance to antibiotics that are medically important to humans in an interconnected and increasing globalized socio-ecosystem. For this reason, national and international strategies

to contain resistance recognize that AMR is a cross-sectoral problem requiring global collective action by international agencies, national governments, and NGOs (Hoffman et al., 2015).

A comprehensive set of multisectoral policies to tackle AMR set out by WHO in 2001 (WHO, 2001b), and reiterated in 2012 (WHO, 2012), helped trigger a series of policy responses at international and national levels. Nonetheless, implementation of these policies during this period remained partial and uncoordinated (Carlet, Pulcini, & Piddock, 2014; Gelband & Delahoy, 2014; WHO, 2015d).

Political interest and policy action in AMR has grown significantly since the 2010s. The initial focus on antibiotic R&D and other economic issues pertinent to AMR has gradually expanded to encompass a wide-ranging set of global health activities involving international agencies, national governments, and NGOs around the world (Carlet et al., 2014; Outterson, Powers, Daniel, & McClellan, 2015; Review on Antimicrobial Resistance, 2016). Political momentum around AMR culminated in the adoption of a Political Declaration to tackle AMR at the General Assembly of the United Nations in 2016 (OPGA/WHO/FAO/OIE joint news release, 2016), and progress continues to be made in the implementation of WHO's Global Action Plan on AMR (WHO, 2015a) and Global AMR Surveillance System (WHO, 2015b), and in efforts to enable monitoring and evaluation of policies related to AMR (**EXHIBIT 16-12**).

EXHIBIT 16-12 Objectives of WHO's Global Action Plan on AMR

- To improve awareness and understanding of AMR through effective communication, education, and training
- To strengthen the knowledge and evidence base through surveillance and research
- To reduce the incidence of infection through effective sanitation, hygiene, and infection prevention measures
- To optimize the use of antimicrobial medicines in human and animal health
- To develop the economic case for sustainable investment that takes account of the needs of all countries, and to increase investment in new medicines, diagnostic tools, vaccines, and other interventions

Reprinted from World Health Organization (WHO). (2015). Global Action Plan on antimicrobial resistance. Retrieved from: http://www.who.int/antimicrobial-resistance/publications/global-action-plan/en/

Pharmaceuticals for Noncommunicable Diseases

As discussed in detail in the *Chronic Diseases and Risks* chapter, noncommunicable diseases (NCDs) are now the leading cause of morbidity and mortality globally, and account for half of the burden of disease in lower-middle-income countries and one-third of that burden in low-income countries. Nevertheless, health systems, especially in LMICs, have struggled to adapt to the demands of this changing burden of disease. Ensuring availability, affordability, and rational use of pharmaceuticals for NCDs is one of the challenges yet to be conquered.

This section addresses the specific issues of medicines for NCDs, focusing primarily on cancers, diabetes, and cardiovascular and respiratory diseases as the most prevalent of these conditions, and as those for which pharmaceutical intervention is clearly indicated in the majority of cases. WHO's (2017g) Model List of Essential Medicines (MLEM) identifies antineoplastic and immunosuppressive medicines used in cancer treatment (Section 8), cardiovascular conditions (Section 12), insulins and other medicines used for diabetes (Section 18.5), and asthmatic conditions and chronic obstructive pulmonary disorder (COPD) (Section 25.1); these are the medicines considered here.

Concordance of National Essential Medicine Lists

A number of studies have assessed the extent of inclusion of those medicines recommended by the MLEM in national formularies. The 2015 MLEM added a significant number of new cancer medicines (which were retained in the 2017 list), recognizing the rising burden of disease from common cancers. Analysis of national medicines lists indicates that the degree of concordance with the MLEM varies considerably, with low-income countries including fewer of the recommended anticancer medications in their own national formularies (Cuomo & Mackey, 2017; Robertson, 2016).

Concordance between the MLEM and national formularies may be considered a necessary but not sufficient condition for rational prescribing: Consistency with treatment guidelines and training are also needed, as well as inclusion of medication specifications in procurement lists, reimbursement lists, and management guidelines (Bissell, 2015). For instance, a study in Papua New Guinea identified a range of inconsistencies between the medicines included in the Medical and Dental Catalogue and those included in standard treatment guidelines (Joshua, Passmore, & Sunderland,

2016), creating a source of confusion that may contribute to the high levels of irrational prescribing that have been documented in the country.

Availability and Affordability of NCD Medicines

The WHO/HAI database of medicine prices, availability, affordability, and price components (HAI, 2017) has been used to benchmark progress toward a target of 80% availability of affordable essential NCD medicines by 2025 (Ewen, Zweekhorst, Regeer, Laing, 2017). The surveys included in the database cover 50 medicines, 24 of which are indicated for cardiovascular diseases, diabetes, and COPD. The lowest-price generic medicines were available in a median of between 25% and 75% of public-sector medicine outlets and between 44% and 87% of private-sector medicine outlets surveyed across the categories of low-, lower-middle-, and upper-middle-income countries.

Affordability is measured in the WHO/HAI surveys in terms of the number of days' wages of the lowest-paid unskilled government worker. The level of 1 day's wage for a monthly supply of the medicine in question is used as a benchmark of affordability. On this basis, most low-price generics are classified as affordable when sourced from the public sector in the median medicine outlet, except for medicines for COPD in lower-middle-income countries. On the same basis, most were not affordable when sourced from the private sector, except for cardiovascular disease medicines in low-income countries (marginally at a median of 0.9 day's wage), and for diabetes and COPD medicines in upper-middle-income countries. Collectively, the proportion of public-sector outlets that had available and affordable the lowest-price generic medicine when surveyed ranged from 11% for cardiovascular medicines in low-income countries to 46% for diabetes medicines in upper-middle-income countries; in the private sector, these figures ranged from 9% for COPD medicines in lower-middle-income countries to 51% for cardiovascular medicines in upper-middle-income countries. Hence, the bulk of experience across the 30 countries included in the WHO/HAI database demonstrate the need for improvement to meet WHO's newly proposed target.

Babar et al. (2013) conducted their own survey using the WHO/HAI methodology, examining the availability and affordability of corticosteroids and a bronchodilator in 52 countries. These researchers found large variations between countries in terms of availability, pricing levels, and practices. For example, some countries subsidize asthma medicines purchases to a significant extent (e.g., Mozambique provides an 83% subsidy for generic salbutamol), while others leverage substantial margins on the sale of similar products (e.g., Peru charges 206% of the procurement price for generic budesonide). Other multicountry and national-level affordability analyses include those conducted by Mendis et al. (2007), Beran et al. (2015), Nguyen et al. (2009), Balasubramanian et al. (2014), and Jingi et al. (2014).

Other approaches have been used to measure the same concepts. For example, Jacobs et al. (2016) evaluated the capacity of the district level in Cambodia to manage NCDs, including an assessment of availability of key medicines for hypertension and diabetes. Unsurprisingly, given the absence of these medicines from the national formulary as described by Robertson et al. (2016), these researchers found the district system ill equipped to manage NCDs, and specifically found that the quantity of medicines provided allowed for only a few patients to be treated for short periods. Katende et al. (2015) used a similar approach to establish the lack of readiness of Ugandan health services, including the nonavailability of essential medicines for dealing with NCDs, which was most pronounced at the primary level. Choudhry et al. (2014) used data obtained from 4,000 "stockists" (wholesalers) of statins in India to track the rising rate of prescriptions of statins between 2006 and 2010, noting that this rate remained well below the estimate of need based on rates in high-income countries.

Pandey and Meltzer (2016) modeled the financial burden and impoverishment that would occur if households purchased cardiovascular medications according to need in India. They estimated that 17% of rural households and 32% of urban households would benefit from using these medications, but suggested that 17 million rural households and 10 million urban households would be impoverished by doing so.

In the private sector, the retail price of pharmaceuticals presents even more extreme issues for affordability. A recent survey examined the monthly cost of a basket of four commonly used cardiovascular disease medicines (aspirin, beta blocker, angiotensin-converting enzyme inhibitor, and statin) in retail pharmacies across 596 communities in 18 countries participating in the Prospective Urban Rural Epidemiology (Khatib et al., 2016). Using the catastrophic payment method described earlier, the study found that the combined monthly cost of all four medicines was potentially unaffordable for 0.14% of households in high-income countries, 25% in upper-middle-income countries, 33% in lower-middle-income countries, and 60% in low-income countries.

Robertson et al. (2016) compared the results of various approaches to measuring the availability of diabetes and hypertension medicines. As well as providing a range of data points from three methodologies, all of which suggested problems of low availability of NCD medicines, particularly in the public sector, these authors clarified the need for a standard approach on the basis of the divergence of estimates of the same outcome.

Causes of Poor Availability and Affordability of NCD Medicines

Mackay and Liang (2012) have established that patent protection is not the source of high medicine prices and low affordability of NCD medicines. Indeed, all of the active ingredients of medicines listed by the MLEM as treatments for cardiovascular and respiratory disease, cancers, and diabetes at the time of their research were available in generic versions.

Irrational prescribing has been measured in a range of settings such as East and South India (Bhavika, Prasanna, & Swathi, 2016; Prasad, Pradhan, Datta, Samajdar, & Panda, 2015), Pakistan (Riaz et al., 2016), and Moldova (Ferrario et al., 2016). The last study describes the inclusion of salbutamol syrup for children in the list of reimbursable medicines in Moldova's health insurance system, and in clinical protocols for family physicians, despite its higher price and lower effectiveness than inhaled formulations of salbutamol. Clearly, if the most cost-effective medicines are not purchased, then public-sector availability and/or affordability must be reduced; if private patients are prescribed less cost-effective medicines, then affordability is reduced.

Irrational prescribing and procuring may be caused by poor knowledge of the most cost-effective options on the part of prescribers or procurers. Some have alleged the deliberate promotion of inappropriate medicines by an industry that profits from their greater use. For example, Yudkin (2012) argues that analog insulins are being promoted to low-income countries as appropriate medicines, even though evidence suggests they are not a cost-effective option. Gill et al. (2017) resist an attempt to get analog insulins included in the MLEM on the same grounds. In some cases, resistance to the use of generics by both physicians and pharmacy staff drives the continued prescription and dispensing of medicines that are less cost-effective. For example, Flood et al. (2017) describe this situation in Guatemala. Patient pressure favoring brand-name products can also play a role.

A number of studies have explored the role played by the expansion of insurance systems on the availability and affordability of NCD medicines (Bigdeli et al., 2016; Ferrario et al., 2016; Garabedian, Ross-Degnan, Ratanawijitrasin, Stephens, & Wagner, 2012). Increasing insurance coverage might be expected to increase the availability and affordability of medicines if it is associated with increasing public subsidy; it might also increase the purchasing power of those who procure medicines for the insurance system, and thereby enhance their leverage over prices. Conversely, if purchasing is not well organized, the increased demand for NCD medicines may increase price; or if insurance is not universal, those excluded from coverage may face lower availability and higher prices.

In Thailand, Garabedian et al. (2012) found that the Universal Coverage Scheme increased access to medicines for noncommunicable diseases in primary care but not in secondary or tertiary care. In Moldova, Ferrario et al. (2016) found improvements over the period since introduction of national health insurance in the affordability of subsidized cardiovascular medicines, but suggested that this trend was due more to rising incomes of the period than to expanding coverage. Bigdeli et al. (2016) compared the patterns of access to NCD treatments and medicines under health equity funds and insurance schemes with those of two specific NCD-focused interventions, finding that coverage for NCDs under the two financing schemes was limited, leaving people shopping around in the private sector and often accessing higher-cost, less effective treatments. The disease-focused interventions were better at supporting NCD management using cost-effective strategies.

Strategies for Resolving the Issues of Access and Affordability of NCD Medicines

A number of commentators have suggested that the strategies used to extend access to antiretroviral drugs for HIV patients could be replicated for NCDs (e.g., Hogerzeil et al., 2013; Kishore, Vedanthan, & Fuster, 2011). These strategies have been enumerated by Kishore et al. (2011, p. 1982) as follows:

1. Enhancing capacity for generic substitution.
2. Expediting generic availability by overcoming legal barriers related to patents and licenses.
3. Optimizing local procurement practices in the public sector.
4. Broadening global procurement via third-party price negotiations.

5. Engaging the private sector to differentially price CVD medicines in LMICs.
6. Regulating retail markups in the supply chain.
7. Eliminating tariffs on medicines.
8. Developing a fixed-dose combination (FDC) for CVD (the "Polypill").

An example of the implementation of this thinking is the Asthma Drug Facility. This program was established by the International Union Against Tuberculosis and Lung Disease to organize the purchase of generic asthma inhalers for LMICs while assuring quality and reducing the price of these medicines. Although it succeeded in approximately halving the price of inhalers (Beran et al., 2015), funding ran out in 2014 and a new funding source has not been secured (Bissell, Perrin, & Beran, 2016).

Other strategies are implied by the problems mentioned earlier in this section. For example, reducing irrational prescription and procurement, and "joining up" national formularies, standard treatment guidelines, and training content all have the capacity to improve availability and affordability. When Barber et al. (2017) modeled the impact of full use of generic medicines for cancers in India, they found that a 56% price reduction in average price across 18 medicines compared could be achieved.

Agodokpessi et al. (2015) make the case for a revolving drug fund for essential asthma medicines in a country such as Benin where universal health coverage is far from achieved.

Conclusion

This chapter has identified a number of challenges to ensuring access to pharmaceuticals. These issues include the need to finance R&D for new medicines and vaccines for neglected diseases; the drive to strengthen national pharmaceutical distribution and management to ensure that the right pharmaceutical products are consistently available throughout the health system; the problems of affordability and the need to generate sustainable financing for medicines that are reasonably priced; and the challenges of ensuring safe and rational medicine use.

In many ways, these unmet needs are "old" problems that have challenged pharmaceutical systems in low- and middle-income settings for decades. In addition, a number of new challenges will demand further flexibility and responsiveness from these systems in the future. These issues include AMR and the need to both incentivize the development of new medicines and protect the effectiveness of existing products, and the epidemiologic transition away from infectious diseases and toward NCDs, which will mean that growing numbers of patients will need medicines on a lifelong basis.

The environment within which pharmaceuticals are developed, procured, and distributed is becoming increasingly more complex, particularly in the context of the epidemiologic transition and the move toward universal health coverage. A plethora of new actors are involved in this sometimes chaotic environment. At the national level, insurance funds and regulators will increasingly exert influence over pharmaceutical decisions that once were the main responsibility of national health ministries. At the international level, new actors include new corporate and philanthropic funding sources, new funding mechanisms and intermediaries, multinational pharmaceutical firms based in the Global South, and regional and international regulatory bodies. These agents may have different incentives for their engagement and sometimes face conflicting priorities. While this new environment has had a generally positive impact on the availability of pharmaceuticals to meet the needs of the low- and middle-income world, it also poses clear risks to the strength and stability of national pharmaceutical systems, and the new initiatives have at times created duplication, inefficiency, and resulting inequities.

From the material presented in this chapter, it should be clear that the functioning of the pharmaceutical system provides a snapshot of the performance of the health system as a whole. The measures that are needed to strengthen health systems in low- and middle-income settings—such as improved governance, universal coverage of equitable healthcare financing mechanisms, and an adequate, high-performing, and appropriately distributed workforce—will all have positive repercussions for the pharmaceutical system, and represent an important step toward ensuring access to pharmaceuticals for all. At the same time, many actions specific to the pharmaceutical sector can be taken across the patient, provider, and system levels to improve availability, affordability, and use of medicines.

Discussion Questions

1. Availability, affordability, and safe and effective use are the main challenges to improving pharmaceutical access. Define what is meant by each, and explain how they can be addressed.
2. What are "neglected diseases" and why are they neglected? Which policy interventions are

available to stimulate investment in pharmaceuticals to address these conditions?
3. Describe three approaches to improving affordability of medicines and some of the challenges of implementing them.
4. In what ways have global health initiatives had positive and negative impacts on medicine access?
5. Which strategies can be employed to improve supply chain management?
6. What are the implications of AMR? Which strategies can be used to contain the problem?
7. What is the impact of AMR on global health security?
8. What are some of the strategies for addressing AMR?
9. What are some of the challenges to continuous availability, affordability, and effective use of medicines to manage noncommunicable diseases?

References

Abuya, T. O., Mutemi, W., Karisa, B., Ochola, S. A., Fegan, G., & Marsh, V. (2007). Use of over-the-counter malaria medicines in children and adults in three districts in Kenya: Implications for private medicine retailer interventions. *Malaria Journal, 6*, 57.

Access to Medicine Foundation. (2017). *First independent framework for assessing pharmaceutical company action: Antimicrobial Resistance Benchmark 2018*. Amsterdam, Netherlands: Author.

Agodokpessi, G., Ait-Khaled, N., Gninafon, M., Tawo, L., Bekou, W., Perrin, C., . . . Chiang, C. Y. (2015). Assessment of a revolving drug fund for essential asthma medicines in Benin. *Journal of Pharmaceutical Policy and Practice, 8*, 12.

Alfadl, A. A., Ali, G. K. M., Yousif, M. A., & Babekir, M. F. A. (2016). Pharmacy practice in Sudan. In A. I. Fathelrahman, M. I. M. Ibrahim, & A. I. Wertheimer (Eds.), *Pharmacy practice in developing countries: Achievements and challenges* (p. 743). London, UK: Academic Press.

Anderson, S., & Huss, R. (2004). Issues in the management of pharmaceuticals in international health. In S. Anderson, R. Huss, R. Summers, & K. Wiedenmayer (Eds.), *Managing pharmaceuticals in international health* (pp. 1–17). Basel, Switzerland: Birkhauser Verlag.

Årdal, C., Outterson, K., Hoffman, S. J., Ghafur, A., Sharland, M., Ranganathan, N., . . . Røttingen, J. A. (2016). International cooperation to improve access to and sustain effectiveness of antimicrobials. *The Lancet, 387*, 296–307.

Arinaminpathy, N., Cordier-Lassalle, T., Lunte, K., & Dye, C. (2015). The Global Drug Facility as an intervention in the market for tuberculosis drugs. *Bulletin of the World Health Organization, 93*, 237–248a.

Arnold, S. R., & Straus, S. E. (2005). Interventions to improve antibiotic prescribing practices in ambulatory care. *Cochrane Database of Systematic Reviews, 1*, CD003539.

Aronson, J. K. (2006). Rare diseases and orphan drugs. *British Journal of Clinical Pharmacology, 61*, 243–245.

Attaran, A., Barry, D., Basheer, S., Bate, R., Benton, D., Chauvin, J., . . . McKee, M. (2012). How to achieve international action on falsified and substandard medicines. *British Medical Journal, 345*, e7381.

Babar, Z. U., Lessing, C., Mace, C., & Bissell, K. (2013). The availability, pricing and affordability of three essential asthma medicines in 52 low- and middle-income countries. *Pharmacoeconomics, 31*, 1063–1082.

Balasubramaniam, R., Benaragama, B. V., & Sri Ranganathan, S. (2014). A national survey of price and affordability of key essential medicines for children in Sri Lanka. *Indian Journal of Pediatrics, 81*, 362–367.

Ballou-Aares, D., Frietas, A., Kopezak, L. R., Kraiselburd, S., Laverty, M., & Macharia, E. (2008). The private sector's role in health supply chains: Review of the role and potential for private sector engagement in developing country health supply chains. Technical partner paper 13. Essential Medicines and Health Products Information Portal. Retrieved from http://apps.who.int/medicinedocs/en/d/Js16323e/

Barber, M., Gotham, D., & Hill, A. (2017). Potential price reductions for cancer medicines on the WHO Essential Medicines List. *European Journal of Cancer, 72*, S119.

Belongia, E. A., & Schwartz, B. (1998). Strategies for promoting judicious use of antibiotics by doctors and patients. *British Medical Journal, 317*, 668–671.

Beran, D., Zar, H. J., Perrin, C., Menezes, A. M., & Burney, P. (2015). Burden of asthma and chronic obstructive pulmonary disease and access to essential medicines in low-income and middle-income countries. *The Lancet Respiratory Medicine, 3*, 159–170.

Bermudez, J., & 't Hoen, E. (2010). The UNITAID patent pool initiative: Bringing patents together for the common good. *Open AIDS Journal, 4*, 37–40.

Bhavika, D., Prasanna, V., & Swathi, B. (2016). Drug utilization of anti-hypertensive drugs in a tertiary care hospital. *International Journal of Basic and Clinical Pharmacology, 5*, 1580–1585.

Bigdeli, M., Jacobs, B., Men, C. R., Nilsen, K., Van Damme, W., & Dujardin, B. (2016). Access to treatment for diabetes and hypertension in rural Cambodia: Performance of existing social health protection schemes. *PLoS One, 11*, e0146147.

Bigdeli, M., Peters, D. H., & Wagner, A. K. (2014). *Medicines in health systems: Advancing access, affordability and appropriate use*. Geneva, Switzerland: Alliance for Health Policy and Systems Research, World Health Organization.

Bissell, K. (2015). Essential medicines for chronic respiratory diseases: Can people breathe easily in low-income countries? *International Journal of Tuberculosis and Lung Disease, 19*, 1.

Bissell, K., Perrin, C., & Beran, D. (2016). Access to essential medicines to treat chronic respiratory disease in low-income countries. *International Journal of Tuberculosis and Lung Disease, 20*, 717–28.

Boucher, H. W., Talbot, G. H., Bradley, J. S., Edwards, J. E., Gilbert, D., Rice, L. B., . . . Bartlett, J. (2009). Bad bugs, no drugs: No ESKAPE! An update from the Infectious Diseases Society of America. *Clinical Infectious Diseases, 48*, 1–12.

Burki, T. (2010). The real cost of counterfeit medicines. *The Lancet Infectious Diseases, 10*, 585–586.

Cameron, A., Ewen, M., Auton, M., & Abegunde, D. (2011). Medicines prices, availability and affordability. In *The world medicines situation 2011* (3rd ed.). Geneva, Switzerland: World Health Organization. Retrieved from http://apps.who.int/medicinedocs/documents/s18065en/s18065en.pdf.

Cameron, A., Ewen, M., Ross-Degnan, D., Ball, D., & Laing, R. (2009). Medicine prices, availability, and affordability in 36 developing and middle-income countries: A secondary analysis. *The Lancet, 373*, 240–249.

Carlet, J., Pulcini, C., & Piddock, L. J. V. (2014). Antibiotic resistance: A geopolitical issue. *Clinical Microbiology and Infection, 20*, 949–953.

Caudron, J. M., Ford, N., Henkens, M., Mace, C., Kiddle-Monroe, R., & Pinel, J. (2008). Substandard medicines in resource-poor settings: A problem that can no longer be ignored. *Tropical Medicine & International Health, 13*, 1062–1072.

Centers tor Disease Control and Prevention (CDC). (2011). World Health Day: Media fact sheet. Retrieved from https://www.cdc.gov/media/releases/2011/f0407_antimicrobialresistance.pdf

Chaiyakunapruk, N., & Jones, S. M. (2016). *Pharmacy practice in Thailand*. London, UK: Academic Press.

Chauveau, J., Meiners, C. M., Luchini, S., & Moatti, J.-P. (2008). *Evolution of prices and quantities of ARV drugs in African countries: From emerging to strategic markets*, Cheltenham, UK/Northampton, MA: Elgar.

Chiguzo, A. N., Mugo, R. W., Wacira, D. G., Mwenda, J. M., & Njuguna, E. W. (2008). Delivering new malaria drugs through grassroots private sector. *East African Medical Journal, 85*, 425–431.

Choudhry, N. K., Dugani, S., Shrank, W. H., Polinski, J. M., Stark, C. E., Gupta, R., . . . Jha, P. (2014). Despite increased use and sales of statins in India, per capita prescription rates remain far below high-income countries. *Health Affairs (Millwood), 33*, 273–282.

Connelly, P., & Rosen, S. (2006). Treatment of HIV/AIDS at South Africa's largest employers: Myth and reality. *South African Medical Journal, 96*, 128–133.

Coriat, B. (2008). *The political economy of HIV/AIDS in developing countries: TRIPS, public health systems and free access*. Cheltenham, UK/Northampton, MA: Edward Elgar Publishing.

Creese, A. (2011, May). *WHO/HAI project on medicine prices and availability. Review series on pharmaceutical pricing and policies and interventions (Working paper 5: sales taxes on medicines)*. Amsterdam, Netherlands: Health Action International.

Cuomo, R. E., & Mackey, T. K. (2017). The availability of essential cancer medication: An analysis of national formularies. *Journal of Cancer Policy, 12*, 49–54.

Dar, O. A., Hasan, R., Schlundt, J., Harbarth, S., Caleo, G., Dar, F. K., . . . Heymann, D. L. (2015). Exploring the evidence base for national and regional policy interventions to combat resistance. *The Lancet, 387*, 285–295.

Davey, P., Wilcox, M., Irving, W., & Thwaites, G. (2015). *Antimicrobial chemotherapy* (pp. 206–226). Oxford, U.K.: Oxford University Press.

De Luca, C., & Calzavara, J. (2016). The Medicines Patent Pool: Five years on. Pharmafile. Retrieved from http://www.pharmafile.com/news/503074/medicines-patent-pool-five-years

Deisingh, A. K. (2005). Pharmaceutical counterfeiting. *Analyst, 130*, 271–279.

DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics, 47*, 20–33.

European Centre for Disease Control. (2009). *The bacterial challenge: Time to react*. Stockholm, Sweden: Author.

European Commission. (2017). Medicinal products for human use: Falsified medicines. Retrieved from https://ec.europa.eu/health/human-use/falsified_medicines_en

Ewen, M., Zweekhorst, M., Regeer, B., & Laing, R. (2017). Baseline assessment of WHO's target for both availability and affordability of essential medicines to treat non-communicable diseases. *PLoS One, 12*, e0171284.

Ferrario, A., Chitan, E., Seicas, R., Sautenkova, N., Bezverhni, Z., Kluge, H., & Habicht, J. (2016). Progress in increasing affordability of medicines for non-communicable diseases since the introduction of mandatory health insurance in the Republic of Moldova. *Health Policy and Planning, 31*, 793–800.

Flood, D., Mathieu, I., Chary, A., Garcia, P., & Rohloff, P. (2017). Perceptions and utilization of generic medicines in Guatemala: A mixed-methods study with physicians and pharmacy staff. *BMC Health Services Research, 17*, 27.

Flynn, M. (2008). Public production of anti-retroviral medicines in Brazil, 1990–2007. *Development and Change, 39*, 513–536.

Friedman, N. D., Temkin, E., & Carmeli, Y. (2016). The negative impact of antibiotic resistance. *Clinical Microbiology and Infection 22,* 416–422.

Garabedian, L. F., Ross-Degnan, D., Ratanawijitrasin, S., Stephens, P., & Wagner, A. K. (2012). Impact of universal health insurance coverage in Thailand on sales and market share of medicines for non-communicable diseases: An interrupted time series study. *British Medical Journal Open*, *2*(6), e001686.

GAVI. (2016). *Keeping children healthy: The Vaccine Alliance progress report 2015*. Geneva, Switzerland: GAVI, Vaccine Alliance.

Gelband, H., & Delahoy, M. (2014). *Policies to address antibiotic resistance in low-and middle-income countries*. Washington, DC: The Center for Disease Dynamics, Economics and Policy.

Gill, G., Yudkin, J. S., Tesfaye, S., De Courten, M., Gale, E., Motala, A., . . . Wild, S. (2017). Essential medicines and access to medicine. *The Lancet Diabetes and Endocrinology, 5*, 324–5.

Global Fund. (2017a). Price and quality reporting. Retrieved from https://www.theglobalfund.org/en/sourcing-management/price-quality-reporting

Global Fund. (2017b). *Results report 2016*. Geneva, Switzerland: Global Fund to Fight AIDS, Tuberculosis and Malaria.

Global Pharma Health Fund (GPHF). (2017). The GPHF-Minilab: Protection against counterfeit medicines. Retrieved from https://www.gphf.org/en/minilab/

Gulland, A. (2017). WHO targets antimicrobial resistance in new essential medicines list. *British Medical Journal, 357*, j2809.

Haak, H., & Radyowijati, A. (2010). Determinants of antimicrobial use: Poorly understood—poorly researched. In A. Sosa, D. Byarugaba, C. Amábile-Cuevas, P. R. Hsueh, S. Kariuki, & I. Okeke (Eds). *Antimicrobial resistance in developing countries* (pp. 283–300). New York, NY: Springer.

Hagmann, M. (2001). TB drug prices slashed for poor countries. *Bulletin of the World Health Organization, 79*, 904–905.

Haynes, C., & Mcleod, C. (2015). *A review of reviews of educational interventions designed to change the public's knowledge and behaviour in relation to antimicrobial use and antimicrobial resistance that target healthcare professionals and patients*. London, England: Social Care Centre, National Institute for Health and Care Excellence.

Health Action International (HAI). (2017). Database of medicine prices, availability, affordability and price components. Retrieved from http://www.haiweb.org/MedPriceDatabase/

Hoffman, S. J., Caleo, G. M., Daulaire, N., Elbe, S., Mossialos, E., & Røttingen, J.-A. (2015). Strategies for achieving global collective action on antimicrobial resistance. *Bulletin of the World Health Organization*, *93*, 867–876.

Hoffman, S. J., & Outterson, K. (2015). What will it take to address the global threat of antibiotic resistance? *Journal of Law, Medicine and Ethics, 43*, 363–368.

Hogerzeil, H. V., Liberman, J., Wirtz, V. J., Kishore, S. P., Selvaraj, S., Kiddell-Monroe, R., . . . & Von Schoen-Angerer, T. (2013). Promotion of access to essential medicines for non-communicable diseases: Practical implications of the UN political declaration. *The Lancet, 381*, 680–689.

Holloway, K. A., & Henry, D. (2014). WHO essential medicines policies and use in developing and transitional countries: An analysis of reported policy implementation and medicines use surveys. *PLoS Medicine, 11*, e1001724.

Holloway, K. A., Rosella, L., & Henry, D. (2016). The impact of WHO essential medicines policies on inappropriate use of antibiotics. *PloS one, 11*, e0152020.

Holmes, A. H., Moore, L. S. P., Sundsfjord, A., Steinbakk, M., Regmi, S., Karkey, A., Guerin, P. J., & Piddock, L. J. V. (2016). Understanding the mechanisms and drivers of antimicrobial resistance. *The Lancet, 387*, 176–187.

Hulscher, M. E. J. L., Grol, R. P. T. M., & Van Der Meer, J. W. M. (2010). Antibiotic prescribing in hospitals: A social and behavioural scientific approach. *The Lancet Infectious Diseases, 10*, 167–175.

Hussain, A., & Jamshed, S. (2016). *Pharmacy practice in Pakistan*. London, UK: Academic Press.

International Council for Harmonization of Technical Requirements. (1996). Guidelines for good clinical practice. Retrieved from http://www.ich.org/products/guidelines/efficacy/efficacy-single/article/good-clinical-practice.html

International Dispensary Association. (2017). Product indicator 2017. Retrieved from https://www.idafoundation.org/media/wysiwyg/PDF/IDA1601_Price_indicator_zonder_prijzen.pdf

International Federation of Pharmaceutical Manufacturers Associations (IFPMA). (2017). *The pharmaceutical industry and gobal health: Facts and figures 2017*. Geneva, Switzerland: Author.

International Medical Products Anti-Counterfeiting Taskforce (IMPACT). (2006). Counterfeit drugs kill. Retrieved from https://www.gphf.org/images/downloads/impactbrochure.pdf

International Medical Products Anti-Counterfeiting Taskforce (IMPACT). (2011). The handbook: Facts, activities, documents developed by the Assembly and the Working Groups, 2006–2010. Retrieved from http://apps.who.int/medicinedocs/en/d/Js20967en/

International Network for Rational Use of Drugs (INRUD). (2017). Bibliographies. Retrieved from https://sites.google.com/a/msh.org/inrud-archive/bibliographies

Jacobs, B., Hill, P., Bigdeli, M., & Men, C. (2016). Managing non-communicable diseases at health district level in Cambodia: A systems analysis and suggestions for improvement. *BMC Health Services Research, 16*, 32.

Jingi, A. M., Noubiap, J. J., Ewane Onana, A., Nansseu, J. R., Wang, B., Kingue, S., & Kengne, A. P. (2014). Access to diagnostic tests and essential medicines for cardiovascular diseases and diabetes care: Cost, availability and affordability in the West Region of Cameroon. *PLoS One, 9*, e111812.

Joshua, I. B., Passmore, P. R., & Sunderland, B. V. (2016). An evaluation of the essential medicines list, standard treatment guidelines and prescribing restrictions, as an integrated strategy to enhance quality, efficacy and safety of and improve access to essential medicines in Papua New Guinea. *Health Policy and Planning, 31*, 538–546.

Kanjanarat, P., Winterstein, A. G., Johns, T. E., Hatton, R. C., Gonzalez-Rothi, R., & Segal, R. (2003). Nature of preventable adverse drug events in hospitals: A literature review. *American Journal of Health-System Pharmacy, 60*, 1750–1759.

Kardas, P., Devine, S., Golembesky, A., & Roberts, C. (2005). A systematic review and meta-analysis of misuse of antibiotic therapies in the community. *International Journal of Antimicrobial Agents, 26*, 106–113.

Katende, D., Mutungi, G., Baisley, K., Biraro, S., Ikoona, E., Peck, R., . . . Grosskurth, H. (2015). Readiness of Ugandan health services for the management of outpatients with chronic diseases. *Tropical Medicine & International Health, 20*, 1385–1395.

Kerry, V. B., & Lee, K. (2007). TRIPS, the Doha Declaration and Paragraph 6 decision: What are the remaining steps for protecting access to medicines? *Global Health, 3*, 3.

Khatib, R., McKee, M., Shannon, H., Chow, C., Rangarajan, S., Teo, K., . . .Yusuf, S. (2016). Availability and affordability of cardiovascular disease medicines and their effect on use in high-income, middle-income, and low-income countries: An analysis of the PURE study data. *The Lancet, 387*, 61–69.

Kishore, S. P., Vedanthan, R., & Fuster, V. (2011). Promoting global cardiovascular health ensuring access to essential cardiovascular medicines in low- and middle-income countries. *Journal of the American College of Cardiologists, 57*, 1980–1987.

Laing, R., Hogerzeil, H., & Ross-Degnan, D. (2001). Ten recommendations to improve use of medicines in developing countries. *Health Policy and Planning, 16*, 13–20.

Laing, R. O., & McGoldrick, K. M. (2000). Tuberculosis drug issues: Prices, fixed-dose combination products and second-line drugs. *International Journal of Tuberculosis and Lung Disease, 4*, S194–S207.

Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A. K. M., Wertheim, H. F. L., Sumpradit, N., . . . Cars, O. (2013). Antibiotic resistance: The need for global solutions. *The Lancet Infectious Diseases, 13*, 1057–1098.

Laxminarayan, R., Matsoso, P., Pant, S., Brower, C., Røttingen, J.-A., Klugman, K., & Davies, S. C. (2016). Access to effective antimicrobials: A worldwide challenge. *The Lancet, 387*, 168–175.

Laxminarayan, R., Sridhar, D., Blaser, M., Wang, M., & Woolhouse, M. (2016). Achieving global targets for antimicrobial resistance. *Science, 353*, 1057–875.

Laxminarayan, R., & Weitzman, M. L. (2002). On the implications of endogenous resistance to medications. *Journal of Health Economics, 21*, 709–718.

Lee, C. R., Cho, I. H., Jeong, B. C., & Lee, S. H. (2013). Strategies to minimize antibiotic resistance. *International Journal of Environmental Research and Public Health, 10*, 4274–4305.

Leoni, P., & Sandroni, A. (2016). Can patent duration hinder medical innovation. *International Journal of Health Economics and Management, 16*, 397–406.

Levy, M., & Rizansky, A. (2014). Market failure in the pharmaceutical industry and how it can be overcome: The CureShare mechanism. *European Journal of Health Economics, 15*, 143–156.

Light, D. W., & Lexchin, J. R. (2012). Pharmaceutical research and development: What do we get for all that money? *British Medical Journal, 345*, e4348.

Lu, Y., Hernandez, P., Abegunde, D., & Edejer, T. (2011). Medicine expenditures. In *The world medicines situation 2011* (3rd ed.). Geneva, Switzerland: World Health Organization. Retrieved

from http://apps.who.int/medicinedocs/documents/s18767en/s18767en.pdf

Mackey, T. K., & Liang, B. A. (2012). Patent and exclusivity status of essential medicines for non-communicable disease. *PLoS One, 7*, e51022.

Mackintosh, M., Banda, G., Wamae, W., & Tibandebage, P. (2016). *Making medicines in Africa: The political economy of industrializing for local health*. (International Political Economy Series). London, England: Palgrave Macmillan.

Mahmood, K. T., Tahir, F. A. M., & Haq, I. U. (2011). Pharmacovigilance: A need for best patient care in Pakistan: A review. *Journal of Pharmaceutical Sciences and Research, 3*, 1566–1584.

Management Sciences for Health. (2010). ADDO (accredited drug dispensing outlets) program. Retrieved from https://www.msh.org/journal-tags/addo-accredited-drug-dispensing-outlets-program

Management Sciences for Health. (2012). *MDS-3: Managing access to medicines and health technologies.* Arlington, VA: Management Sciences for Health & World Health Organization. Retrieved from https://www.msh.org/resources/mds-3-managing-access-to-medicines-and-health-technologies

Management Sciences for Health. (2017). International drug price indicator guide. Retrieved from https://www.msh.org/resources/international-drug-price-indicator-guide

Marsh, V. M., Mutemi, W. M., Willetts, A., Bayah, K., Were, S., Ross, A., & Marsh, K. (2004). Improving malaria home treatment by training drug retailers in rural Kenya. *Tropical Medicine & International Health, 9*, 451–460.

Mendelson, M. (2017). Antibiotic resistance has a language problem. *Nature, 545*, 23–25.

Mendelson, M., Røttingen, J. A., Gopinathan, U., Hamer, D. H., Wertheim, H., Basnyat, B., . . . Balasegaram, M. (2016). Maximising access to achieve appropriate human antimicrobial use in low-income and middle-income countries. *The Lancet, 387*, 188–198.

Mendis, S., Fukino, K., Cameron, A., Laing, R., Filipe, A., Jr., Khatib, O., . . . Ewen, M. (2007). The availability and affordability of selected essential medicines for chronic diseases in six low- and middle-income countries. *Bulletin of the World Health Organization, 85*, 279–288.

Monnet, D. L. (2005). Antibiotic development and the changing role of the pharmaceutical industry. *International Journal of Risk and Safety in Medicine, 17*, 133–145.

Morgan, D. J., Okeke, I. N., Laxminarayan, R., Perencevich, E. N., & Weisenberg, S. (2011). Non-prescription antimicrobial use worldwide: A systematic review. *The Lancet Infectious Diseases 11*, 692–701.

Mossialos, E., Morel, C. M., Edwards, S. E., Berensen, J., Gemmill, M., & Brogen, D. (2010). Policies and incentives for promoting innovation in antibiotic research. *European Observatory on Health Systems and Policies*, 1–197.

Munos, B. (2009). Lessons from 60 years of pharmaceutical innovation. *Nature Reviews Drug Discovery, 8*, 959–968.

Musungu, S. F., & Oh, C. (2006). *The use of flexibilities in TRIPS by developing countries: Can they promote access to medicines?* Geneva, Switzerland: World Health Organization.

Nebeker, J. R., Barach, P., & Samore, M. H. (2004). Clarifying adverse drug events: A clinician's guide to terminology, documentation, and reporting. *Annals of Internal Medicine, 140*, 795–801.

Nguyen, T. A., Knight, R., Mant, A., Cao, Q. M., & Auton, M. (2009). Medicine prices, availability, and affordability in Vietnam. *Southern Medical Review, 2*, 2–9.

Nguyen, T. A., Knight, R., Roughead, E. E., Brooks, G., & Mant, A. (2015). Policy options for pharmaceutical pricing and purchasing: Issues for low- and middle-income countries. *Health Policy and Planning, 30*, 267–280.

NICE. (2015). Antimicrobial stewardship: Systems and processes for effective antimicrobial medicine use. Retrieved from https://www.nice.org.uk/guidance/ng15

Nicol, D., & Owoeye, O. (2013). Using TRIPS flexibilities to facilitate access to medicines. *Bulletin of the World Health Organization, 91*, 533–539.

Niëns, L., Poel, E. V. D., Cameron, A., Ewen, M., Laing, R., & Brouwer, W. (2012). Practical measurement of affordability: An application to medicines. *Bulletin of the World Health Organization, 90*, 219–227.

Olcay, M., & Laing, R. (2005). *Pharmaceutical tariffs: What is their effect on prices, protection of local industry and revenue generation?* Geneva, Switzerland: World Health Organization & Commission on Intellectual Property Rights, Innovation and Public Health.

OPGA/WHO/FAO/OIE joint news release. (2016). Press release: High-level meeting on antimicrobial resistance. General Assembly of the United Nations. Retrieved from https://www.un.org/pga/71/2016/09/21/press-release-hl-meeting-on-antimicrobial-resistance/

Outterson, K., Gopinathan, U., Clift, C., So, A. D., Morel, C. M., & Røttingen, J. A. (2016). Delinking investment in antibiotic research and development from sales revenues: The challenges of transforming a promising idea into reality. *PLoS Medicine, 13*, e1002043.

Outterson, K., Powers, J. H., Daniel, G. W., & McClellan, M. B. (2015). Repairing the broken market for antibiotic innovation. *Health Affairs, 34*, 277–285.

Outterson, K., Samora, J. B., & Keller-Cuda, K. (2007). Will longer antimicrobial patents improve global public health? *The Lancet Infectious Diseases, 7*, 559–566.

Pandey, K. R., & Meltzer, D. O. (2016). Financial burden and impoverishment due to cardiovascular medications in low and middle income countries: An illustration from India. *PLoS One, 11*, e0155293.

Patouillard, E., Hanson, K. G., & Goodman, C. A. (2010). Retail sector distribution chains for malaria treatment in the developing world: A review of the literature. *Malaria Journal, 9*, 50.

Petrova, E. (2014). Innovation in the pharmaceutical industry: The process of drug discovery and development. In M. Ding, J. Eliashberg, & S. Stremersch (Eds.), *Innovation and marketing in the pharmaceutical industry* (pp. 19–81). New York, NY: Springer.

Power, E. (2006). Impact of antibiotic restrictions: The pharmaceutical perspective. *Clinical Microbiology and Infection, 12*, 25–34.

Prasad, A., Pradhan, S. P., Datta, P. P., Samajdar, S. S., & Panda, P. (2015). Drug prescription pattern for bronchial asthma in a tertiary-care hospital in Eastern India. *National Journal of Physiology, Pharmacy and Pharmacology, 5*, 263–266.

Pulcini, C., Beovic, B., Béraud, G., Carlet, J., Cars, O., Howard, P., . . . Sharland, M. (2017). Ensuring universal access to old antibiotics: A critical but neglected priority. *Clinical Microbiology and Infection*, *23*, 590—592.

Pulcini, C., Mohrs, S., Beovic, B., Gyssens, I., Theuretzbacher, U., Cars, O., & ESCMID Study Group for Antibiotic Policies. (2017). Forgotten antibiotics: A follow-up inventory study in Europe, the USA, Canada and Australia. *International Journal of Antimicrobial Agents, 49*, 98–101.

Radyowijati, A., & Haak, H. (2002). Determinants of antimicrobial use in the developing world. *Child Health Research Project Special Report, 4*, 1–36.

Radyowijati, A., & Haak, H. (2003). Improving antibiotic use in low-income countries: An overview of evidence on determinants. *Social Science and Medicine, 57*, 733–744.

Renwick, M. J., Brogan, D. M., & Mossialos, E. (2016). A systematic review and critical assessment of incentive strategies for discovery and development of novel antibiotics. *Journal of Antibiotics, 69*, 73–88.

Renwick, M. J., Simpkin, V., & Mossialos, E. (2016). Targeting innovation in antibiotic drug discovery and development: The need for a One Health–One Europe–One World Framework. World Health Organization (acting as the host organization for, and secretariat of, the European Observatory on Health Systems and Policies). London, England: European Observatory on Health Systems and Policies. Retrieved from http://apps.who.int/medicinedocs/documents/s23164en/s23164en.pdf

Review on Antimicrobial Resistance. (2016). *Tackling drug-resistant infections globally: Final report and recommendations*. Retrieved from https://amr-review.org/sites/default/files/160518_Final%20paper_with%20cover.pdf

Rex, J. H., & Outterson, K. (2016). Antibiotic reimbursement in a model delinked from sales: A benchmark-based worldwide approach. *The Lancet Infectious Diseases*, *16*, 500–505. London https://amr-review.org/sites/default/files/160518_Final%20paper_with%20cover.pdf

Riaz, H., Godman, B., Bashir, S., Hussain, S., Mahmood, S., Waseem, D., . . . Raza, S. A. (2016). Evaluation of drug use indicators for non-communicable diseases in Pakistan. *Acta Poloniae Pharmaceutica, 73*, 787–794.

Robertson, J., Barr, R., Shulman, L. N., Forte, G. B., & Magrini, N. (2016). Essential medicines for cancer: WHO recommendations and national priorities. *Bulletin of the World Health Organization, 94*, 735–742.

Robinson, J. (2016). What is the falsified medicines directive? *Pharmaceutical Journal, 297*, 243.

Rutta, E. (2013). Flagship report 2014: Medicines in health systems: Advancing access, affordability, and appropriate use. Chapter 5: Accredited drug dispensing outlets. Retrieved from http://www.who.int/alliance-hpsr/resources/FR_Ch5_Annex1.pdf

Saha, C. N., & Bhattacharya, S. (2011). Intellectual property rights: An overview and implications in pharmaceutical industry. *Journal of Advanced Pharmaceutical Technology & Research, 2*, 88–93.

Samb, B., Evans, T., Dybul, M., Atun, R., Moatti, J. P., Nishtar, S., . . . Etienne, C. (2009). An assessment of interactions between global health initiatives and country health systems. *The Lancet, 373*, 2137–2169.

Santa-Ana-Tellez, Y., Mantel-Teeuwisse, A. K., Dreser, A., Leufkens, H. G. M., & Wirtz, V. J. (2013). Impact of over-the-counter restrictions on antibiotic consumption in Brazil and Mexico. *PLoS One, 8*, e75550.

Scannell, J. W., & Bosley, J. (2016). When quality beats quantity: Decision theory, drug discovery, and the reproducibility crisis. *PLoS One, 11*, e0147215.

Scudellari, M. (2011). Are advance market commitments for drugs a real advance? *Nature Medicine, 17*, 139–139.

Sharma, D. C. (2003). ARV prices nosedive after Clinton brokering. *The Lancet, 362*, 1467.

Sjöstedt, M. (2013). Aid effectiveness and the Paris Declaration: A mismatch between ownership and results-based management? *Public Administration and Development, 33*, 143–155.

Smith, R. D., & Coast, J. (2013). The true cost of antimicrobial resistance. *British Medical Journal, 346*, f1493.

So, A. D., & Shah, T. A. (2014). New business models for antibiotic innovation. *Upsala Journal of Medical Sciences, 119*, 176–180.

Stillman, K. & Bennett, S. (2005). *Systemwide effects of the Global Fund: Interim findings from three country studies*. Bethesda, MD: Partners for Health Reformplus Project, Abt Associates.

Stop Stockouts Campaign. (2015). Third annual national stop stockouts survey. Retrieved from http://stockouts.org/Download/2015_stock_outs_national_survey.pdf

Stop TB Partnership. (2017). Global Drug Facility. Retrieved from http://www.stoptb.org/gdf/

Stop TB Partnership & World Health Organization (WHO). (2014). *Stop TB Partnership Global Drug Facility activity report 2012–2013*. Geneva, Switzerland: WHO.

Strengthening Pharmaceutical Systems Program. (2011). *Safety of medicines in sub-Saharan Africa: Assessment of pharmacovigilance systems and their performance*. Arlington, VA: Management Sciences for Health. Retrieved from http://apps.who.int/medicinedocs/documents/s19152en/s19152en.pdf

Teixeira Rodrigues, A., Roque, F., Falcão, A., Figueiras, A., & Herdeiro, M. T. (2013). Understanding physician antibiotic prescribing behaviour: A systematic review of qualitative studies. *International Journal of Antimicrobial Agents, 41*, 203–212.

't Hoen, E. (2016). *Private patents and public health: Changing intellectual property rules for access to medicines*. Amsterdam, Netherlands: Health Action International.

Tonkin-Crine, S., Yardley, L., & Little, P. (2011). Antibiotic prescribing for acute respiratory tract infections in primary care: A systematic review and meta-ethnography. *Journal of Antimicrobial Chemotherapy, 66*, 2215–2223.

Trouiller, P., Olliaro, P., Torreele, E., Orbinski, J., Laing, R., & Ford, N. (2002). Drug development for neglected diseases: A deficient market and a public-health policy failure. *The Lancet, 359*, 2188–2194.

United Nations. (2005). *Consolidated list of products whose consumption and/or sale have been banned, withdrawn, severely restricted or not approved by governments*. New York, NY: Author.

United Nations International Children's Emergency Fund. (2017). Supply catalogue. Retrieved from https://supply.unicef.org/unicef_b2c/app/displayApp/(layout=7.0-12_1_66_67_115&carea=%24ROOT)/.do?rf=y

Vaccine Confidence Project. (2015). The state of vaccine confidence. Retrieved from http://www.vaccineconfidence.org/The-State-of-Vaccine-Confidence-2015.pdf

Van Boeckel, T. P., Gandra, S., Ashok, A., Caudron, Q., Grenfell, B. T., Levin, S. A., & Laxminarayan, R. (2014). Global antibiotic consumption 2000 to 2010: An analysis of national pharmaceutical sales data. *The Lancet Infectious Diseases, 14*, 742–750.

Viana, A. L. D. Á., Silva, H. P. D., Ibañez, N., Iozzi, F. L., Viana, A. L. D. Á., Silva, H. P. D., . . . & Iozzi, F. L. (2016). A política de desenvolvimento produtivo da saúde e a capacitação dos laboratórios públicos nacionais. *Cadernos de Saúde Pública*, 32.

Village Reach. (2012, February 27). Starting at the last mile: Rethinking medical supply chains in low-income countries. Retrieved from http://www.villagereach.org/wp-content/uploads/2012/03/02272012.Starting-at-the-Last-Mile-Whitepaper.pdf

Vrueh, R. L. A. D., & Crommelin, D. J. A. (2017). Reflections on the future of pharmaceutical public–private partnerships: From input to impact. *Pharmaceutical Research*, *34*(10): 1985–1999.

Walsh, C., & Wencemicz, T. (2016). *Antibiotics: Challenges, mechanisms, opportunities*. Washington, DC: American Society for Microbiology Press.

Wiedenmayer, K. (2004). Rational use of medicines. In S. C. Anderson, R. Summers, & K. Wiedenmayer (Eds.), *Managing pharmaceuticals in international health* (pp. 141–152). Basel, Switzerland: Birkhauser Verlag.

Wirtz, V. J., Hogerzeil, H. V., Gray, A. L., Bigdeli, M., De Joncheere, C. P., Ewen, M. A., . . . Reich, M. R. (2017). Essential medicines for universal health coverage. *The Lancet, 389*, 403–476.

World Health Organization (WHO). (2001a). *Macroeconomics and health: Investing in health for economic development: Report of the Commission on Macroeconomics and Health*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2001b). *WHO global strategy for containment of antimicrobial resistance*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2002). *The importance of pharmacovigilance: Safety monitoring of medicinal products*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2005). Access to medicines: Intellectual property protection: impact on public health. *WHO Drug Information*, *19*(3), 236–241.

World Health Organization (WHO). (2007). Quality assurance of pharmaceuticals: A compendium of guidelines and related materials. Vol. 2, 2nd updated ed., Good manufacturing practices and inspection. Retrieved from http://www.who.int/medicines/areas/quality_safety/quality_assurance/QualityAssurancePharmVol2.pdf

World Health Organization (WHO). (2010a). Combating counterfeit, falsified and substandard, medicines: Defining the way forward? Retrieved from http://apps.who.int/medicinedocs/documents/s17549en/s17549en.pdf

World Health Organization (WHO). (2010b). Essential medicines and health products. Retrieved from http://www.who.int/medicines/services/essmedicines_def/en/

World Health Organization (WHO). (2010c). Pharmaceuticals: Restrictions in use and availability. Retrieved from http://apps.who.int/medicinedocs/documents/s17126e/s17126e.pdf

World Health Organization (WHO). (2010d). Promoting rational use of medicines: Core components. Retrieved from http://apps.who.int/medicinedocs/en/d/Jh3011e/1.html

World Health Organization (WHO). (2010e). Substandard and falsified medicines. Retrieved from http://www.who.int/mediacentre/factsheets/fs275/en/

World Health Organization (WHO). (2012). *The evolving threat of antimicrobial resistance: Options for action*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2014). *Antimicrobial resistance: Global report on surveillance*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2015a). *Global action plan on antimicrobial resistance*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2015b). *Global Antimicrobial Resistance Surveillance System: Manual for early implementation*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2015c). *WHO guideline on country pharmaceutical pricing policies*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2015d). *Worldwide country situation analysis: Response to antimicrobial resistance*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2016). *The role of intellectual property in local production in developing countries*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2017a). Adverse drug reactions monitoring. Retrieved from http://www.who.int/medicines/areas/quality_safety/safety_efficacy/advdrugreactions/en/

World Health Organization (WHO). (2017b). *Antibacterial agents in clinical development: an analysis of the antibacterial clinical development pipeline, including tuberculosis*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2017c). The international pharmacopoeia. Retrieved from http://www.who.int/medicines/publications/pharmacopoeia/en/

World Health Organization (WHO). (2017d). Pharmacovigilance. Retrieved from http://www.who.int/medicines/areas/quality_safety/safety_efficacy/pharmvigi/en/

World Health Organization (WHO). (2017e). Prequalification of medicines by WHO. Retrieved from http://www.who.int/mediacentre/factsheets/fs278/en/

World Health Organization (WHO). (2017f, March; amended August 2017). WHO model list of essential medicines for children, 6th list. Retrieved from http://www.who.int/medicines/publications/essentialmedicines/6th_EMLc2017_FINAL_amendedAug2017.pdf?ua=1

World Health Organization (WHO). (2017g). WHO model list of essential medicines: 20th list. Retrieved from http://www.who.int/medicines/publications/essentialmedicines/en/

World Health Organization (WHO). (2017h). The WHO Programme for International Drug Monitoring. Retrieved from http://www.who.int/medicines/areas/quality_safety/safety_efficacy/National_PV_Centres_Map/en/

World Health Organization (WHO) & Health Action International (HAI). (2008). *Measuring medicine prices, availability, affordability and price components* (2nd ed.). Geneva, Switzerland: Author.

Yudkin, J. S. (2012). The drug industry is a barrier to diabetes care in poor countries. *British Medical Journal, 344*, e3018.

Zurovac, D., Talisuna, A. O., & Snow, R. W. (2012). Mobile phone text messaging: Tool for malaria control in Africa. *PLoS Medicine, 9*, e1001176.

CHAPTER 17

Innovation, Technology, and Design

Brady Hunt, Mary Natoli, Eric Richardson, Maria Oden, and Rebecca Richards-Kortum

Introduction

Why is it so challenging to use technology to solve critical global health challenges? In this chapter, we address this question by examining the process of technical innovation and engineering design, with a focus on how technology can be used to develop solutions that equitably improve lives throughout the world. Although we emphasize the design of medical technologies, the same general principles can be applied to design technical solutions to meet other global health needs, such as better sanitation and agriculture.

We begin by considering the importance of innovation and technology for the student of global health. We then explore three promising areas of technical innovation that are impacting global health: (1) mobile health technologies, (2) point-of-care diagnostic technologies, and (3) improved access technologies. In the third section, we give an overview of the innovation process, with an eye toward approaches that address global health challenges. Because technical progress alone does not guarantee equitable health outcomes (Fong & Harris, 2015), we also consider the important question of how to scale up successful technologies. Finally, we conclude by discussing some of the high-priority gaps that could be addressed with new innovations.

Innovation: A Catalyst for Change in Global Health

What is innovation, and what does it look like when applied to a vast global health problem? We begin our discussion of innovation and global health by defining technical innovation, and then extend that definition into the context of global health.

Innovation is one of the hallmarks of human civilization. The success of the human race as a species is innately tied to our ability to understand a problem, explore possible solutions, and then find new and better ways to solve it. Innovation is often accompanied by scientific discovery—a better understanding that opens up new possibilities. For our discussion of innovation and technology in the context of global health, we define *technical innovation* as the use of science to meet a social need in the face of economic constraints (Niemeier, Gombachika, & Richards-Kortum, 2014).

Technical innovation is responsible for many significant advances in medical care and public health in high-income countries (**FIGURE 17-1**). Technology to

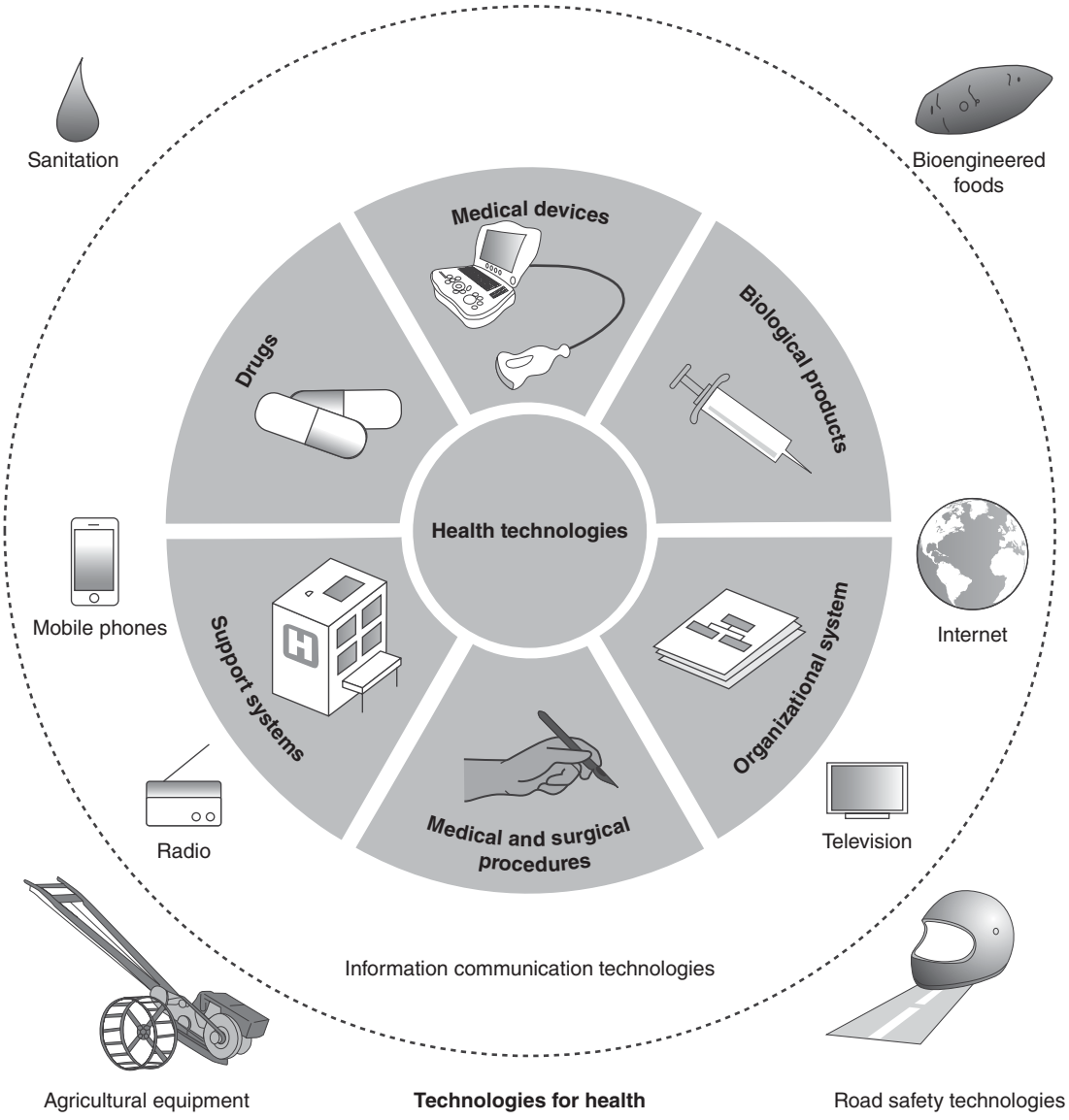


FIGURE 17-1 Overview of technologies that could improve global health.

Reprinted from Howitt, P., Darzi, A., Yang, G. Z., Ashrafian, H., Atun, R., Barlow, J., . . . Wilson, E. (2012). Technologies for global health. *The Lancet, 380*(9840), 507–535. Copyright © 2012, with permission from Elsevier.

improve global health includes medical tools such as antibiotics, vaccines, and medical devices, as well as tools to improve road safety, sanitation, and agricultural productivity (Howitt et al., 2012). Yet, with the exception of vaccines and pharmaceuticals, medical technology and innovation have not yet played a major role in global public health (Piot, 2012). For example, engineers have known how to produce safe drinking water and build toilets for more than 100 years, yet 2.5 billion people still lack basic sanitation (Niemeier et al., 2014). Indeed, more of the world's population has access to cell phones than to basic sanitation facilities (Niemeier et al., 2014). Yet, despite the ubiquity of medical technology, it was not until 2007 that the World Health Organization (WHO) issued the first global directive on medical devices (Sinha & Barry, 2011).

In achieving global health goals, improving equitable access to technical progress is just as important as the technical progress itself. In 2013, the Commission on Investing in Health called for a reduction in preventable infectious disease, maternal, and child deaths to universally low levels by 2035 (Yamey & Morel, 2016). Modeling shows that scaling up existing health technologies will bring low-income countries only two-thirds of the way to this "grand convergence in global health"; full convergence can be achieved only through the development and deployment of new health technologies (Yamey & Morel, 2016). These predictions are consistent with historical evidence of the benefits of adopting innovation in low-resource settings; data show that countries that rapidly adopt new technical innovations reduce their child

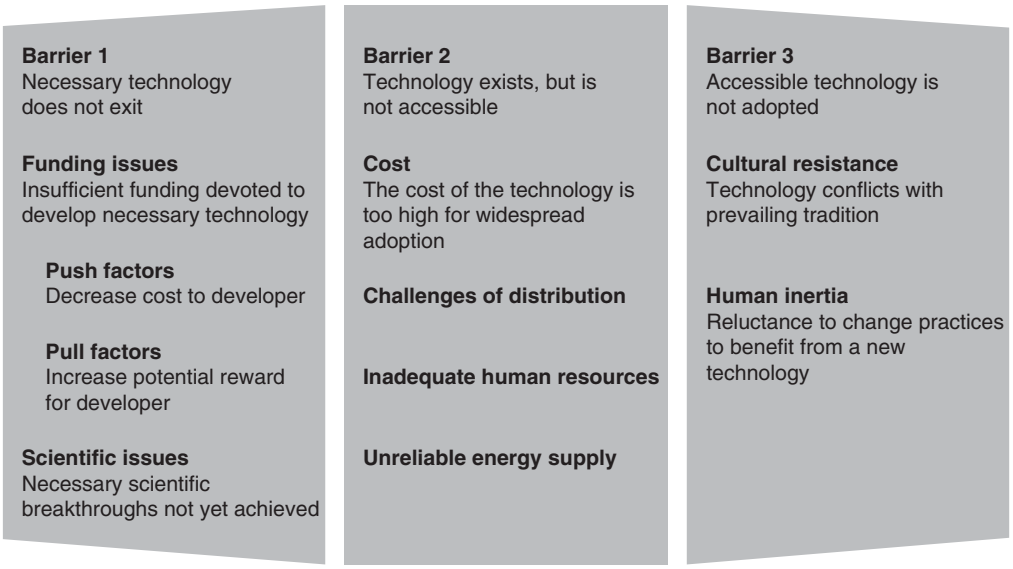


FIGURE 17-2 Three fundamental barriers to the use of technology to improve global health.

Reprinted from Howitt, P., Darzi, A., Yang, G. Z., Ashrafian, H., Atun, R., Barlow, J., . . . Wilson, E. (2012). Technologies for global health. *The Lancet, 380*(9840), 507–535. Copyright © 2012, with permission from Elsevier.

mortality rates by an additional 2% per year compared to non-adopting countries (Yamey & Morel, 2016).

FIGURE 17-2 summarizes the three major barriers to increased use of medical technology to improve global health—namely, the necessary technology (1) does not exist, (2) is not accessible, or (3) has not been adopted (Howitt et al., 2012). Where the necessary technology does not yet exist, successful innovation requires that innovators have a deep understanding of the local context in which the solution will be used (Sinha & Barry, 2011). This context dictates constraints that a successful solution must satisfy, including, for example, any limitations on cost, the supply chain for consumables, the available infrastructure to support a solution, and the educational level and number of personnel who will use and maintain the solution. These considerations are paramount for medical equipment.

Most medical equipment in many low-income countries consists of devices donated from high-resource settings (Howitt et al., 2012). Often, these donated devices fail because of line voltage incompatibilities, unavailability of consumables and spare parts, and harsh environmental conditions (Howitt et al., 2012); studies show that as much as three-fourths of donated devices do not work (Richards-Kortum & Oden, 2013). Thus, there is an important need for medical technologies that are explicitly designed to meet the needs of low-resource settings, with full understanding of the constraints that the technologies must deal with. The process of technical innovation discussed later in this chapter is designed to guide innovators through a process that addresses barriers to access and adoption through contextually specific needs finding, as well as development of locally relevant business and marketing strategies.

Society as a whole must value and foster innovation if all people are to reap its benefits. Because economic disparity is the primary driver of health disparity, where innovation takes place will play an important role in determining whether the gap of health disparity widens or closes. There is a great need for innovation leaders who are willing to engage in solving global health challenges, not only to create solutions for today's challenges, but also to train the leaders who will solve the challenges of the future.

Technology Platforms Impacting the Global Health Landscape

Low-resource areas are often physically separated from good healthcare facilities by poor and unreliable transportation networks. Furthermore, deep-rooted economic inequality leads to very low numbers of healthcare workers per capita. As such, global health problems have a number of unique needs, which are discussed throughout this chapter. Platforms that alleviate human resources requirements and other economic constraints can significantly impact the global health landscape. Here, we highlight three such platforms: mobile health technologies, point-of-care diagnostic technologies, and improved access technologies.

Mobile Health Technologies

Nearly 90% of the world's population now has access to wireless phone service; building on this opportunity, many efforts are under way to use mobile phone technology to address health challenges (mHealth). mHealth approaches have shown promise in a variety of low- and high-income settings, ranging from helping patients in Toronto, Canada, reduce their blood pressure and monitor blood glucose levels to reducing the cost of influenza surveillance in Kenya (Hampton, 2012).

A variety of mHealth applications have been developed to help community health workers provide improved services with a particular focus on maternal and child health, human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS), and sexual and reproductive health (Braun, Catalani, Wimbush, & Israelski, 2013). mHealth applications have enabled community health workers to collect high-quality health data in the field with fewer errors and less data loss compared to paper records (Braun et al., 2013). Moreover, mHealth tools provide real-time access to reminders and decision support tools, facilitating compliance with clinical care guidelines (Braun et al., 2013). For example, a study in Tanzania showed that midwives who received reminders about past-due patient visits by cell phone improved the number of timely visits to pregnant women (Braun et al., 2013).

In recent years, pharmaceutical companies and other medical research institutions have begun utilizing mHealth to conduct large-scale clinical research (Check Hayden, 2016). Both the iOS and Android communities have launched open-source projects to facilitate creation of clinical research apps for specific diseases. For example, the mPower study team launched an app for people with Parkinson's disease and enrolled more than 6,800 participants—three times more participants than were included in the previous largest Parkinson's disease study (Check Hayden, 2016). The mPower app transforms the smartphone's built-in accelerometer and microphone into a measurement tool for evaluating the steadiness of participants' gait and speech, respectively (Check Hayden, 2016). **FIGURE 17-3** compares mHealth research to standard clinical research and highlights some of the advantages and disadvantages of each approach. Although initial app-based mHealth studies have provided a proof of concept for large-scale clinical trials, most studies have encountered a decline in participant engagement over time and have patient demographics that are biased toward those who own a smartphone (Kvedar & Fogel, 2017).

Smartphone-linked wearable sensors have the potential to significantly expand the impact of mHealth applications (Steinhubl, Muse, & Topol, 2015). As shown in **FIGURE 17-4**, a wide variety of wearable sensors are under development, including tools to monitor electrocardiography (ECG), blood pressure, galvanic skin response, heart rate variability, and blood glucose (Steinhubl et al., 2015). Many of these sensors are produced using traditional microfabrication methods that support low-cost production at high volume. In addition, a number of new processes to develop flexible, wearable electronic sensors have recently been reported (Wang, Liu, & Zhang, 2017). While promising, the majority of these sensors are in early stages of development.

In addition to wearable sensors, a number of lab-on-a-chip mHealth technologies are in development. Many of them couple microfluidic devices that process small volumes of biological fluids together with cell phone–based readers, leveraging miniature light-emitting diodes (LEDs), optical elements, cell phone cameras, and smartphone screens to create low-cost sensors with sophisticated detection capabilities (Steinhubl et al., 2015).

	Advantages	Disadvantages
mHealth 	√ Significantly easier to recruit participants √ Lower cost to scale up and maintain √ Continuous data collection	x Requires software development expertise x Self reported data x Less clinician and patient engagement
Traditional clinical research 	√ More robust patient selection and data validation processes √ Lower cost and complexity to initiate √ More clinician and patient engagement	x Significantly more labor and time intensive x Limited geographic reach x Discrete data collection

FIGURE 17-3 Advantages and disadvantages of app-based mHealth versus standard clinical research.

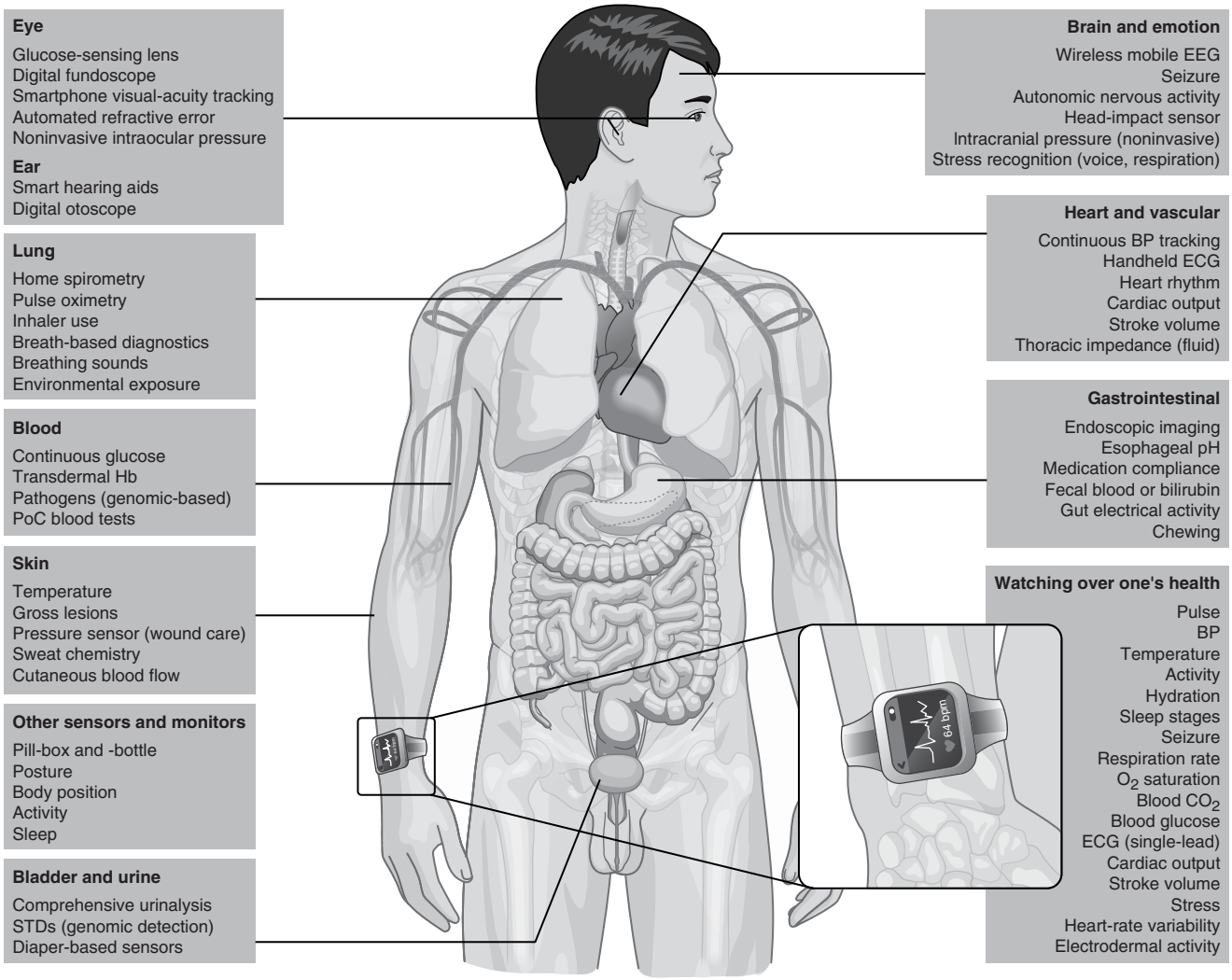


FIGURE 17-4 A wide variety of wearable sensors are in development for use in mHealth applications.

Reproduced from Steinhubl, S. R., Muse, E. D., & Topol, E. J. (2015). The emerging field of mobile health. *Science Translation Medicine, 7*(283):283rv283. Reprinted with permission from AAAS.

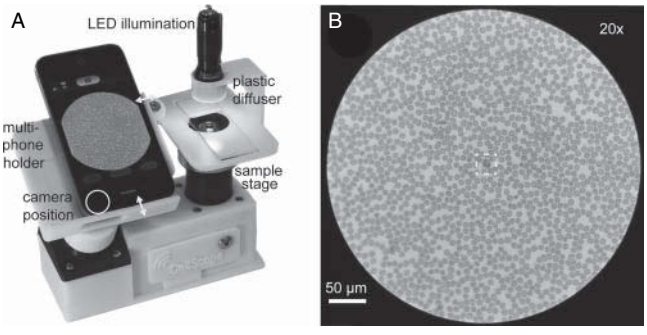


FIGURE 17-5 A. A mobile phone based microscope. **B.** An image containing a blue-stained granulocyte and surrounding red blood cells taken with the mobile phone microscope using an iPhone 4.

Reproduced from Skandarajah, A., Reber, C. D., Switz, N. A., & Fletcher, D. A. (2014). Quantitative imaging with a mobile phone microscope. *PLoS One, 9*(5), e96906. doi:10.1371/journal.pone.0096906, https://creativecommons.org/licenses/by/4.0/

Microscopes are one of the most commonly used pieces of equipment in medical labs; however, their cost and complexity often limit the utility of these devices in low-income countries (Miller et al., 2010). As shown in **FIGURE 17-5**, Fletcher and colleagues transformed a smartphone into a high-resolution microscope that could provide image in both bright-field and fluorescence modes, with comparable spatial resolution and image quality to much more expensive laboratory-grade microscopes (Breslauer, Maamari, Switz, Lam, & Fletcher, 2009; Skandarajah, Reber, Switz, & Fletcher, 2014). In more recent years, the smartphone-based microscope has been used to quantify levels of parasitic worms in an unprocessed patient blood sample, without need for an expert clinician to analyze the sample (D'Ambrosio et al., 2015).

The smartphone quantified the parasites by taking a video of the blood sample and detecting the motion of the live parasites in the blood, highlighting how increasingly powerful processors available in smartphones can be used for medical diagnosis.

To date, relatively few rigorous, well-controlled trials have evaluated the health impact of mHealth tools (Shuchman, 2014). Moreover, mHealth developers often use platforms that are not mutually compatible, limiting the technologies' potential impact in the field. Although small-scale efforts and pilot projects to measure program outcomes are increasing (Braun et al., 2013), significant gaps in large-scale, rigorous assessment remain. Finally, experts question whether existing mHealth applications provide sufficient utility to convince patients to use them; a recent survey of 1,700 mHealth projects found none that provided essential, actionable guidance for patients facing the acute health conditions that are common in low-resource settings (Royston et al., 2015). Despite these limitations, as smartphones continue to be adopted globally, smartphone-related innovation will likely play an increasingly important role in solving global health challenges.

Point-of-Care Diagnostic Technologies

Throughout history, medicine has typically involved testing within a laboratory of a well-resourced facility such as a hospital, clinic, or diagnostic laboratory. This often requires sending specimens off-site for analysis and waiting hours or days to learn the results. Because care must continue without diagnostic information in the meantime, this type of wait can have unfavorable results for patients. Furthermore, in low-resource settings where clinics are sparse, patients are often unable or reluctant to return to the clinic to receive their test results. In response to this need for more rapid testing, point-of-care diagnostics has developed as an area of research and innovation in the past several decades. Point-of-care diagnostic tests are performed with few resources, and can provide results within a single patient visit. Tests can also be performed by minimally trained technicians, helping alleviate some of the problems created by the chronic shortage of healthcare workers in many countries.

The lateral flow test is the simplest and most successful point-of-care diagnostic technology to date (Sharma, Zapatero-Rodriguez, Estrela, & O'Kennedy, 2015). This technology is based on a series of porous membranes that absorb and wick a fluid laterally. When reagents are included within these membranes, a reaction takes place as the fluid wicks along, and visible markings appear. Typically, a lateral flow test includes one control line that indicates that the test worked properly, and one or more test lines that indicate the presence or absence of an analyte or analytes of interest. One well-known application of lateral flow technology is the home pregnancy test, which functions by capturing the pregnancy marker human chorionic gonadotropin (hCG) from a urine sample (Cole, 2012). Based on technology developed in the 1950s, lateral flow assays helped decentralize laboratory testing in high-resource settings in the 1970s (Sharma et al., 2015). In low-resource settings, lateral flow assays have primarily been used to detect infectious diseases, including HIV and malaria (Sharma et al., 2015).

Adult HIV diagnostics have been one of the greatest global health successes of the past several decades. Whereas someone who acquired HIV in the pretreatment era could expect to live only 12.5 years, a young person in an industrialized country who becomes infected today can expect to live a near normal lifespan (Collaborative Group on AIDS Incubation and HIV Survival, 2000; Samji et al., 2013). Research suggests that comparable results are achievable in low-resource settings as well, due to two factors: (1) the advance of point-of-care diagnostics for HIV and (2) a massive global scale-up of access to life-saving antiretroviral therapy.

In the early 2000s, point-of-care tests to detect HIV-related antibodies and proteins enabled rapid HIV diagnosis. After reaching a peak in 2005, the number of AIDS-related deaths declined from approximately 2 million per year to the current 1.1 million per year (**FIGURE 17-6**). HIV testing continues to become more convenient as HIV self-testing increases in popularity (Hurt & Powers, 2014). The impact of these simpler diagnostics would have been minimal, however, if critical treatments remained difficult to access. After stakeholder consultations in all regions of the world, tremendous efforts were undertaken to reduce the cost of medications and improve ease of access in remote areas. As a result, in the region most affected by HIV—that is, eastern and southern Africa—the number of HIV-positive people receiving treatment has more than doubled since 2010, reaching nearly 10.3 million people today (UNAIDS, 2016). AIDS-related deaths in this region have decreased by 36% during the same time period. The widespread successes seen so far with HIV illustrate what can be accomplished when technical innovation is paired with access pathways that put the innovations in the hands of the users most in need. According to the United Nations Programme on HIV/AIDS, the quest to end the AIDS epidemic "will inspire broader global health and development efforts, demonstrating

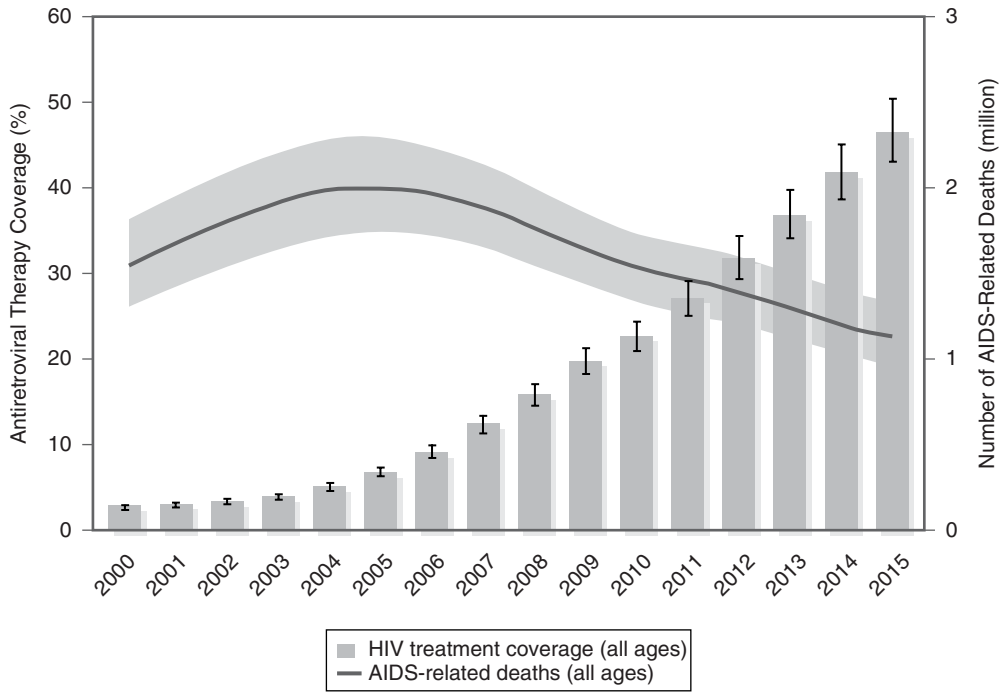


FIGURE 17-6 Antiretroviral therapy coverage and number of AIDS-related deaths, global, 2000–2015.

Reproduced from UNAIDS. (2016). Global AIDS update 2016. Retrieved from http://www.unaids.org/sites/default/files/media_asset/global-AIDS-update-2016_en.pdf

what can be achieved through global solidarity, evidence-based action, and multi-sectoral partnerships" (UNAIDS, 2017).

The development of point-of-care diagnostics for malaria has also dramatically transformed the global health landscape. Traditionally, malaria was diagnosed by direct observation of malaria parasites within peripheral blood smears under microscopy (Wilson, 2012). However, many settings do not have access to good-quality microscopy services. Rapid diagnostic tests for malaria take the form of a typical lateral flow test, which captures antibodies produced in response to the disease. Implementation of these tests has been widespread, with manufacturers reporting a total of 270 million sales of malaria rapid diagnostic tests in 2015 (WHO, 2016b). Most of these sales are in Africa, where the disease burden of malaria is highest. Malaria mortality rates are estimated to have decreased by 29% globally between 2010 and 2015 (WHO, 2016b). In the WHO African Region, reduced mortality due to malaria alone added 1.2 years to average life expectancy at birth between 2000 and 2015.

The last decade has seen rapid progress in the development of novel technologies that can provide molecular diagnostics at the point of care in low-resource settings, moving from lateral flow assays to instrumented systems for blood chemistry to detection of nucleic acids (Sharma et al., 2015) (**FIGURE 17-7**). For example, Cepheid's GeneXpert provides sample-to-answer nucleic acid testing for *Mycobacterium tuberculosis* (MTB) and drug resistance in a fully automated platform that can be used with minimal technical expertise (Lawn et al., 2013). The assay has better sensitivity than microscopy, although it is not as sensitive as liquid culture. A four-module GeneXpert platform costs approximately $17,000 and, with subsidization from a variety of funders, each single test cartridge costs $9.98, significantly more than the per-test cost of microscopy (Lawn et al., 2013). Even at this higher cost, initial analyses suggest the GeneXpert assay will be cost-effective in settings with a high burden of tuberculosis (TB). However, its high relative cost, sophisticated hardware, and need for infrastructure (wall power and climate-controlled environment) are likely to limit implementation of this technology (Lawn et al., 2013).

Despite the promise of the currently available affordable, sensitive point-of-care diagnostics, a number of important unmet needs persist. Indeed, the 2013 Ebola epidemic highlighted the need for continued development of point-of-care diagnostic technologies. During the epidemic, diagnosis was performed in the laboratory using polymerase chain reaction (PCR) at a cost of approximately $100 per test. Although PCR requires only 2–6 hours to perform, test delays of up to 1 week were reported at the peak of the epidemic, likely contributing to increased risk of transmission (Nouvellet et al., 2015). In late 2014, WHO called for rapid, simple, and sensitive Ebola diagnostics; four tests have been approved that detect Ebola-specific RNA or virus

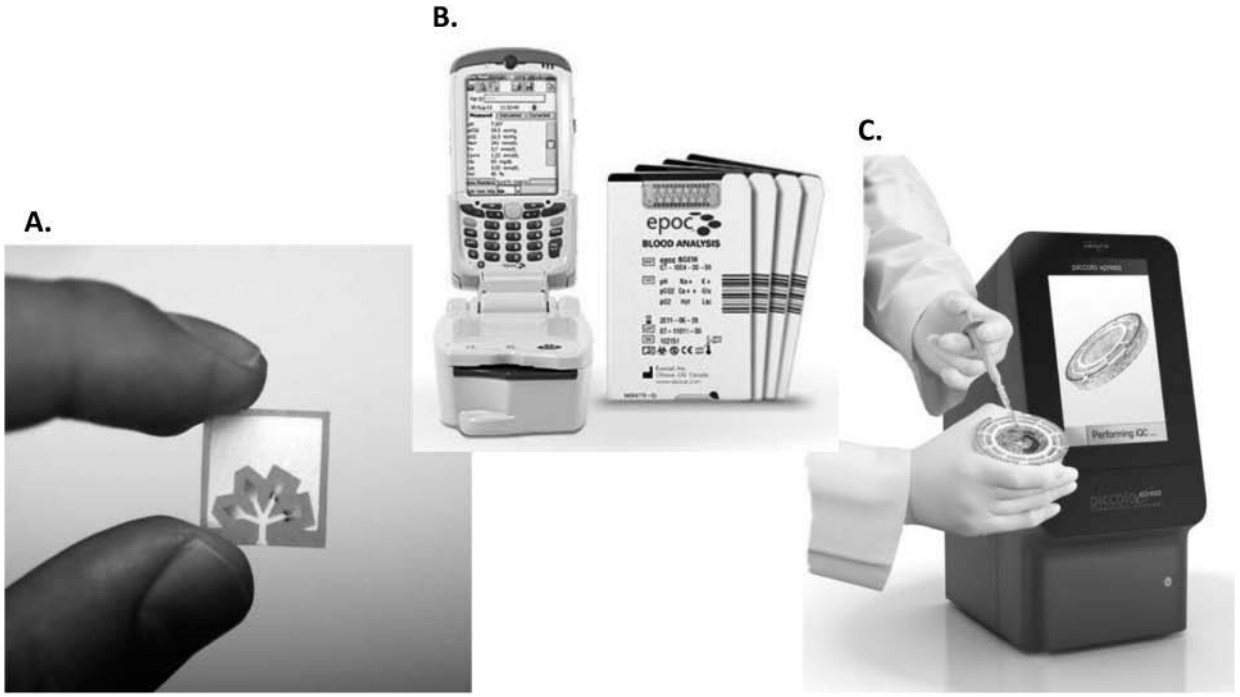


FIGURE 17-7 Examples of point-of-care diagnostic devices under development. **A.** Two-dimensional paper network to quantify glucose and other analytes in urine. **B.** Lab-on-a-chip system to profile blood chemistry developed by Alere. **C.** Lab on a disk device developed by Abaxis.

antigen. Although antigen-based rapid tests typically have lower sensitivity and specificity compared to PCR, modeling shows an approach using rapid tests for screening together with confirmatory PCR for those patients who test positive could have reduced the scale of the 2013 epidemic by more than one-third (Nouvellet et al., 2015). Further technical innovation is needed to support Ebola testing at the point of care; the newly approved tests require cold storage, high biosafety levels, and, in some cases, access to a PCR machine.

Improved Access Technologies

> Imagine the challenge of getting a vaccine from a pharmaceutical company to a remote village in Africa. The vaccine leaves by truck, is flown to Africa, and then is carried across dirt roads by truck or bicycle to eventually reach a refrigerator in a rural clinic that has sporadic electricity. The journey may take days or weeks, during which the vaccine is constantly at risk of being exposed to too much heat, sapping its potency.

This scenario, described by the Program for Appropriate Technology in Health (PATH), illustrates why technologies that improve access are among the most important global health innovations.

Vaccines can be damaged when the cold chain is interrupted during delivery or storage, causing heat-sensitive proteins to denature (WHO, 2012). The stability of a vaccine must be determined through testing, and careful handling must ensure that the vaccine's potency is preserved. The general temperature sensitivity of vaccines has resulted in cold-chain requirements for vaccines. However, in low-resource settings where vaccines must be removed from the cold chain for the last step of distribution, another monitoring method was developed. The vaccine vial monitor (VVM) is a technology that measures cumulative exposure to heat on vaccine vials. A colored disk indicates to health professionals whether a vaccine is still effective and can be used, or whether it should be discarded. Since their introduction in 1996, VVMs have become a critical part of vaccination efforts (Eriksson, Gessner, Jaillard, Morgan, & Le Gargasson, 2017). VVMs have allowed low- and middle-income countries (LMICs) to increase the availability of life-saving vaccines to their populations and reduce under-5 mortality (Eriksson et al., 2017).

Another example of a technology that improves access to critical health care is the Pratt Pouch. During childbirth, the transmission of HIV from mother to child is a significant risk, but can be prevented if the infant receives antiretroviral medication shortly after birth. Antiretroviral therapy has been shown to

decrease the risk of mother-to-child transmission from 21% to less than 7% (Torpey et al., 2010). However, a large percentage of mothers in sub-Saharan Africa deliver at home, and it is difficult to keep antiretroviral drugs outside of the clinic due to their short shelf life. Without treatment, millions of children are at risk of contracting HIV during delivery or breastfeeding.

In response to this problem, the Pratt School of Engineering at Duke University developed a packaging method that extends the life of two common antiretroviral medications by as much as 12 months (Choy, Ortiz, & Malkin, 2015). The Pratt Pouch is made of polyethylene, a heat-stable material that has been used in the food industry for many years (Malkin & Howard, 2012). After heat-sealing, this choice of material preserves the shelf life of the contents far longer than a foil pouch would (Malkin & Howard, 2012). The Pratt Pouch resembles a fast-food ketchup pouch in appearance, and provides a mother with a 14-day to 6-week supply of pre-dosed medication pouches. At home, a caretaker simply tears the pre-dosed package and delivers the drugs orally to the infant. Mothers studied in clinical trials reported that the device was easy to use, and access to medication for HIV-exposed infants born outside of a health facility increased from 35% before the introduction of the pouch to 94% after (Dahinten & Malkin, 2016).

Often, cost is the primary barrier to implementation of technologies in low-resource settings. In response, academic and commercial groups have sought to build medical tools using the cheapest materials possible. One example of an ultra-low-cost tool is the Foldscope, an origami-based paper microscope developed at Stanford University (Cybulski, Clements, & Prakash, 2014). The Foldscope is constructed from three pieces of paper, a spherical ball lens, lens-holder apertures, an LED with a diffuser or condenser lens, a battery, and an electrical switch (**FIGURE 17-8**). Early tests have shown that the device can show bright-field images of several organisms that cause serious infections, including *Giardia lamblia* and *Leishmania donovani*, as well as the commonly studied *Escherichia coli*. A low-magnification version (140×) costs only $0.58, while a high-magnification version (2,180×) costs $0.97.

The Foldscope is also extremely rugged, self-aligning itself during folding, and using components that can survive being stepped on. Taken together, these qualities give the device the potential to be applied not only to diagnostics, but also to education. Many children around the world, even in developed countries, have never had the chance to use a microscope due to these devices' cost and complexity. The Foldscope has the potential to expand science/technology/engineering/mathematics (STEM) education, fostering curiosity and early interest in science that could provide a more sustainable model for global health.

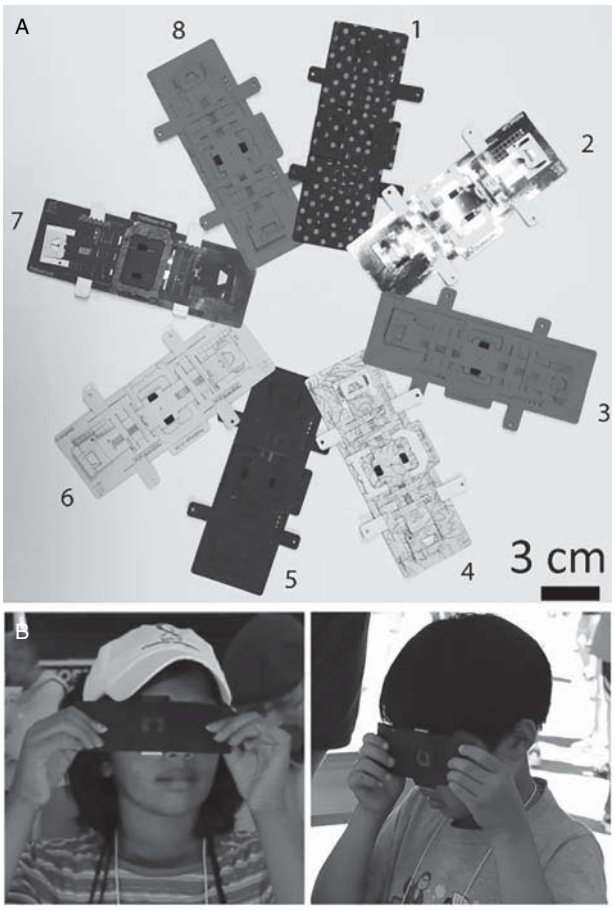


FIGURE 17-8 **A.** Multiple Foldscopes constructed out of different colored cardstock. **B.** Novice users using the Foldscope.

Mobile health, point-of-care, and improved access technologies each address a specific difficulty of delivering health care to remote settings. Furthermore, each of the technologies within the platforms highlighted in the preceding discussion was developed in response to a specific global health need. To be as effective as possible, the technology was developed with the end user in mind, and multiple iterations were, or are still, being produced. In the next section, we investigate the steps typically involved in the innovation process for global health problems.

Global Health Innovation Processes

How do innovators identify unmet clinical needs and begin to develop new technologies that meet the needs of patients and healthcare providers? How do

implementers take a promising innovation from an early-stage prototype to a commercial product that is delivered at scale? Technical innovation can be divided into four phases: (1) identifying needs; (2) solution development; (3) business plan creation; and (4) scaling up. In the next sections, we consider each step in the innovation process in more detail.

Identifying Needs

Innovators from high-income countries are often not familiar with the unique needs of LMICs, which are home to the majority of the world's population. The tendency of many innovators is to assume that the needs of LMICs are similar to the needs of high-income countries, but with the additional constraint of cost. Many multinational companies have failed in their attempts to penetrate these markets because they do not take the time to study differences in market dynamics (Khanna, Palepu, & Sinha, 2005). LMICs represent a diverse set of needs, many of which are unique to the particular country and do not involve cost constraints.

To develop sustainable, impactful technologies in global health, innovators spend a significant amount of time identifying and analyzing market needs prior to generating concepts. Importantly, the needs-finding process is solution-agnostic, and innovators need to resist the temptation to jump into solutions before needs are studied closely. A common cause of failure of new ventures is that there is no market need for the venture's product; to mitigate this risk, many methods have been developed to identify needs (Blank, 2013). The focus on market needs (a market-pull approach) instead of searching for technology applications (a technology-push approach) is particularly important in a global setting where the innovator may not fully understand the complexities of healthcare systems and cultures that are not her or his own.

As a first step toward understanding a complex healthcare system, the innovator can map the patient journey using a patient flow diagram; each step of the patient journey is captured as a block in the diagram (Yock et al., 2015). **FIGURE 17-9** shows an example of a patient's journey through the diagnosis of chronic kidney disease (CKD) and treatment of peritoneal dialysis in Costa Rica using a methodology described by Knapp, Zeratsky, and Kowitz (2016). The players (or stakeholders) in the process are listed on the left, and the desired outcome on the far right. This example shows the one successful pathway the patient can take; more complex patient flow diagrams have decision points and other bifurcations that represent several possible pathways for the patient. It is often helpful to map the most common patient pathway at the start of the needs-finding process, then continue to add branches as alternative pathways are discovered through additional research or observation.

While the patient is of primary importance, equally important to a sustainable innovation is the flow of money through the healthcare system (Yock et al., 2015). A money flow diagram can be created to

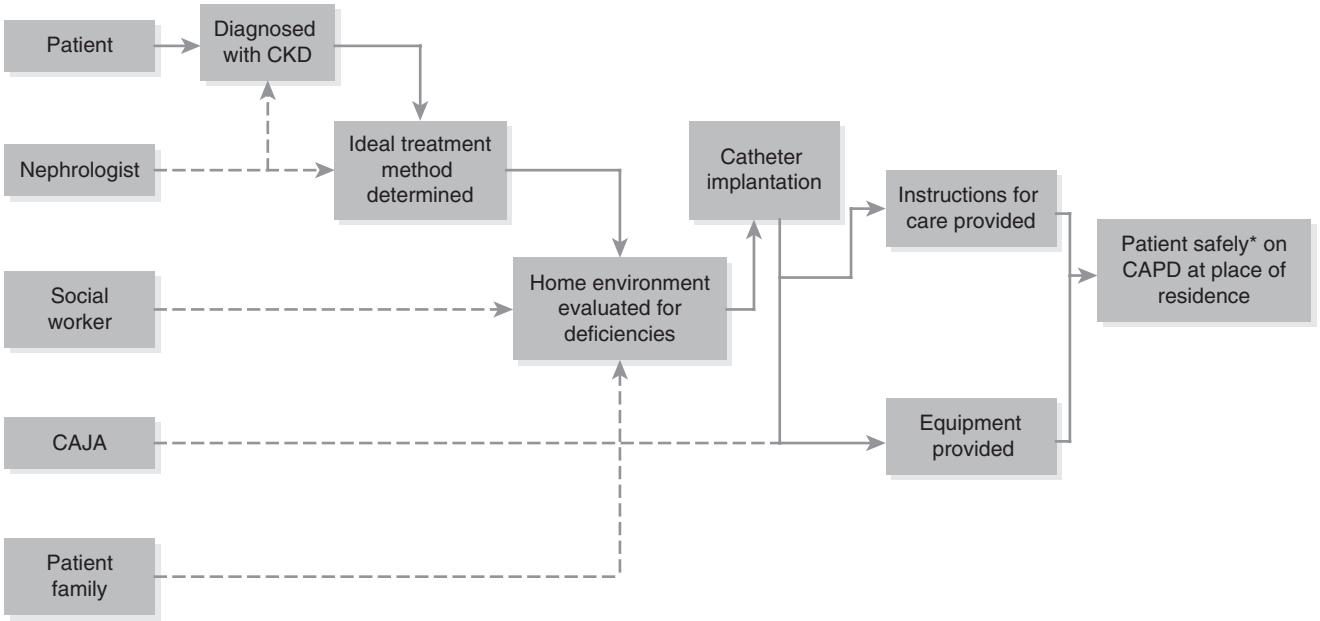


FIGURE 17-9 A flow diagram for a patient with chronic kidney disease, who is treated with peritoneal dialysis in Costa Rica.

Abbreviations: CAJA = Costa Rican Department of Social Security, which is primarily responsible for public health care; CAPD = continuous peritoneal ambulatory dialysis.
The diagram was developed by a team of graduate students at Rice University.

show where, when, and how money is exchanged in the healthcare system. These diagrams may be equally (or even more) complex as patient flow diagrams. If the innovator is not familiar with a country's healthcare payment structure, she or he may find it helpful to create separate diagrams for the public and private sectors, as well as separate diagrams for different phases of care (e.g., screening, diagnosis, and treatment).

Both patient and money flow diagrams will indicate stakeholders in patient diagnosis and care. Stakeholders can be conveniently categorized into patients, providers (such as physicians and nurses), payers (such as insurance companies or government institutions), and facilities (such as hospitals and clinics). While a formal stakeholder analysis is typically done later in the process, it is helpful to assess the amount of stakeholder need in different phases of care to focus the needs-finding efforts.

Once a potential area of interest has been identified, innovators can use a variety of techniques to gather as many needs as possible. As with other divergent phases of product development (such as brainstorming concepts), quantity is key to finding quality. It is not uncommon for a team to find hundreds, and sometime thousands, of needs before narrowing its options down to one choice that continues into the concept generation phase.

Needs-finding strategies that are particularly helpful in global health can be organized into three categories: observation, interviewing, and secondary research. A mix of these three strategies used in parallel is likely to yield the best results. Other forms of primary market research, such as questionnaires, focus groups, video ethnography, and competitive analysis, may be useful in the context of global health, but are beyond the scope of this chapter.

Direct observation of human behavior, or ethnographic research, is an excellent way to uncover needs that a stakeholder or user may not be able to articulate in an interview. It also serves as an effective way to develop empathy for the patient or provider. Efforts should be made to observe the work of all stakeholders, including nurses, administrators, supporting hospital staff, and patient families. If possible, stakeholders should be observed in as many stages of care as possible to get a holistic view of the experience and to validate a patient flow map. During the observation process, the innovator looks carefully for "systems" problems such as bottlenecks, inefficiencies, and cost drivers.

Interviews should likewise be conducted with as many stakeholders as possible. As described in conjunction with the contextual inquiry method (Privitera, 2015), it is often best to follow an observation session with interviews to confirm the understanding of what occurred in the clinical setting, ask clarifying questions, and uncover the reasons for certain behaviors. While interview techniques have been widely described in the literature (Burchill & Brodie, 1997; Privitera, 2015; Yock et al., 2015), the following technique may be especially helpful in needs-finding interviews:

1. Encourage storytelling to get specific examples and details.
2. Ask open-ended questions that cannot be answered by a simple "yes" or "no."
3. Continue to ask "why" until you can get to the root cause of a problem.
4. Ask people to discuss problems that their colleagues are facing, as it is often easier to describe others' challenges rather than your own challenges.

It is helpful to schedule time immediately after the observation or interview to analyze the findings. Indeed, it is not uncommon for a team to spend an hour of analysis for every hour of clinical observation.

Causes or problems identified through this process can be translated into need statements. A good need statement defines a problem, population, and outcome (Yock et al., 2015). The need can then easily be articulated with a statement of the following form: "A way to (problem), in (population), in order to (outcome)." For example, a need statement might be "A way to access the pericardium without ultrasound in rural patients with acute cardiac tamponade in order to quickly and safely remove fluid."

To ensure that a need statement is not too general or too broad, the innovator can scope the need up (make it more general) or down (make it more specific) to confirm that the need is at an appropriate level of resolution. For example, the need statement given previously could be scoped up to "A way to treat cardiac tamponade in patients in order to restore their hemodynamic stability." It could be scoped down to "A way to quickly access the pericardium in the emergency department in patients with acute cardiac tamponade in order to quickly and safely remove 100 mL of fluids."

Likewise, all need statements should be checked for an embedded solution. An embedded solution prescribes the manner in which to solve the problem, unnecessarily constraining the innovator during concept generation. For example, a need statement that begins with "A catheter that . . ." has an embedded solution because it prescribes that a catheter must be used. Perhaps other solutions could address the need without the need of a catheter.

To select a small handful of needs to begin concept generation, a screening or prioritization process is helpful. This is done by ranking potential needs against a set of criteria developed by the team. After screening leads to a smaller group of needs, performing more observations can validate these needs and offer further insights.

Solution Development

Once a need has been identified, refined, and validated, it is appropriate to move on to the solution development stage of innovation. Throughout this section, we use examples from a global health design project completed by undergraduate students at Rice University to illustrate various steps of the design process (Brown et al., 2013) (**EXHIBIT 17-1**).

The first step of solution development is to carefully define design criteria and constraints that must be met (**EXHIBIT 17-2**). This is essentially a translation of the needs (of the customer) into the specifications (to be met by the designer). Design criteria should be stated in a manner that is measurable. For example, "The system must cost less than $200" is a measurable criterion, whereas "The system should be low cost" is not. Some important design criteria should become

EXHIBIT 17-1 BubbleCPAP: The Need

Half of all premature infants suffer from respiratory distress and need support to breathe in the first few days of life. Continuous positive airway pressure (CPAP) is a method of respiratory support used in neonates who struggle to breathe. CPAP maintains inspiratory and expiratory pressures above the ambient pressure to increase a baby's functional residual capacity and improve static lung compliance, thereby making it easier for an infant to breathe.

Overall goal: Design a CPAP machine that is safe, simple, easy to use, and less expensive to make than conventional CPAP machines.

EXHIBIT 17-2 BubbleCPAP: Design Criteria

After defining their problem, the CPAP team identified a set of design criteria and specified testable, quantifiable targets for their solution. This table shows the most important design criteria for CPAP.

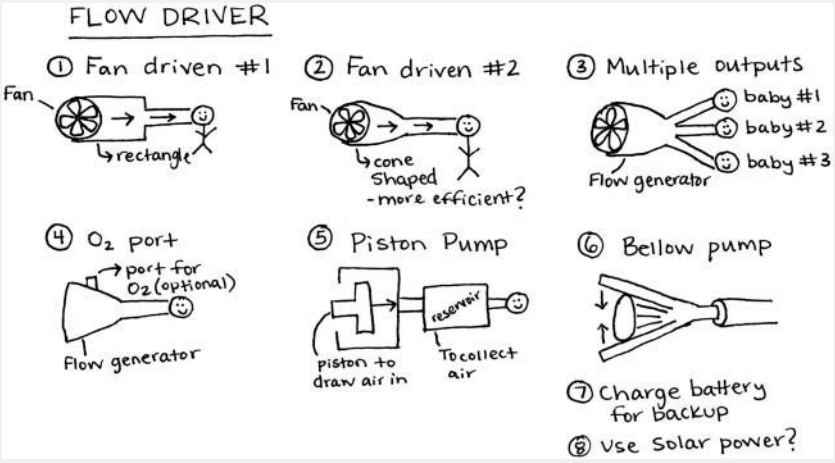
Criterion	Design Goal and Evaluation Method
Controllability	Ambient air flow must be adjustable over a range of 5–10 L/min and flow from optional O_2 port must be adjustable over a range of 1–10 L/min. CPAP must be adjustable over a range of 5–10 cm H_2O.
Safety	Flow from the generator must not exceed 10 L/min and CPAP must not exceed 10 cm H_2O. The device must alert the user if pressure in the circuit drops to atmospheric pressure.
Ease of use	The time required to set up the system on a patient for treatment must be less than 15 minutes.
Power requirements	The system must require a maximum of 240 V to operate and require less than 50 W.
Reusability/durability	The tubing, patient interface, and water bottle must be reusable for a minimum of 5 uses. The flow generator must function properly by producing pressure of 7 L/min after a 5-foot drop onto a concrete floor.
Sterilizability	After sterilization of the used device with 1:4 bleach solution, mechanical properties must remain unchanged.
Cost	The cost of goods, including the flow generator, CPAP circuit, and alarm, must be less than $200. Subsequent costs for tubing and the patient interface must be less than $10 per patient.

Cultural acceptance	The system must be considered culturally acceptable by greater than 80% of surveyed users (doctors, nurses, or other healthcare providers) in developing countries.
Portability	The total size of the system must not exceed 1 ft^3 and the weight must not exceed 5 kg.

EXHIBIT 17-3 BubbleCPAP: Brainstorming

During the brainstorming process, the CPAP team developed solutions for many different major components of their design. First, they brainstormed different methods of delivering CPAP and settled on a bubbleCPAP concept. Next, the team brainstormed concepts for tubing, patient interface, alarms and safety mechanisms, pressure control, and air flow drivers.

To make brainstorming more manageable, the team isolated the flow driver as a subset of a challenge as a topic for brainstorming rather than the whole challenge. The CPAP team knew that the system needed to pump room air at flow rates up to 10 L/min continuously. It was critical to identify or build a pump that was durable, yet low cost and easy to repair. Some of the resulting concepts from a brainstorming session on flow drivers are shown in the figure.



clear from the team's interviews and background research. It is important to note which criteria must be met to ensure success of the project and which may be flexible if tradeoffs are required. In addition, relevant regulatory standards for the category of product should be identified early in the process. A good place to start when seeking the relevant standards for a new product is to review the specific standards that are met by competing products.

Common design criteria used by global health innovators include the following:

- Cost
- Safety considerations
- Performance characteristics
- Ease of use
- Power requirements
- Size/portability
- Compatibility with existing systems or consumables
- Environmental and sustainability requirements
- Government regulations and standards
- Durability

When specifying the criteria, it is helpful to consider how they will be measured or tested. To demonstrate that the design meets all the goals, testing will be required.

After defining the design criteria, innovators brainstorm as many new ideas as possible (Ulrich & Eppinger, 2012). Building on patent searches, existing product reviews, expert consultations, and literature reviews that were previously been completed, the team can brainstorm specific solution concepts (**EXHIBIT 17-3**).

By considering the advantages and disadvantages of each brainstormed solution, it is possible to rank and identify the best options to move forward toward prototyping (**EXHIBIT 17-4**). Sometimes, final

EXHIBIT 17-4 BubbleCPAP: Ranking Conceptual Solutions

Knowing they wanted to focus on a piston pump, the CPAP team members identified three piston pump types and applied a concept scoring matrix to their ideas. The table shows this matrix, where the criteria are weighted and the design options are given a score of 0–5 for each criterion. The weighting and the score for each design and criterion were multiplied and summed to get a final score. Using this technique, the CPAP team determined that a diaphragm pump was the most promising type of piston pump for their need.

Design Criterion Specific to Pump	Weight	Piston Flow Drivers		
		Diaphragm	Tire Inflator	Nebulizer
Cost	25%	4	4	3
Durability	25%	4	3	3
Flow	15%	4	5	5
Pressure	15%	5	5	5
Power	10%	5	4	4
Portability	10%	5	4	4
Loudness	5%	5	2	3
Total Weighted Score		**4.35**	**3.95**	**3.75**

choices cannot be made with only the screening and scoring techniques. This can happen when there is uncertainty—for example, about whether a concept can really work—or where cost and feasibility are not yet determined. In that case, it may be advisable to move more than one concept forward to initial prototype development so that new and innovative concepts can be tested and evaluated.

Next, working prototypes of the most promising solutions are built (**EXHIBIT 17-5**). Successful prototypes aim to demonstrate the concept or concepts in the design solution. Sometimes, more than one option is prototyped for a certain component for comparison purposes. The prototype needs to be of sufficiently high fidelity that testing can be performed to evaluate whether it meets the design criteria. Often initial prototypes are instrumented with data acquisition systems so that data can be collected to demonstrate the function of the system. Later in the process, once design concepts have been validated, electronics can be downsized to a microcontroller to create a smaller, stand-alone form factor.

Prototypes are tested to evaluate how well they meet the design goals for a particular problem, and to identify areas for improvement. At all stages of solution development, the design and prototype should be constantly improved based on learnings from the previous steps. The first round of testing occurs in the laboratory, and after improvements, testing might progress to pilot clinical evaluations. To use the system with patients, the innovator must obtain ethics review board approvals at all institutions involved in the study. Striving to meet all standards in the design and collecting safety information in early laboratory-based studies can support later regulatory applications.

Finally, through this iterative process of design and evaluation, the innovator arrives at an appropriate solution for the clinical need. The innovator is now ready to consider a business plan for delivering the innovation to the user.

Business Plan Creation

A robust business plan is a key component of the strategy to bring a global health innovation to sustainable scale. A comprehensive business plan includes a thorough understanding of the relevant market and user needs, a plan for manufacturing and distribution, and

EXHIBIT 17-5 Bubble CPAP: Prototypes

To evaluate the CPAP prototype, the pressures and air flow from the CPAP system were compared to those of a commercial system in use at a local hospital. The frequency and magnitude of these parameters were the same, indicating that the prototype was able to deliver therapeutic benefit to a patient. At the same time, the prototype was demonstrated to nurses and doctors at numerous hospitals throughout Africa. Focus groups including clinical experts were asked to evaluate the system during trainings and simulations. Data on the user interactions with the device were collected. This led to a series of improvements that were incorporated into the subsequent design in preparation for the pilot clinical evaluation.

The figures show the progression of prototypes for the CPAP system. The first prototype (A) used a low-cost plastic shoebox as an enclosure for the initial device. Of course, this was not suitable as a system that would be used on a patient, but rather represented a quick solution to create a model that could undergo laboratory-based performance testing. The second prototype (B) used a commercially available container that was stronger and suitable for a clinic-ready system for pilot evaluation. Over time and in moving toward the commercial product (C), the CPAP designs become progressively more professional and incorporated improvements to the user interface and improved manufacturability, but the basic design concepts remained fairly unchanged.

A. Early conceptual prototype. **B.** Clinic-ready evaluation unit. **C.** Commercial product.

a plan to obtain regulatory approval from the appropriate governing bodies.

The first step in creating a business plan is to define the value proposition of the innovation—that is, a succinct summary of the value of the innovation in comparison to alternative solutions (Osterwalder, Pigneur, Bernarda, & Smith, 2014). Value is considered from the perspective of the patient, the provider, and the payer. If the proposed solution does not offer significant value to each of these stakeholders, it is unlikely to succeed at scale. In many cultures, society and tradition influence value in a way that it is difficult for innovators from other parts of the world to understand. Because culture often affects the uptake of a new product, it should be researched and considered when developing a business plan. Likely, effective interviews during the needs-finding stage will have illuminated some of these challenges.

As early as possible in the design process, innovators should develop a manufacturing strategy and understand the costs associated with actually producing the innovation. The cost comprises several components: materials, labor, and burden. Material costs include all materials that are used in the production of a single product, including those wasted (or scrapped) in production. The yield—that is, the material used in the final product divided by the total material used in the process—can play a significant role in medical device costs given the high standard of quality that is required. Labor will play a large role in the cost of medical devices that require significant assembly. Labor costs are typically projected through an hourly, per-person rate. They include the time of the person directly assembling or manufacturing the product, but exclude the labor of engineers or managers who support manufacturing (these are accounted for in burden). Burden accounts for the overhead associated with manufacturing. This includes facility costs, utility bills, maintenance costs, the salaries of engineers and managers who support the manufacturing line, consumables, and many other costs.

Medical devices are required to obtain regulatory approval and must meet specific standards before they can be sold for clinical use. These standards normally depend on the complexity and risk profile of the particular technology. Low-risk devices, such as tongue depressors and examination gloves, are subject to fewer requirements than complex, life-saving, or implantable devices such as heart valves. In addition, the specific regulations and regulatory approval

processes vary by country. The United States, Canada, Australia, Japan, and the European Union have fairly well-defined regulations and processes for obtaining regulatory approval (Yock, 2015). In contrast, other countries, including many in Africa, do not have a specific agency or well-defined regulatory process. This, of course, complicates efforts to import and distribute new medical technologies.

Identifying an appropriate regulatory pathway for a newly developed medical product is a key part of the global health innovation process. The chosen regulatory pathway will dictate the exact scope of any clinical validation required. In addition to noting the current state of regulation, good regulatory strategists consider trends that may affect regulation in the months or years before approval can be obtained. Indeed, regulation in most countries, especially those with large, emerging economies, is changing rapidly as the medical technology industry is globalizing. Several case studies of successful introduction of health technologies worldwide are presented in the USAID series "Idea to Impact" (Center for Accelerating Innovation and Impact, 2015).

Overall, the business plan plays an important role in the successful adoption of technical innovations by articulating the return on investment from both a public health perspective and a financial perspective, considering the incentives for all people involved in the purchase and use of the innovation. With the business plan laid out, focus then shifts toward introducing the innovation into the marketplace and scaling up its adoption.

Scaling Up

The final step in creating impact with a global health innovation is to scale up the technology to a broad audience (see the *Culture, Behavior, and Health* chapter for information on the diffusion of health innovations model). Achieving grand convergence in global health would have enormous global health impact, preventing 10 million deaths every year (Kruk et al., 2016). Unfortunately, the landscape of global health is littered with examples of innovations that are known to be effective but remain underutilized. For example, the scientific foundation of oral rehydration therapy (ORT) was developed in the 1960s, but its potential impact was not realized until it was incorporated into community health strategies in the 1980s (Frost & Reich, 2008); three decades later, however, significant coverage gaps persist in use of ORT. **TABLE 17-1** compares current coverage levels of representation maternal and child health interventions to levels needed to

TABLE 17-1 Examples of Scale-Up of Maternal and Child Health Interventions Required to Achieve Grand Convergence by 2035

Intervention	Current Coverage in Low- and Lower-Middle-Income Countries	Coverage Rate by 2035 That Will Be Required to Reach Convergence
Modern family planning methods	30%	50%
Skilled birth assistance in labor	65%	99%
Neonatal resuscitation	28%	84%
Pregnant women sleeping under an insecticide-treated bed net for malaria prevention	26%	100%
Treatment of malaria in pregnant women	55%	100%
Kangaroo care (skin-to-skin contact for the newborn)	4%	95%
Oral rehydration therapy (ORT) for childhood diarrhea	40%	99%

Reproduced from Kruk, M. E., Yamey, G., Angell, S. Y., Beith, A., Cotlear, D., Guanais, F., . . . Goosby, E. (2016). Transforming global health by improving the science of scale-up. *PLoS Biology, 14*(3):e1002360, https://creativecommons.org/licenses/by/4.0/

Note: Coverage is the percent of the population in need receiving the intervention.

achieve convergence. The gaps highlighted serve as a reminder that innovations do not deliver themselves (Yamey & Morel, 2016).

Why is it so challenging to bring successful innovations to scale? Innovators often underestimate the challenges associated of moving from successful pilot test to field evaluation to implementation at scale. For example, innovators developing point-of-care diagnostics did not anticipate the tremendous strain on the healthcare system associated with a move to suddenly decentralize testing from a few central labs to hundreds of point-of-care sites (Peeling, 2015). It can be difficult to secure financial support for scale-up; often, policy makers are more interested in supporting the introduction of new products than in sustaining and scaling access to existing affordable technologies (Piot, 2012). Funding mechanisms are often more likely to emphasize technical innovation over simplicity, resulting in complex and costly technologies that are difficult to scale (Richards-Kortum & Oden, 2013).

Fortunately, there is a precedent for successful scale-up in the global health arena: The scale-up of drugs and vaccines has been successful in recent decades. Since its inception 16 years ago, the Global Alliance for Vaccines Initiative (GAVI) has helped create a healthy global vaccine landscape through four guiding principles: (1) product development, (2) health systems strengthening, (3) financing, and (4) market shaping. All of these principles are critical and often unfilled needs for diagnostics and medical technologies in low-resource settings. Recognizing the success of GAVI, innovators have called for a Global Alliance for Medical Diagnostics Initiative (GAMDI) to coordinate the development, introduction, and utilization of medical diagnostics (Mugambi, Palamountain, Gallarda, & Drain, 2017).

TABLE 17-2 summarizes the factors associated with successful scale-up of innovations based on field work in East Africa (Yamey, 2011). Innovations that are simple and that are accompanied by robust technical policies are easier to scale (Yamey, 2011). Designs that minimize the use of consumables may be most sustainable in the face of weak supply chains (Richards-Kortum & Oden, 2013). Integrated packages of technologies that address focused clinical areas (e.g., surgery) may be easier to scale because they allow implementers to more efficiently navigate implementation barriers for the package as a whole rather than for one technology at a time (Richards-Kortum & Oden, 2013).

Others argue that rather than asking "How do you scale up?", innovators, especially those associated with nonprofit organizations, should begin early in the innovation process to define their endgame—that is, the role they tend to play in the overall solution once they have proved the effectiveness of their innovation (Gugelev & Stern, 2015). In studying nonprofit organizations that successfully scaled interventions, Gugelev and Stern identified six possible endgames, identifying those that are best suited to various types of innovations (**FIGURE 17-10**).

TABLE 17-2 Factors Associated with Successful Scale-Up of an Innovation

Innovation	■ Keep the innovation as simple as possible. ■ Develop robust technical policies and training materials for scale.
Implementers	■ Ensure strong leadership and governance. ■ Engage local implementers and stakeholders. ■ Partner with government and nongovernmental organizations.
Delivery strategy	■ Use a phased approach to scale-up. ■ Tailor scale-up to the local context. ■ Integrate innovations within the health system.
Context	■ Incorporate innovations in national policies. ■ Secure country ownership of innovations.

Produced for this book chapter by the authors at Rice University.

Technological Opportunities: Filling the Gaps in Global Health

While innovations have played a central role in improving health in high-resource settings, more than 50 years later these innovations remain largely unavailable in low-resource settings. Although there is no clear process whereby the global public health community sets an agenda for innovation (Piot, 2012), a number of high-priority areas of need have been identified. We conclude by reviewing four areas where medical device innovations are urgently needed: newborn health, maternal health, surgery, and the growing global burden of chronic disease.

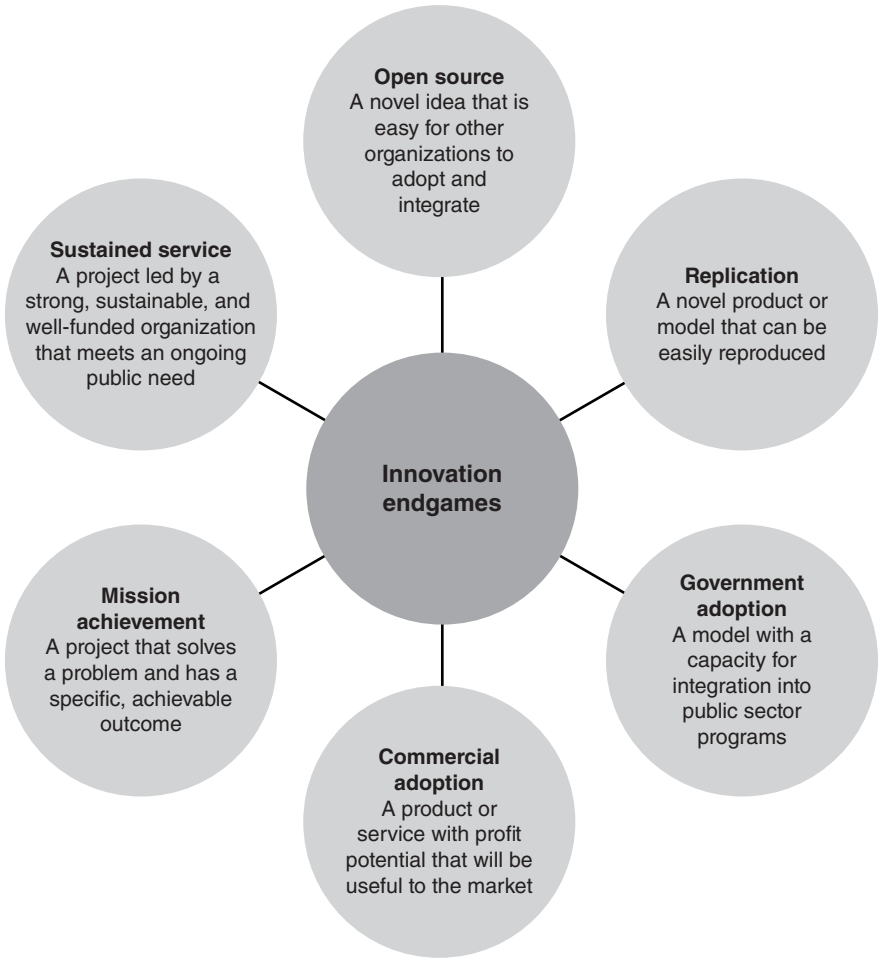


FIGURE 17-10 Six possible endgames for nonprofit organizations to consider as they scale new innovations.

Newborn Health

Since 1990, global child mortality has dropped by half. By comparison, neonatal mortality (death within the first 28 days of life) has declined much more slowly (see the *Reproductive Health* chapter for further detail). Newborn deaths now represent 45% of under-5 deaths (Lawn et al., 2014; Liu et al., 2012) and complications of preterm birth are the world's leading killer of children (Lawn & Kinney, 2014). In Africa, which has one of the highest neonatal mortality rates in the world, more than 1 million newborns die every year, the majority from preventable causes. Evidence shows that providing comprehensive hospital care during labor, delivery, and the first week of life, including for small and sick babies, can reduce newborn deaths by 75% (Bhutta et al., 2014; Kinney et al., 2015; Lawn, 2014; Lawn et al., 2013). Although most African mothers now deliver in hospitals, technologies to provide comprehensive care for newborns (defined as facility-based care with the infrastructure, technology, and skills to manage common complications especially for preterm newborns [WHO, 2016a]) are simply not available in most African hospitals (Bhutta et al., 2014), the region of the world where progress to improve newborn survival is slowest (Opondo et al., 2009).

FIGURE 17-11 shows the three main causes of neonatal death, which can be prevented or treated through seven pathways of care: (1) provide hydration and nutrition; (2) provide temperature stability; (3) monitor and treat jaundice; (4) monitor and treat hypoglycemia; (5) monitor labor and delivery; (6) provide breathing support; (7) prevent and treat infections. Figure 17-11 also identifies technologies required to accomplish these seven functions; these technologies are a high priority for innovation efforts and would enable a low-resource facility to address the majority of causes of neonatal death (Maynard et al., 2015). Improvements in obstetric care are also needed to help prevent birth asphyxia (Engmann, Khan, Moyer, Coffey, & Bhutta, 2016).

In the last 25 years, maternal mortality have dropped by half in high-income settings; in contrast, maternal mortality rates in sub-Saharan Africa have plateaued (Engmann et al., 2016). Using expert consultation together with a 22-criteria decision matrix,

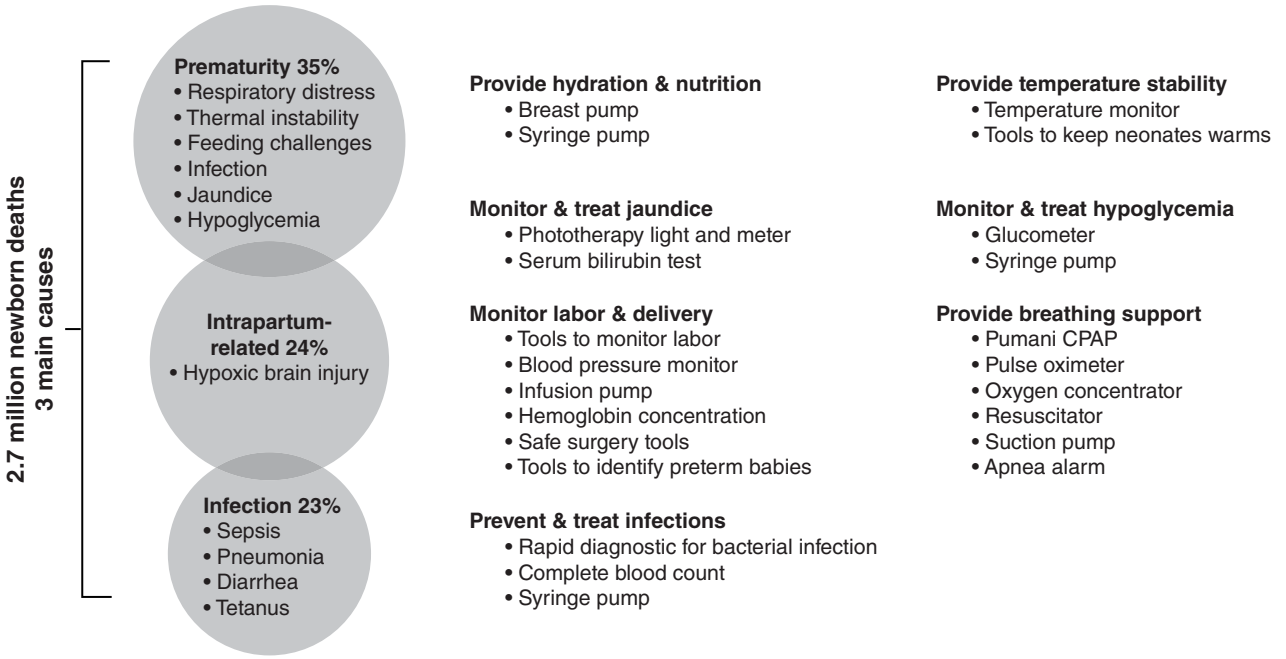


FIGURE 17-11 Leading causes of neonatal death and their corresponding clinical care pathways and management tools.

PATH identified five innovations with high potential to improve maternal health:

- Technologies to improve use of oxytocin to prevent and manage postpartum hemorrhage
- Uterine balloon tamponade to control severe postpartum bleeding
- An improved proteinuria test to diagnose preeclampsia
- Simplified dosing of magnesium sulfate to treat preeclampsia and eclampsia
- Better methods of blood pressure measurement validated for use in pregnant women (Herrick et al., 2014)

TABLE 17-3 details the assessment criteria used to evaluate technologies to address postpartum hemorrhage, preeclampsia, and eclampsia; the criteria span all the steps of the innovation and scale-up process. Other priorities identified to improve maternal health in low-resource settings include point-of-care diagnostics to identify and differentiate bacterial and viral sepsis; biomarkers to identify which women are likely to experience premature labor; tools to help safely deliver blood products; heat-stable and inhalable forms of oxytocin; and low-cost infusion systems for delivery of magnesium sulfate (Engmann et al., 2016).

Chronic Disease

As prevention and treatment of infectious diseases has improved in low-resource settings, noncommunicable diseases (NCDs) have increasingly emerged as a cause of morbidity and mortality (Hunter & Fineberg, 2014). NCDs, including chronic diseases such as cardiovascular disease, diabetes, cancer, and chronic obstructive pulmonary disease (COPD), now account for two-thirds of all global deaths (Checkley et al., 2014). Eighty percent of these deaths occur in LMICs (Checkley et al., 2014), and there is an important need for innovations to identify patients at risk and provide appropriate, effective prevention and treatment strategies.

For example, mHealth applications have the potential to play an important role in efforts to help patients with cardiovascular disease self-manage their disease, especially in combination with community health worker outreach (Checkley et al., 2014). However, a recent review examining the impact of mHealth interventions for NCDs in LMICs found that only 20 of 1,274 studies published in this area evaluated the impact of the intervention in a randomized clinical trial; overall results of these 20 studies showed a positive but modest effect on processes of care and clinical outcomes (Beratarrechea, Moyano, Irazola, & Rubinstein, 2017).

COPD often results from exposure to indoor air pollution associated with cooking fueled by biomass. Better low-cost spirometers are needed for diagnosis of COPD; improved cookstoves that reduce particulate matter to acceptable levels are urgently needed (Checkley et al., 2014).

TABLE 17-3 Assessment Criteria for Technologies that Address Postpartum Hemorrhage and Preeclampsia and Eclampsia

Category	Criteria	Evaluation Approach (high, medium, low)*
Gap	Gap-filling potential for health	Evaluate what layers of the treatment continuum could most greatly reduce mortality (e.g., prevention, diagnostics, treatment), the strength of the available data, and the percentage of cases that could be managed
Technology performance	Clinical evidence (efficacy/effectiveness)	Superior to benchmark,** similar to benchmark, inferior to benchmark
	Safety (patient/healthcare worker)	Superior to benchmark,** similar to benchmark, inferior to benchmark
	Ease of use	Superior to benchmark,** similar to benchmark, inferior to benchmark
	Usage requirements (e.g., durability, shelf life, electricity, storage temperature)	Technology has two or fewer usage requirements, technology has between three and five usage requirements, technology has greater than five usage requirements
Enabling factors	Alignment with internationally recognized guidelines (e.g., the World Health Organization, the International Federation of Gynecology and Obstetrics, and the International Confederation of Midwives)	Technology is recommended by at least one organization, divergent opinions exist, technology is not recommended
	Donor financial support (product development or implementation)	Funding greater than US$5 M, between US$1 M and US$5 M, less than US$1 M (or no funding identified)
	Other nonfinancial support	Placeholder for other nonfinancial supporters that may not be captured elsewhere; not scored
	Acceptability profile	Broadly acceptable, mixed, acceptable in few geographies (South Asia and sub-Saharan Africa)
	Organizational capabilities	All key partners have experience with the technology, a subset of key partners has experience with the technology, none of the key partners has experience with the technology
Market analysis	Manufacturing costs	Superior to benchmark** (lower cost), similar to benchmark, inferior to benchmark (higher cost)
	Distribution system requirements (warehouse, cold chain, transportation factors)	Infrastructure for a distribution channel(s) exists and high likelihood of utilization, infrastructure for distribution channel(s) exists and moderate likelihood of utilization, distribution channel(s) does not exist or a low likelihood of utilization

	Manufacturing plan established	Capture manufacturing information including delays, hurdles, risk, and complexity; not scored
	Target setting (community, primary health care, hospital)	Capture setting(s) where the technology would be used; not scored
	Target provider/administrator	Community healthcare worker, nurse or midwife, physician
	Potential multiple markets (additional value to the healthcare system)	Additional uses identified (high), no additional uses (low)
	Technology readiness level (clinical/regulatory development)	Regulatory or commercialization, confirmatory, discovery, or exploratory
	Cost of clinical development	Less than US$5 M, between US$5 M and US$50 M, greater than US$50 M
	Clarity of regulatory/clinical pathways	Class I US Food and Drug Administration (USFDA) device, class II USFDA device, class III USFDA device or requires efficacy data (e.g., drugs)
Unique considerations	System requirements (disruption)	Low system requirements (training/infrastructure), modest system requirements, high system requirements
	Product bundling	No other technologies required for impact, likely need to bundle with one technology, likely need to bundle with two or more technologies
	Other	Placeholder for other unique considerations; not scored

*Detailed definitions are available at http://sites.path.org/mnhtech/assessment/tool/

**Benchmark based on WHO treatment guidelines or the best benchmark for that technology category.

Surgery

Surgical care is an important component of care for maternal health, many noncommunicable diseases, and injuries, yet until recently it had been a neglected component of the global health agenda (Meara et al., 2016). A recent *The Lancet* Commission on Global Surgery recommended investment to strengthen access to safe surgical and anesthesia care (Meara et al., 2016). **EXHIBIT 17-6** outlines 10 high-priority needs to enable surgical care; technical innovation is needed to meet many of these needs, including infrastructure needs, sterilization capacity, ensuring a safe and adequate blood supply, and better patient monitoring technologies that can be used in settings with limited healthcare personnel (Meara et al., 2016). Expert groups have called for an international device consortium to encourage innovation to support surgical care in low-resource settings (Meara et al., 2016).

Conclusion

While technology has already changed the landscape of global health, there remains significant untapped potential to apply the tools of innovation to address many unmet global health needs. The examples discussed in the areas of mHealth, point-of-care diagnostics, and improved access technologies illustrate both the promise of new innovations and the significant challenges in translating these new tools to the field. Innovators can maximize their chances for

EXHIBIT 17-6 Criteria for Provision of Safe Surgical and Anesthesia Care

1. Trained surgical provider
2. Trained anaesthesia provider
3. Infrastructure, equipment and supplies necessary to perform safe general anaesthesia, loco-regional anaesthesia, laparotomy, caesarean delivery, and treatment of open fracture (including, for example, electricity, water, personal protective equipment for staff, basic laboratories, and HIV-testing capabilities)
4. Decontamination and sterilisation capacity
5. Blodd supply that is safe and affordable (screened and cross-matched blood)
6. Drugs, including antibiotics, pain medicines, and anaesthetics (from the WHO Model List of Essential Medicines)
7. Nursing care, which includes a record of appropriate physiological observations
8. 24 h surgical cover with the ability to review and respond to a deteriorating patient
9. Quality-improvement processes, including audit of perioperative mortality
10. Risk assessment and operation planning for planned procedures

success by following a well-defined process to ensure that they consider the specific needs of users from the beginning and have a pathway to deliver their solution at scale. Multidisciplinary collaborations within the global health, technology, and business communities will be needed if we are to achieve the promise of innovation and close the gaps in global health.

Discussion Questions

1. How do innovation and technology help address the underlying reasons for global health inequities? Give examples.
2. It is commonly observed that medical equipment from high-income counties that is donated to low-income counties tends to collect dust and not be used. Why do you think this is the case? What could be done differently to avoid this problem?
3. What are examples of successful applications of mHealth? Propose two new applications for mobile devices in global health.
4. Point-of-care diagnostic technologies have had a significant impact on global health, but they are not without their drawbacks. What are some potential drawbacks of point-of-care testing?
5. Pick a country and a clinical area you are interested in working on. Which strategies would you employ to identify needs in that clinical area for that particular country?
6. Identify a medical technology used in high-income countries that could be used in a selected low-income country. Describe the barriers that might prevent the technologies to be used in the selected country.

References

Beratarrechea, A., Moyano, D., Irazola, V., & Rubinstein, A. (2017). mHealth interventions to counter noncommunicable diseases in developing countries: Still an uncertain promise. *Cardiology Clinics, 35*(1), 13–30. doi:10.1016/j.ccl.2016.08.009

Bhutta, Z. A., Das, J. K., Bahl, R., Lawn, J. E., Salam, R. A., Paul, V. K., . . . Walker, N. (2014). Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? *The Lancet, 384*(9940), 347–370. doi:10.1016/s0140-6736(14)60792-3

Blank, S. (2013). Why the lean start-up changes everything. *Harvard Business Review, 91*(5), 63–72.

Braun, R., Catalani, C., Wimbush, J., & Israelski, D. (2013). Community health workers and mobile technology: A systematic review of the literature. *PLoS One, 8*(6), e65772. doi:10.1371/journal.pone.0065772

Breslauer, D. N., Maamari, R. N., Switz, N. A., Lam, W. A., & Fletcher, D. A. (2009). Mobile phone based clinical microscopy for global health applications. *PLoS One, 4*(7), e6320. doi:10.1371/journal.pone.0006320

Brown, J., Machen, H., Kawaza, K., Mwanza, Z., Iniguez, S., Lang, H., . . . Oden, M. (2013). A high-value, low-cost bubble continuous positive airway pressure system for low-resource settings: Technical assessment and initial case reports. *PLoS One, 8*(1), e53622. doi:10.1371/journal.pone.0053622

Burchill, G., & Brodie, C. H. (1997). *Voices into choices: Acting on the voice of the customer.* Madison, WI: Oriel Inc.

Center for Accelerating Innovation and Impact. (2015). *Idea to impact: A guide to introduction and scale of global health innovations.* Washington, DC: USAID.

Check Hayden, E. (2016). Mobile-phone health apps deliver data bounty. *Nature, 531*(7595), 422–423. doi:10.1038/531422a

Checkley, W., Ghannem, H., Irazola, V., Kimaiyo, S., Levitt, N. S., Miranda, J. J., . . . Yan, L. L. (2014). Management of NCD in low- and middle-income countries. *Global Heart, 9*(4), 431–443. doi:10.1016/j.gheart.2014.11.003

Choy, A., Ortiz, M., & Malkin, R. (2015). Accurate dosing of antiretrovirals at home using a foilized, polyethylene pouch to prevent the transmission of HIV from mother to child. *Medicine (Baltimore), 94*(25), e1030. doi:10.1097/MD.0000000000001030

Cole, L. A. (2012). The hCG assay or pregnancy test. *Clinical Chemistry and Laboratory Medicine, 50*(4), 617–630. doi:10.1515/CCLM.2011.808

Collaborative Group on AIDS Incubation and HIV Survival. (2000). Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: A collaborative re-analysis. Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action. Concerted Action on SeroConversion to AIDS and Death in Europe. *The Lancet, 355*(9210), 1131–1137.

Cybulski, J. S., Clements, J., & Prakash, M. (2014). Foldscope: Origami-based paper microscope. *PLoS One, 9*(6), e98781. doi:10.1371/journal.pone.0098781

Dahinten, A. P., & Malkin, R. A. (2016). The Pratt Pouch provides a three-fold access increase to antiretroviral medication for births outside health facilities in southern Zambia. *Open Biomedical Engineering Journal, 10*, 12–18. doi:10.2174/1874120701610010012

D'Ambrosio, M. V., Bakalar, M., Bennuru, S., Reber, C., Skandarajah, A., Nilsson, L., . . . Fletcher, D. A. (2015). Point-of-care quantification of blood-borne filarial parasites with a mobile phone microscope. *Science Translational Medicine, 7*(286), 286re284. doi:10.1126/scitranslmed.aaa3480

Engmann, C. M., Khan, S., Moyer, C. A., Coffey, P. S., & Bhutta, Z. A. (2016). Transformative innovations in reproductive, maternal, newborn, and child health over the next 20 years. *PLoS Medicine, 13*(3), e1001969. doi:10.1371/journal.pmed.1001969

Eriksson, P., Gessner, B. D., Jaillard, P., Morgan, C., & Le Gargasson, J. B. (2017). Vaccine vial monitor availability and use in low- and middle-income countries: A systematic review. *Vaccine, 35*(17), 2155–2161. doi:10.1016/j.vaccine.2016.11.102

Fong, H., & Harris, E. (2015). Technology, innovation and health equity. *Bull World Health Organ*, *93*(7), 438-438a.

Frost, L. J., & Reich, M. R. (2008). *Access: How do good health technologies get to poor people in poor countries?* Cambridge, MA: Harvard Center for Population and Development Studies.

Gugelev, A., & Stern, A. (2015, Winter). What's your endgame? *Stanford Social Innovation Review,* 41–47.

Hampton, T. (2012). Recent advances in mobile technology benefit global health, research, and care. *Journal of the American Medical Association, 307*(19), 2013–2014. doi:10.1001/jama.2012.4465

Herrick, T. M., Harner-Jay, C. M., Levisay, A. M., Coffey, P. S., Free, M. J., & LaBarre, P. D. (2014). Prioritizing investments in innovations to protect women from the leading causes of maternal death. *BMC Pregnancy & Childbirth, 14*, 10. doi:10.1186/1471-2393-14-10

Howitt, P., Darzi, A., Yang, G. Z., Ashrafian, H., Atun, R., Barlow, J., . . . Wilson, E. (2012). Technologies for global health. *The Lancet, 380*(9840), 507–535. doi:10.1016/s0140-6736(12)61127-1

Hunter, D. J., & Fineberg, H. V. (2014). Convergence to common purpose in global health. *New England Journal of Medicine, 370*(18), 1753–1755. doi:10.1056/NEJMe1404077

Hurt, C. B., & Powers, K. A. (2014). Self-testing for HIV and its impact on public health. *Sexually Transmitted Diseases, 41*(1), 10–12. doi:10.1097/OLQ.0000000000000076

Khanna, T., Palepu, K. G., & Sinha, J. (2005). Strategies that fit emerging markets. *Harvard Business Review, 83*(6), 4–19.

Kinney, M. V., Cocoman, O., Dickson, K. E., Daelmans, B., Zaka, N., Rhoda, N. R., . . . Darmstadt, G. L. (2015). Implementation of the Every Newborn action plan: Progress and lessons learned. *Seminars in Perinatology, 39*(5), 326–337. doi:10.1053/j.semperi.2015.06.004

Knapp, J., Zeratsky, J., & Kowitz, B. (2016). *Sprint: How to solve big problems and test new ideas in just five days.* New York, NY: Simon and Schuster.

Kruk, M. E., Yamey, G., Angell, S. Y., Beith, A., Cotlear, D., Guanais, F., . . . Goosby, E. (2016). Transforming global health by improving the science of scale-up. *PLoS Biology, 14*(3), e1002360. doi:10.1371/journal.pbio.1002360

Kvedar, J. C., & Fogel, A. L. (2017). mHealth advances clinical research, bit by bit. *Nature Biotechnology, 35*(4), 337–339. doi:10.1038/nbt.3851

Lawn, J. E. (2014). The child survival revolution: What next? *The Lancet, 384*(9947), 931–933. doi:10.1016/s0140-6736(14)61057-6

Lawn, J. E., Blencowe, H., Oza, S., You, D., Lee, A. C., Waiswa, P., . . . Cousens, S. N. (2014). Every Newborn: Progress, priorities, and potential beyond survival. *The Lancet, 384*(9938), 189–205. doi:10.1016/s0140-6736(14)60496-7

Lawn, J. E., Davidge, R., Paul, V. K., von Xylander, S., de Graft Johnson, J., Costello, A., . . . Molyneux, L. (2013). Born too soon: Care for the preterm baby. *Reproductive Health, 10*(suppl 1), S5. doi:10.1186/1742-4755-10-s1-s5

Lawn, J. E., & Kinney, M. (2014). Preterm birth: Now the leading cause of child death worldwide. *Science Translational Medicine, 6*(263), 263ed221. doi:10.1126/scitranslmed.aaa2563

Lawn, S. D., Mwaba, P., Bates, M., Piatek, A., Alexander, H., Marais, B. J., . . . Zumla, A. (2013). Advances in tuberculosis diagnostics: The Xpert MTB/RIF assay and future prospects for a point-of-care test. *The Lancet Infectious Diseases, 13*(4), 349–361. doi:10.1016/s1473-3099(13)70008-2

Liu, L., Johnson, H. L., Cousens, S., Perin, J., Scott, S., Lawn, J. E., . . . Black, R. E. (2012). Global, regional, and national causes of child mortality: An updated systematic analysis for 2010 with time trends since 2000. *The Lancet, 379*(9832), 2151–2161. doi:10.1016/s0140-6736(12)60560-1

Malkin, R., & Howard, C. (2012). A foilized polyethylene pouch for the prevention of transmission of HIV from mother to child. *Open Biomedical Engineering Journal, 6*, 92–97. doi:10.2174/1874120701206010092

Maynard, K. R., Causey, L., Kawaza, K., Dube, Q., Lufesi, N., Maria Oden, Z., . . . Molyneux, E. M. (2015). New technologies for essential newborn care in under-resourced areas: What is needed and how to deliver it. *Paediatric International Child Health, 35*(3), 192–205. doi:10.1179/2046905515y.0000000034

Meara, J. G., Leather, A. J., Hagander, L., Alkire, B. C., Alonso, N., Ameh, E. A., . . . Yip, W. (2016). Global Surgery 2030: Evidence and solutions for achieving health, welfare, and economic development. *International Journal of Obstetric Anesthesia, 25*, 75–78. doi:10.1016/j.ijoa.2015.09.006

Miller, A. R., Davis, G. L., Oden, Z. M., Razavi, M. R., Fateh, A., Ghazanfari, M., . . . Richards-Kortum, R. (2010). Portable, battery-operated, low-cost, bright field and fluorescence microscope. *PLoS One, 5*(8), e11890. doi:10.1371/journal.pone.0011890

Mugambi, M. L., Palamountain, K. M., Gallarda, J., & Drain, P. K. (2017). Exploring the case for a global alliance for medical diagnostics initiative. *Diagnostics (Basel), 7*(1). doi:10.3390/diagnostics7010008

Niemeier, D., Gombachika, H., & Richards-Kortum, R. (2014). How to transform the practice of engineering to meet global health needs. *Science, 345*(6202), 1287–1290. doi:10.1126/science.1257085

Nouvellet, P., Garske, T., Mills, H. L., Nedjati-Gilani, G., Hinsley, W., Blake, I. M., . . . Ferguson, N. M. (2015). The role of rapid diagnostics in managing Ebola epidemics. *Nature, 528*(7580), S109–S116. doi:10.1038/nature16041

Opondo, C., Ntoburi, S., Wagai, J., Wafula, J., Wasunna, A., Were, F., . . . English, M. (2009). Are hospitals prepared to support newborn survival? An evaluation of eight first-referral level hospitals in Kenya. *Tropical Medicine & International Health, 14*(10), 1165–1172. doi:10.1111/j.1365-3156.2009.02358.x

Osterwalder, A., Pigneur, Y., Bernarda, G., & Smith, A. (2014). *Value proposition design: How to create products and services customers want*. Hoboken, NJ: John Wiley & Sons.

Peeling, R. (2015). Bringing diagnostics to developing countries: An interview with Rosanna Peeling. *Expert Review of Molecular Diagnostics, 15*(9), 1107–1110. doi:10.1586/14737159.2015.1081802

Piot, P. (2012). Innovation and technology for global public health. *Global Public Health, 7*(suppl 1), S46–S53. doi:10.1080/17441692.2012.698294

Pollock, N. R., Rolland, J. P., Kumar, S., Beattie, P. D., Jain, S., Noubary, F., . . . Whitesides, G. M. (2012). A paper-based multiplexed transaminase test for low-cost, point-of-care liver function testing. *Science Translational Medicine, 4*(152), 152ra129. doi:10.1126/scitranslmed.3003981

Privitera, M. B. (2015). *Contextual inquiry for medical device design.* Cambridge, MA: Academic Press.

Richards-Kortum, R., & Oden, M. (2013). Devices for low-resource health care. *Science, 342*(6162), 1055–1057. doi:10.1126/science.1243473

Royston, G., Hagar, C., Long, L. A., McMahon, D., Pakenham-Walsh, N., & Wadhwani, N. (2015). Mobile health-care information for all: A global challenge. *The Lancet Global Health, 3*(7), e356–e357. doi:10.1016/s2214-109x(15)00054-6

Samji, H., Cescon, A., Hogg, R. S., Modur, S. P., Althoff, K. N., Buchacz, K., . . . The North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD) of IeDEA. (2013). Closing the gap: Increases in life expectancy among treated HIV-positive individuals in the United States and Canada. *PLoS One, 8*(12), e81355. doi:10.1371/journal.pone.0081355

Sharma, S., Zapatero-Rodriguez, J., Estrela, P., & O'Kennedy, R. (2015). Point-of-care diagnostics in low resource settings: Present status and future role of microfluidics. *Biosensors (Basel), 5*(3), 577–601. doi:10.3390/bios5030577

Shuchman, M. (2014). The promise and pitfalls of global mHealth. *Canadian Medical Association Journal, 186*(15), 1134–1135. doi:10.1503/cmaj.109-4861

Sinha, S. R., & Barry, M. (2011). Health technologies and innovation in the global health arena. *New England Journal of Medicine, 365*(9), 779–782. doi:10.1056/NEJMp1108040

Skandarajah, A., Reber, C. D., Switz, N. A., & Fletcher, D. A. (2014). Quantitative imaging with a mobile phone microscope. *PLoS One, 9*(5), e96906. doi:10.1371/journal.pone.0096906

Steinhubl, S. R., Muse, E. D., & Topol, E. J. (2015). The emerging field of mobile health. *Science Translational Medicine, 7*(283), 283rv283. doi:10.1126/scitranslmed.aaa3487

Torpey, K., Kasonde, P., Kabaso, M., Weaver, M. A., Bryan, G., Mukonka, V., . . . Colebunders, R. (2010). Reducing pediatric HIV infection: Estimating mother-to-child transmission rates in a program setting in Zambia. *Journal of Acquired Immune Deficiency Syndrome, 54*(4), 415–422. doi:10.1097/QAI.0b013e3181e36616

Ulrich, K. T., & Eppinger, S. D. (2012). *Product design and development*. New York, NY: McGraw-Hill/Irwin.

UNAIDS. (2016). Global AIDS update 2016. Retrieved from http://www.unaids.org/sites/default/files/media_asset/global-AIDS-update-2016_en.pdf

UNAIDS. (2017). 90–90–90: An ambitious treatment target to help end the AIDS epidemic. Retrieved from http://www.unaids.org/sites/default/files/media_asset/90-90-90_en.pdf

Wang, X., Liu, Z., & Zhang, T. (2017). Flexible sensing electronics for wearable/attachable health monitoring. *Small, 13*(25). doi:10.1002/smll.201602790

Wilson, M. L. (2012). Malaria rapid diagnostic tests. *Clinical Infectious Diseases, 54*(11), 1637–1641. doi:10.1093/cid/cis228

World Health Organization (WHO). (2012). Stability of vaccines. Retrieved from http://www.who.int/biologicals/vaccines/stability_of_vaccines_ref_mats/en/

World Health Organization (WHO). (2016a). *Standards for improving quality of maternal and newborn care in health facilities*. Geneva, Switzerland: World Health Organization.

World Health Organization (WHO). (2016b). *World malaria report 2016.* Geneva, Switzerland: Author.

Yamey, G. (2011). Scaling up global health interventions: A proposed framework for success. *PLoS Medicine, 8*(6), e1001049. doi:10.1371/journal.pmed.1001049

Yamey, G., & Morel, C. (2016). Investing in health innovation: A cornerstone to achieving global health convergence. *PLoS Biology, 14*(3), e1002389. doi:10.1371/journal.pbio.1002389

Yock, P. G. (2015). *Biodesign: The process of innovating medical technologies*. Cambridge, UK: Cambridge University Press.

Yock, P. G., Zenios, S., Makower, J., Brinton, T. J., Kumar, U. N., Watkins, F. J., . . . Krummel, T. M. (2015). *Biodesign*. Cambridge, UK: Cambridge University Press.

CHAPTER 18

Evaluations of Large-Scale Health Programs

Melinda K. Munos, Benjamin Johns, and Cesar G. Victora

Why We Need Large-Scale Impact Evaluations

In spite of large investments aimed at improving health outcomes in low- and middle-income countries (LMICs), there is a growing realization that few such programs and initiatives have been properly evaluated ("Evaluation," 2011; Evaluation Gap Working Group, 2006; Oxman et al., 2010). Global targets like the Millennium Development Goals (MDGs; http://www.un.org/millenniumgoals/) and Sustainable Development Goals (SDGs; http://www.un.org/sustainabledevelopment/) have also highlighted the need for more accountability in global health. In response, accountability initiatives for global health proliferated during the MDG era, including the Commission on Information and Accountability for Women's and Children's Health (World Health Organization [WHO], 2014), the Global Strategy (Every Woman Every Child, 2015a), the Unified Accountability Framework (Every Woman Every Child, 2015b), and Countdown to 2015 (http://countdown2030.org/2015/2015-final-report). Accountability initiatives were renewed in the SDG era, now including the Independent Accountability Panel (https://www.everywomaneverychild.org/2016/07/28/independent-accountability-panel/) and the Countdown to 2030 (https://www.countdown2030.org).

Notably, the emphasis has been primarily on tracking progress toward global and country targets, rather than understanding whether and why programs are improving health outcomes in LMICs. The problem of large programs not being evaluated (or only superficially evaluated) remains, making it difficult to understand what does and does not work to reduce mortality and improve health. The Evaluation Gap Working Group (2006) has summarized the issue well:

> Each year billions of dollars are spent on thousands of programs to improve health, education, and other social sector outcomes in the developing world. But very few programs benefit from studies that could determine whether or not they actually made a difference. This absence of evidence is an urgent problem: It not only wastes money but denies poor people crucial support to improve their lives.

This chapter covers a specific area of evaluation science: The rationale for and design and implementation of summative evaluations of programs being scaled up—that is, being delivered to large populations

(Mangham & Hanson, 2010). The chapter does not cover formative evaluations, which are aimed at improving program design prior to large-scale implementation, nor does it address program evaluations using qualitative methods.

The focus here is on evaluating the implementation of programs at scale that are aimed at delivering one or more biological and/or behavioral interventions simultaneously. Most, if not all, interventions packaged in these programs have been submitted to randomized trials that established their efficacy—for example, vaccines, antibiotic or antiviral drugs, micronutrients, and insecticide-treated nets. Thus, little doubt exists that these interventions would reduce mortality and improve health status, if they were delivered with appropriate quality to those who need them.

International and bilateral organizations often pay lip service to the need for evaluations of their large-scale initiatives. Some even pledge a given proportion of their program budgets for evaluation. Despite these lofty statements, few comprehensive evaluations of large health programs have been carried out, perhaps because independent evaluations can be threatening, given their potential for revealing shortcomings or lack of impact of a program.

Such evaluations often include several players with different—and sometimes conflicting—interests: the funders of a program or initiative (e.g., foundations or bilateral organizations), program implementers (e.g., international organizations such as WHO or the United Nations Children's Fund [UNICEF], their country missions, and national government counterparts), and the external evaluation team. In this chapter, we provide our perspective as independent evaluators, but many of the concepts and methods presented are applicable to internal evaluations as well. Although we focus on the technical aspects of design, implementation, and analyses, we also use our experience in a number of multi-country studies to highlight the political tensions that inevitably underlie evaluation science.

Throughout this chapter, three evaluations are consistently used as examples: the evaluation of Integrated Community Case Management in Burkina Faso, Ethiopia, and Malawi (**EXHIBIT 18-1**); a retrospective evaluation of the Accelerated Child Survival and Development Initiative (**EXHIBIT 18-2**); and an evaluation of the Tanzanian national voucher scheme for insecticide-treated nets (**EXHIBIT 18-3**). Key evaluation terms used throughout the chapter are defined in **EXHIBIT 18-4**.

EXHIBIT 18-1 The Evaluation of Integrated Community Case Management in Burkina Faso, Ethiopia, and Malawi

Integrated Community Case Management (iCCM) of childhood illness is a strategy that uses trained lay workers (often called community health workers [CHWs]) to manage uncomplicated cases of childhood illness and to refer more serious cases to health facilities. The strategy aims to treat sick children in their communities so as to reduce delays in care seeking and to reach sick children who might not otherwise be taken to a health facility (WHO & UNICEF, 2012). A scale-up of iCCM for childhood illness was implemented by the Ministries of Health (MoH) in Burkina Faso and Malawi with technical support from UNICEF and WHO, respectively, from 2009 to 2013. Simultaneously, the Government of Ethiopia expanded its iCCM strategy to include case management of pneumonia, in addition to diarrhea and malaria.

The programs in Burkina Faso and Malawi were funded by the Bill & Melinda Gates Foundation (BMGF) through the Partnership for Maternal, Newborn, and Child Health (PMNCH). In Ethiopia, support was provided by Canada's Department of Foreign Affairs, Trade and Development (DFATD) through UNICEF. The funders also commissioned a prospective independent evaluation of the program in these countries by the Johns Hopkins Bloomberg School of Public Health.

Three different evaluation designs were used to account for differences in program design and implementation in the three settings. In Burkina Faso, the evaluation used a quasi-experimental design, as the MoH had already identified the nine districts that would receive the program; seven comparison districts were selected using group matching. In Ethiopia, the evaluation was commissioned early enough that the evaluation team was able to randomize woredas (small administrative areas) to receive the expanded iCCM program or the existing program. In Malawi, where iCCM was scaled up to every district in the country and supported by different partners, the evaluation made use of a dose–response design, which assessed the extent to which changes in under-5 mortality at the district level were associated with the strength of implementation of the iCCM program in the district.

Detailed descriptions of the evaluations and results are available in several publications (Amouzou et al., 2016; Amouzou et al., 2016; Munos et al., 2016).

Data from Amouzou et al., 2016a; Amouzou et al., 2016b; and Munos et al., 2016.

EXHIBIT 18-2 The Retrospective Evaluation of the Accelerated Child Survival and Development Program

The Accelerated Child Survival and Development program (ACSD) was implemented by UNICEF in 11 West African countries between 2001 and 2005 with support from the Canada International Development Agency. The aim of ACSD was to increase coverage for proven interventions grouped in three broad packages: EPI+ (immunizations, vitamin A supplementation, insecticide-treated nets), IMCI+ (case management of malaria, diarrhea, and pneumonia; breastfeeding promotion), and ANC+ (antenatal and delivery care, postnatal vitamin A supplementation, intermittent presumptive treatment of malaria). The ultimate aim of the program was to reduce under-5 mortality by 20%.

In each participating country, ACSD high-impact districts were selected to scale up these interventions rapidly to full coverage. At the same time, a smaller set of interventions was supported, to a greater or lesser extent, in additional districts in each country—the so-called ACSD expansion areas. The countries that first moved ahead with rapid implementation were Benin, Ghana, Mali, and Senegal. More information on ACSD is available at the following website: https://www.unicef.org/health/23958_survivaldevelopment.html.

UNICEF commissioned the Bloomberg School of Public Health at Johns Hopkins University to conduct an independent retrospective evaluation of the ACSD project. The objective was to provide valid and timely evidence to child health planners and policy makers about the effectiveness of the ACSD project in reducing child mortality and improving child nutritional status, as a basis for strengthening child health programming in the future. From 2006 to 2008, the evaluation team reviewed documentation on ACSD implementation, evaluated intervention coverage using standard international indicators, and measured the program's impact, focusing on under-5 mortality in Benin, Ghana, and Mali. An evaluation in Senegal was also started but could not be completed due to a lack of information on comparison areas. Because the evaluation was retrospective, information on coverage and mortality was obtained through reanalyses of national surveys, comparing the ACSD high-impact districts with the remaining rural districts in the country.

The evaluation results are available in a publication (Bryce, Gilroy, et al., 2010) and at the following website: https://www.jhsph.edu/research/centers-and-institutes/institute-for-international-programs/completed-projects/accelerated-child-survival-and-development/.

EXHIBIT 18-3 Evaluation of the Tanzanian National Voucher Scheme for Insecticide-Treated Nets

Insecticide-treated nets (ITNs) are a cost-effective means of preventing malaria, which remains one of the major killers of children younger than the age of 5 in Africa. The Tanzania national voucher scheme is a targeted subsidy program that relies on the commercial sector to distribute ITNs throughout the country. Every pregnant woman who attends antenatal care in a government organization or nongovernmental organization (NGO) facility is eligible to receive a voucher that can be used as partial payment for an ITN.

The Tanzania national voucher scheme was rolled out gradually, starting in late 2004 in districts near Dar es Salaam, and ending with some of the most remote districts in the country 18 months later. Program managers decided that a phased rollout was necessary on programmatic grounds because of the size of the country. It was not feasible or acceptable to randomize districts by rollout phase. Nevertheless, the gradual scale-up allowed the evaluation team to compare areas with and without the voucher scheme at three points in time using cross-sectional household and facility surveys. Within-district changes were compared in the analysis with the length of time that the program had been operating in each district to determine whether program duration was associated with increases in net coverage.

The objectives of this study were to assess the changes in the level and socioeconomic distribution of ITN coverage over the period 2005–2007, during which the voucher scheme was initiated and expanded to a national scale, and to examine the link between program duration and change in household net ownership as a measure of program impact.

The evaluation was carried out by a team of U.K.- and Tanzania-based investigators (Hanson et al., 2009; Hanson et al., 2008; Mulligan, Yukich, & Hanson, 2008). Particular emphasis was given to describing the intermediate steps in the process leading to high ITN coverage, as well as to the assessment of costs and of impact of the program on reducing socioeconomic inequalities in net ownership.

EXHIBIT 18-4 Key Terms

Formative evaluation: An evaluation that aims to describe, quantitatively and/or qualitatively, *how* a program is being implemented, and its immediate outputs, with the aim of improving program implementation.
Summative evaluation: An evaluation that aims to quantify the impact or effect of a program on a population.
Input: An item necessary for the implementation of a program, such as funding, policies, plans, and human resources.
Process: Program activities such as training, systems or management improvements, infrastructure improvements, or demand generation.
Output: Short- or medium-term result of program processes, including changes in utilization, service quality, health worker competence or knowledge, and population knowledge or attitudes.
Outcome: Change in population-level intervention coverage or behavior.
Impact: Change in health status, such as mortality, morbidity, nutritional status, or fertility.
Contextual factor: A factor external to the program being evaluated that either *confounds* (masks) the program's true effect or *modifies* the effect of the program in the setting in which it is implemented.

These definitions apply to the terminology used in this chapter. Different agencies and researchers may vary in the way they define these concepts.

Planning the Evaluation

Who Will Carry Out the Evaluation?

Internal evaluations are carried out by the implementing institutions themselves, sometimes with the help of external consultants for specific tasks. Such evaluations often address levels of inputs and utilization (e.g., whether expected quantities of drugs or mosquito nets were procured and distributed) and issues related to process and outputs (e.g., surveys of the quality of care being provided or the frequency of supervision). Less frequently, internal evaluations include population-based coverage surveys, which may be subcontracted to consultants. Internal evaluations are classified into two main categories: (1) formative—that is, aimed at improving the program in its early implementation stage or (2) summative—that is, attempting to document an effect of the program on coverage levels or on health indicators. They may or may not include a comparison with other areas without the program.

The main difference between internal and external evaluations is the lack of independence between implementers and evaluators in internal evaluations, which creates an obvious conflict of interest because the continuity of funding for the program is affected by evaluation results. Nevertheless, internal evaluations are essential for fine-tuning a program, and are increasingly required or strongly encouraged by funders. For example, USAID's (2016) Child Survival and Health Grants Program has strongly encouraged grantees to conduct internal evaluations.

External, or independent, evaluations are carried out by researchers who are not involved in implementation. As a rule, the evaluation is funded by a third party—either the institution that provided funds for the implementation agency or an outside agency, such as the International Initiative for Impact Evaluation (3ie; http://www.3ieimpact.org/). Although external evaluators must collaborate closely with the implementation team—as will be discussed later in this chapter—they retain a level of independence that is essential for ensuring the credibility of the evaluation findings and can contribute to both formative and summative evaluations. In the ACSD evaluation (UNICEF, 2005), the results of internal evaluations proved to be more optimistic than later results from the independent team's external evaluation. This chapter focuses on external evaluations.

In-country research institutions that are independent of program evaluation can and should play an important role in these evaluations, bringing context-relevant experience and expertise to the team, serving as a first point of contact with program implementers, and providing continuity in the dissemination and eventual uptake of evaluation findings. The selection of the in-country research partners should be based on objective criteria that include research capacity relevant to the evaluation design and credibility with the ministry of health. Partnerships between in-country and external research institutions can result in expanded capacity and objectivity that will contribute to the success of the evaluation.

What Are the Evaluation Objectives?

The first task of the evaluation team is to review the available documentation on program objectives and goals, and to turn these items into evaluation

objectives. This transformation is best achieved by involving implementers and funders at the beginning of the evaluation design process. The evaluation objectives should be informed by the program *impact model* (described in more detail later in this chapter), which details the pathways from program inputs and processes to hypothesized impacts. A good impact model requires implementers and evaluators to make explicit the assumptions and hypotheses about how the program will work.

An early decision is to agree upon what is being implemented and what is being evaluated. For example, in the Burkina Faso iCCM evaluation (Exhibit 18-1), the program included both iCCM and some facility-based maternal and newborn activities, including training health workers on emergency obstetric and newborn care (EmONC) and acquiring materials such as obstetric delivery tables. However, discussions with implementers indicated that the program's objective was to reduce under-5 mortality by 25%, that the maternal and newborn health (MNH) activities were being implemented in other regions of the country with funding from other sources, and that there were no foreseen synergies between the iCCM and MNH components. In addition, modeling indicated that most of the impact of the program would be attributable to the iCCM interventions. For all these reasons, the evaluation team decided to focus the evaluation on the implementation and impact of the iCCM interventions, while also documenting MNH activities and assessing changes in coverage of MNH interventions.

The ACSD strategy (Exhibit 18-2) included three packages—EPI+, IMCI+, and ANC+—that were implemented with variable intensities in different countries. Although it may be tempting to try to design an evaluation of such a program in ways that disentangle the effects of different components, in practice this parsing is often impossible, because one cannot predict which components will end up being more strongly implemented, and in which part of the country. Also, many strategies are designed to take advantage of the potential synergies among its components—for example, ACSD—so that trying to break out their effects separately contradicts the entire concept underlying the program.

A related issue is the challenge of trying to attribute program impact to a single donor when a program is funded by several sources, each of which supports one or more components. Thus, the definition of what is being evaluated—and what is not being evaluated—must be agreed upon with evaluation funders and program implementers early in the process of evaluation design.

Large-scale programs and initiatives often establish quantitative goals in terms of what they expect to achieve. These aims are frequently expressed in terms of coverage (e.g., 70% of all pregnant women receiving a long-lasting mosquito net) or impact (e.g., a 25% reduction in under-5 mortality). Program goals are often expressed in ways that allow different interpretations. For example, the target coverage of a program based in health facilities could be calculated for the whole population or only for those persons living within a given distance (catchment area) of the facility.

A more complex example is how to interpret a "25% reduction in mortality." In this case, four interpretations are possible:

- Endline mortality being 25% lower in program areas than baseline levels in the same areas
- Endline mortality being 25% lower in program areas than in comparison areas (regardless of baseline differences)
- The reduction in mortality from baseline to endline in program areas being 25% greater than the corresponding reduction in comparison areas
- The rate of decline over time in the program areas being 25% faster than the rate of decline in comparison areas (e.g., 5% annual decline in areas covered by the program and 4% annual decline in comparison areas)

These four options have different implications for study design and sample size calculations, and it is essential to reach agreement on a common interpretation at the beginning of the evaluation. For example, in the Burkina Faso iCCM evaluation, the Ministry of Health's objective was to reduce under-5 mortality by 25% from baseline to endline, whereas the donor wanted to know whether under-5 mortality was reduced by 25% from baseline to endline relative to a comparison area. The evaluation team had to assess whether each of these targets was achieved and power the evaluation accordingly.

Once the overall evaluation objectives are defined, they must be broken down into specific evaluation questions that the study will try to answer. These questions are detailed later in this chapter. The ultimate objective of an evaluation is to influence decisions. Evaluation objectives and design depend on who the decision maker is and which types of decisions will be taken as a consequence of the findings. Program funders—for example, high-income-country organizations providing aid to health programs in LMICs—are typically interested in providing evidence to their governments and taxpayers that their funds led to a measurable impact on health outcomes, such as mortality, at a reasonable cost.

The decisions to be made in such cases include whether to continue funding the program and at what level, and whether the program strategy needs to be reformulated. Funders often work on a short timeline: They need results soon, in accordance with funding and electoral cycles. They also tend to be more interested in impact measurement than in intermediate results such as inputs, processes, outputs, or outcomes, although there is increasing interest is using intermediate data (e.g., quality of service provision) to improve programs.

In contrast, local implementers—for example, senior officials at the ministry of health—may seek reassurance that the program is moving in the right direction, that the quality of services is adequate, and that high population coverage is being or will be achieved. Their decisions primarily focus on improving the program through specific actions. Evaluation results may also have an important role in advocacy and provide political gains at country level.

Later in this chapter, we describe the importance of building a conceptual model that takes into account the different needs of different partners and the decisions they must make as a result of the evaluation findings. Obtaining information on health impact measures such as mortality is important, but just as important is to understand why the program had—or failed to have—a measurable impact.

When Should Planning for the Evaluation Occur?

Many advantages accrue from planning the evaluation at the time the program is being designed. Developing an evaluation plan and conducting a high-quality baseline survey can take 6 to 12 months, depending on the complexity of the program, the number of implementation and evaluation partners, and the size and scope of the survey. Early-onset, prospective evaluations allow collection of baseline data before implementation starts. They also allow thorough, continuing documentation of program inputs and the contextual variables that may affect the program's impact. Under some circumstances, early planning may enable the evaluation team to influence how the program is rolled out, thereby improving the validity of future comparisons. A disadvantage of prospective evaluations that include a comparison area is that program implementation may change over time for reasons that are outside the control of the evaluation team—for example, similar activities may be implemented in the comparison districts, or some of the program districts may have insufficient implementation.

In reality, evaluation is often an afterthought. Studies are frequently launched when implementation is already underway—and in some cases after the program cycle is completed. Such retrospective evaluations have important limitations. Documentation requires the reconstruction of project assumptions and activities by requesting the assistance of project implementers to produce records of activities and inputs. In such cases, the resulting information is often incomplete, inconsistent, and difficult to verify. Baseline (and sometimes endline) data are often unavailable. Even when they exist, they may be of poor quality, be based on sample sizes that are too small to address the evaluation questions, or lack information on all needed indicators. The importance of carrying out evaluations prospectively cannot be overemphasized.

How Long Will the Evaluation Take?

The answer to this question depends on whether the evaluation is retrospective, prospective, or a mixture of both techniques (ambispective) (Kleinbaum, Kupper, & Morgenstern, 1982). Fully prospective evaluations should include several sequential steps:

1. Collect baseline information.
2. Implement a system to document program activities and contextual factors.
3. Wait until the large-scale program is fully implemented and reaches high population coverage.
4. Conduct assessments of implementation strength (IS) and/or the quality of care (QoC).
5. Allow time for program adjustments based on IS/QoC assessment and documentation results, if time allows and implementers are receptive.
6. Allow time for a biological effect to take place in participating individuals.
7. Wait until such an effect can be measured in an endline survey.
8. Clean the data and conduct the analysis.

Each step can easily take one year, often longer. In our experience, steps 3 and 5 usually take longer than initially anticipated, due to delays in staff deployment, training, and commodity procurement. Step 6 can be fairly short for interventions with a rapid effect on outcomes. For example, case management of disease episodes or interventions that prevent disease transmission—such as insecticide-treated nets or indoor residual spraying for the prevention of malaria—can reduce mortality rapidly. In contrast, the time needed for an intervention to achieve its

biological effect can be much longer for interventions that require changes in behavior, such as changing breastfeeding practices to improve nutrition.

Step 7 can also slow down the production of final results. For example, assessment of child mortality is retrospective, and often the calendar midpoint of the mortality estimate is one to two years prior to the date of the endline survey.

Cleaning the data and conducting a full analysis (step 8) is often overlooked when planning an evaluation, but requires considerable time and effort to do well. The introduction of computer-assisted personal interviewing (CAPI) technology, in which paper questionnaires are replaced with data input through tablets and other devices, has greatly reduced the time needed to clean data, but evaluators need to allocate sufficient time both to program the CAPI questionnaires prior to data collection and to conduct data quality checks during and after data collection.

Donors, policy makers, and evaluation funders frequently organize their work in periodic cycles of 5 years or less. Because of delays in awarding contracts and implementing programs, this can lead to evaluation timelines of 12 to 36 months. Prospective impact evaluations of complex programs under real-life conditions, however, often require longer time frames. Because few policy makers are willing to wait so long for study results, it is important that data on intermediate outcomes, such as quality of care or achieved coverage, become available in the first year or two of the evaluation. This issue is discussed in further detail in the section on implementation.

Where Will the Evaluation Be Carried Out?

Many large-scale programs are implemented simultaneously in more than one country. For these programs, evaluations may include all countries covered by the program or only a subset of countries. The type of evaluation—for example, formative versus summative; adequacy versus plausibility or probability—may also vary by country. This decision is often made by the donor, with or without input from the evaluation team or implementers; less frequently the decision is made by the evaluation team in consultation with the donor or implementing agencies.

Decisions about which countries to select should be based on factors such as the likely strength of the implementation (or, for retrospective or ambispective evaluations, the strength of implementation to date), the availability of local research partners, and the need for the group of countries to be reasonably representative of geographic areas or epidemiologic patterns (e.g., the presence or absence of high levels of human immunodeficiency virus/acquired immunodeficiency syndrome [HIV/AIDS] or malaria). For example, when evaluating a program being rolled out in many LMICs, selection criteria should include both characteristics that are desirable in all participating countries (e.g., political stability, availability of local research partners, likelihood of strong implementation of the program) and characteristics that are desirable in the set of countries as a whole (e.g., reflecting the diversity of all implementing countries in terms of geography, health system strength, and epidemiologic profiles).

Whether country selection is done by evaluators, donors, or implementing agencies, it is important to keep track of the rationale for selecting some countries and not others, because these criteria will affect the external validity—or generalizability—of the evaluation findings (Bryce, Victora, Habicht, Vaughan, & Black, 2004). For example, if the evaluation is conducted only in countries that have performed well in implementing the program, then it is more likely that an impact of the program may be detected, but extrapolating the results to all implementing countries may not be justified. Similarly, evaluation teams may be more likely to select stable, non-conflict-affected settings for evaluation, but their results then may not be generalizable to implementations of the program in conflict settings. In cases where the decision about which countries to select is made by donors or implementers, it is particularly important for the evaluation team to examine and document the rationale behind country selection and to be alert to the possibility of conscious or unconscious bias—that is, the possibility that donors and implementers might tend to select better-performing countries (or countries where the program is likely to perform better) for evaluation.

The prospective multi-country evaluation (MCE) of Integrated Management of Childhood Illness (IMCI) is an example of an evaluation where country selection was conducted by the independent evaluation team based in part on inputs from implementers. IMCI is a strategy for reducing mortality among children younger than 5 years of age that includes case management guidelines for the integrated management of sick children in first-level health facilities, improvement of health systems support, and promotion of key family and community practices related to child health. IMCI was first introduced at the country level in 1996 by Tanzania and Uganda; by 2006, it had been adopted at the national level by ministries of health in more than 100 countries. The MCE aimed to provide information about the barriers to IMCI implementation, the effects of the strategy on health services and communities, the cost of the program, and the number of lives saved. Based on

early reports of strong implementation at the national level, 12 countries participated in the initial round of MCE assessments and data collection about IMCI implementation. More in-depth evaluations were conducted in 5 countries: Bangladesh, Brazil, Peru, Tanzania, and Uganda. Although the 5 countries selected for in-depth evaluation were judged to best meet the selection criteria among the 12 countries originally assessed, there were wide variations in the strength of implementation and, therefore, reasonable external validity to support generalization of the findings (Bryce, Victora, Habicht, Black, & Scherpbier, 2005).

In contrast, the ACSD evaluation in West Africa was retrospective, and UNICEF preselected 4 countries among the 11 participating in the program based on the fact that they were judged to have the strongest implementation of all program components. This decision likely introduced a positive bias into the design and limited the extent to which the results could be generalized to the remaining countries (Bryce, Gilroy, et al., 2010).

Once the country or countries for the evaluation are defined, there is a need to select which districts will be included in the study or, less commonly, to design an evaluation that covers all districts in the country; these options are discussed later in this chapter. As discussed in the context of country selection, the criteria by which districts are chosen may strongly influence the external validity of the evaluation results. Criteria for selecting countries and districts within countries should be reported in the main results publications arising from the evaluation.

Developing an Impact Model

The development of an impact model—also known as an operational or logic model, a model of change, or a conceptual framework—for the program's assumed effects is an essential step in designing an evaluation of a large-scale program. This model lays out how the program inputs are expected to lead to an effect on health. It is typically developed by the program evaluation team through discussions with program implementers prior to the start of the evaluation. The impact model may go through many iterations over the course of the evaluation to account for changes in the program itself, as well as changes in implementers' and evaluators' understanding of the program and its intermediate effects. Recording these changes in the impact model is an important aspect of documenting program implementation.

A commonly used framework includes the following sequence: inputs → process → outputs → outcomes → impact (Bryce et al., 2011). This generic framework must then be adapted to the particular program. **EXHIBIT 18-5** describes the IMCI strategy and its impact model developed by the evaluation team as part of the multi-country evaluation of IMCI (discussed in the previous section) (Bryce et al., 2005; Bryce et al., 2004). Tools are now available to facilitate the mechanics of developing, revising, and documenting versions of the impact model (for example, the Evaluation Planning Tool [hframe.io]).

In the Tanzania ITN evaluation (Exhibit 18-3), a detailed model was designed that included several necessary steps for the voucher scheme processes to translate into effective coverage: Vouchers needed to be in stock; women needed to attend antenatal services; women had to receive and redeem their voucher; and women needed to sleep under the voucher net. Evidence from three sequential surveys was used to assess intermediate steps in the program pathway so as to inform judgments about its effects (Hanson et al., 2009).

Impact models are essential for several reasons. For instance, they help clarify the expectations of program planners and implementers, by laying out these anticipated outcomes clearly. They help ensure that program implementers and evaluators have the same understanding of program activities and the pathways by which these activities are hypothesized to lead to specific impacts. They help estimate sample sizes, by defining what needs to be measured and what the magnitude of effects is likely to be. Most importantly, models contribute to the development of the evaluation proposal. For example, each box in Figure 18-1 generated relevant evaluation questions, indicators, and data collection strategies. The need to measure improvements in health worker performance, for instance, led to the development of health facility surveys based on the IMCI clinical guidelines, which were then implemented in several countries participating in the evaluation. Also, impact models help guide the analyses and attribution of the results. For example, if the evaluation finds a reduction in mortality in the absence of changes in the intermediate outcomes laid out in the model, it is unlikely that the observed impact can be attributed to the program.

Two other reasons for building impact models upfront can be cited. First, in a dynamic evaluation setting with feedback to implementers (described later in this chapter), models can help track changes in assumptions as these evolve in response to early evaluation findings. Second, models force implementers and evaluators to stay honest about what was expected. Unfortunately, program managers may be tempted to rewrite history by changing expectations as programs are rolled out and initial expectations are shown to be unrealistic.

EXHIBIT 18-5 The IMCI Impact Model

FIGURE 18-1 shows a simplified model of the pathways through which the developers of the strategy believed that IMCI would improve child health and nutrition. This model was originally developed by the evaluation team and refined through consultations with IMCI developers. For each box in the model, the evaluation team developed appropriate indicators and data collection procedures. The evaluation questions addressed the assumptions underlying this model, shown by the arrows in the figure. The impact model formed the backbone of the evaluation, serving as a common frame of reference for the evaluation team and program implementers, supporting estimates of needed sample sizes for various evaluation components, and providing a map for the data analysis that showed clearly the expected associations to be tested.

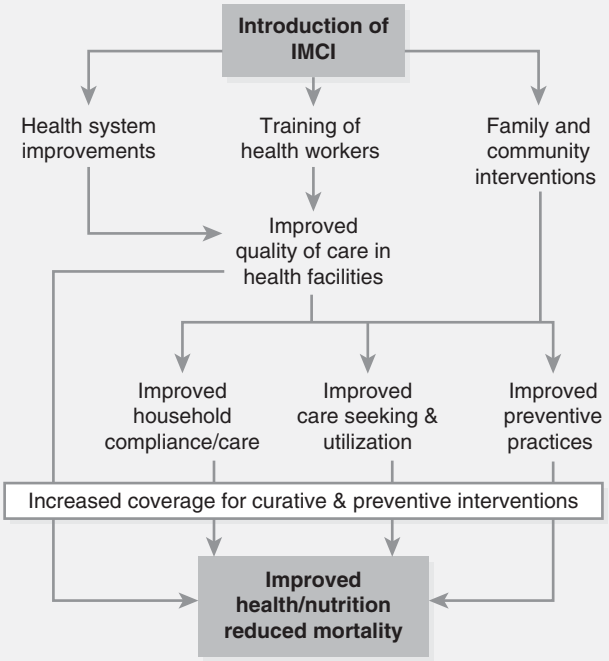


FIGURE 18-1 Example of an impact model: The multi-country evaluation of IMCI.

Reproduced from Bryce, J., Cesar, V. G., Habicht, J-P., Black, R. E., & Scherpbier R. W., on behalf of the MCE-IMCI Technical Advisors. (2005). Programmatic pathways to child survival: Results of a multi-country evaluation of Integrated Management of Childhood. Illness (Figure 2, pp. 4). *Health Policy and Planning, 20*(suppl 1).

The steps involved in developing and checking an impact model are summarized in **TABLE 18-1**. In essence, an impact model describes how a given program is expected to affect health. A later section of this chapter discusses how such models fit within the broader conceptual framework of factors influencing health status, which go beyond any given program or specific attempt to capture other determinants of health.

Stepwise Approach to Impact Evaluations

FIGURE 18-2 summarizes the basic stepwise approach to conducting an impact evaluation. Use of a stepwise approach allows evaluators to provide early feedback to program implementers and to halt the evaluation at midcourse if the implementation is not sufficiently strong to produce a measurable impact, thereby avoiding the costly surveys needed for assessing coverage and impact. The six steps are summarized next.

1. *Assess the technical soundness of implementation plans in light of local epidemiologic and health services characteristics.* It is not unusual for programs to be imported into a country without a proper assessment of whether they address the leading causes of morbidity and mortality, or without a realistic evaluation of the health system characteristics required for making the program work. When considering whether to evaluate a health facility–based program in Niger in the late 1990s, for example, the proposed study was dismissed when a desk review showed an average of 0.5 annual visit per under-5 child in the country, suggesting that the health systems were not strong enough to deliver the program in a short time frame. In recent years, a variety

TABLE 18-1 Steps in the Development of an Impact Model

Step	Details
Learn about the program	Read documents. Interview planners and implementers. Carry out field visits to future implementation areas. Use special techniques as needed: ■ Card-sorting exercise
Develop drafts of the model	Focus on intentions and assumptions. Document responses from implementers. Record iterations and changes as model develops.
Quantify and check assumptions	Review existing evidence and literature. Identify early results from the evaluation: ■ Documentation: What was actually done? ■ Outcomes: Are assumptions confirmed?
Use and evaluate the model	Develop an evaluation design, testing each assumption if possible. Plan for analysis, including contextual factors. Analyze the results. Interpret the results with participation by the implementers.

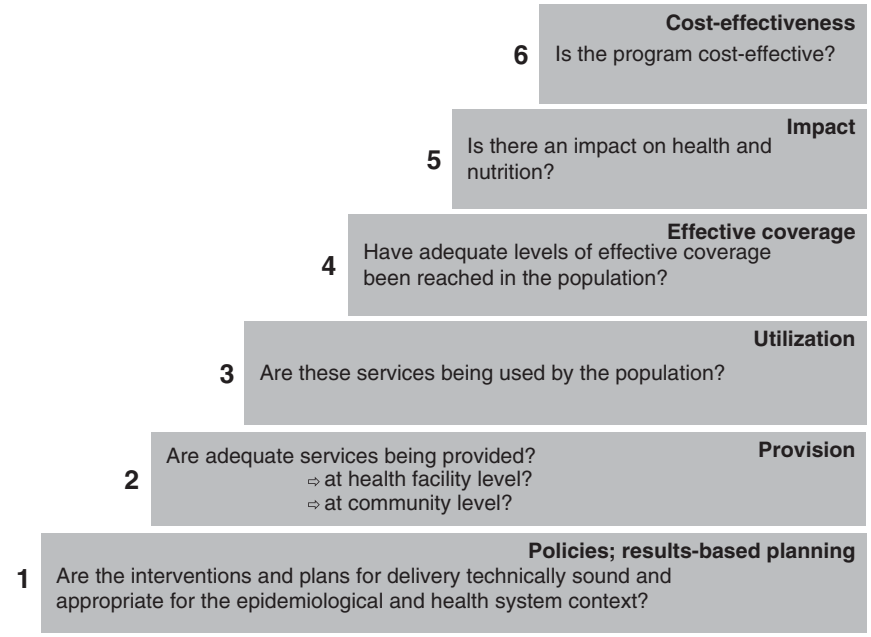


FIGURE 18-2 A stepwise approach to large-scale impact evaluations.

of tools for estimating the impact of interventions in different epidemiologic contexts have been developed. One of those instruments, the Lives Saved Tool (LiST), which is described in detail later in this chapter, can contribute to this kind of initial assessment (Johns Hopkins Bloomberg School of Public Health, n.d.a).

2. *Investigate whether the quantity and quality of the program being provided are compatible with a potential impact*. The evaluation team makes this judgment using the documentation of program activities, combined—if applicable—with surveys assessing the quality of care being provided. For example, in the iCCM evaluation in Burkina Faso, a

survey of CHWs found widespread stockouts of essential drugs and poor quality of care among trained CHWs. Based on these results (and low CHW utilization), it seemed that the program might not achieve an impact, and the endline coverage survey confirmed that the program did not lead to improvements in intervention coverage (Munos et al., 2016). A mortality survey was not deemed necessary to confirm the lack of impact.

3. *Assess whether data on outputs or utilization suggest that an impact is likely*. Even if the program inputs seem sufficient in terms of quantity and quality, uptake by the population may be limited. For example, if the number of insecticide-treated nets effectively distributed in a country is much smaller than the number necessary to achieve adequate coverage in the target population, continuing the study to measure health impact would be unwarranted. This step is usually assessed through documentation of program outputs and/or provider surveys; it does not require household surveys. For example, in the iCCM evaluation in Ethiopia, a survey of health extension workers (HEWs) showed low levels of utilization of such workers by sick children, despite finding that the HEWs provided high quality of care (Miller et al., 2014). As a result, it seemed unlikely that the program would have a population-level impact on mortality, and this suspicion was confirmed by the mortality survey conducted at endline (Amouzou et al., 2016).
4. *Check whether adequate intervention coverage has been reached*. Even if outputs appear to be adequate, it is necessary to check whether they have effectively reached the population and are being used by the target groups. Making this determination usually requires a population survey, which will cost much more than assessing the preceding steps on the basis of documentation. In the ITN voucher evaluation, for example, the nets that were distributed may have been sold to persons who did not reside in the intervention area, or they may have been used by the heads of households instead of by the pregnant women and young children who constituted the primary targets of the program. Assessment of coverage applies not only to interventions based on commodities such as ITNs or vaccines, but also to health-related behaviors—for example, condom use, breastfeeding, or care seeking.
5. *Assess the impact on health*. Measuring health impact often requires costly surveys, and the stepwise approach may reveal shortcomings in the preceding steps that suggest an impact is unlikely. For cost, timing, and logistical reasons, impact is often measured in the same household survey as coverage. However, the inclusion of impact measures—particularly mortality—can greatly increase the sample size and cost of the survey. If evaluators have identified shortcomings related to the program plan, implementation, or utilization and do not expect to see an impact on health status, they may decide to measure only coverage in the endline survey, as was the case in the Burkina Faso iCCM evaluation. Impact attribution also requires ruling out alternative explanations for the observed findings. This may prove to be difficult, or even impossible, no matter how much funders and implementers would like to produce an unequivocal statement. These issues are discussed in greater detail later in this chapter.
6. *Measure cost-effectiveness*. If there is evidence of an impact, the next question relates to cost-effectiveness. Measurement of program costs is discussed in subsequent sections of this chapter.

Types of Inference and Choice of Design

There is no single "best" design for evaluations of large-scale programs, as different types of decisions require different degrees of certainty to support their decisions. Whereas some decisions require randomized trials, other decisions may be adequately taken with observational studies. Also, the manner in which programs are rolled out often limits the scope of possible approaches to evaluation.

In real-world evaluations, the program of interest often accounts for a small part of the variability in the outcomes. **FIGURE 18-3** presents a simplified framework showing that health outcomes may be also influenced by socioeconomic factors, by changes in existing health services in the public and private sectors (which are outside the scope of the program of interest), and by other programs in the health and other sectors present in the same geographic areas.

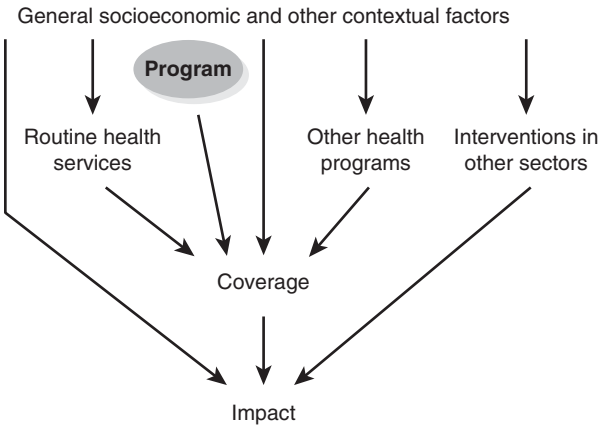


FIGURE 18-3 Simplified conceptual framework of factors affecting health, from the standpoint of evaluation design.

We will refer to this framework when describing different evaluation designs.

Habicht et al. (1999) proposed that evaluations can be classified as one of three types in terms of strength of inference, according to the ways in which the factors described in Figure 18-3 are handled. These types of inference are discussed next in relation to possible evaluation designs.

Adequacy Evaluations

Adequacy evaluations assess whether initial targets were met in the program areas or, at the very least, whether trends moved in the expected direction in terms of coverage or impact measures. For example, for a health manager it may be sufficient to establish that 80% coverage with an intervention (e.g., a vaccine) has been reached in the district; this evidence may be enough to support the continuation of the strategy for delivering this intervention.

Adequacy evaluations are often based on before-and-after studies in the program areas only, aimed at measuring progress in coverage or impact indicators. If the intervention included in the program is new to the geographic area, however, adequacy evaluations may be limited to an endline survey. Such a design is particularly well suited to studies of coverage, in which a direct link exists between the program and the frequency of the intervention's use in the population.

Adequacy evaluations assess whether targets were met or whether trends for outcome or impact measures moved in the expected direction. In contrast, use of adequacy evaluations to attribute changes in coverage or health status to a specific program is more complex. Adequacy evaluations do not include a concurrent comparison group and do not attempt to rule out other explanations for changes in coverage or health status; as a result, in most cases they cannot be used to attribute observed changes in outcomes or impact to a specific program. Even so, adequacy evaluations may produce valid inferences about the effects of a program if the following criteria are met: (1) the causal pathway from program to impact is relatively short and simple; (2) the expected effect is large; and (3) confounding is unlikely—that is, other factors in Figure 18-3 are either unchanged or are unrelated to the impact measure. For example, introduction of the *Haemophilus influenzae* type b (Hib) vaccine in Uruguay was associated with virtual disappearance of Hib disease (Pan American Health Organization [PAHO], 1996). In this example, (1) administration of Hib vaccine to the target population would be expected to lead directly to a reduction in incidence; (2) the vaccine is effective and, therefore, the expected effect was large; and (3) there were no other known factors that could lead to a large drop in Hib incidence. For most health programs and interventions, however, these conditions do not hold, and more complex evaluation designs are needed.

Nevertheless, adequacy evaluations are often the only alternative for programs that are scaled up rapidly and reach whole populations. Also, adequacy evaluations are often the only approach available to estimate the impact on health of policy changes effected through legislation—for example, salt iodization or changes in antimalarial drug regimens. These evaluation designs are also typically less expensive and complex to implement, so they may be used in situations where the program or donor does not require inferences about the effect of the program on coverage or health status.

Even if more complex analyses are being planned, adequacy of program coverage should always be reported. For example, in the Tanzania voucher evaluation (Exhibit 18-3), steady increases in national coverage of any type of net and ITNs were observed over the study period. ITN use among infants increased from 16% in the baseline survey to 34% in the endline survey (Hanson et al., 2009), showing important progress in this area but also revealing considerable room for improvement. Subsequent analyses determined that increases were significantly stronger in early-implementation districts.

In summary, adequacy evaluations are useful for assessing coverage outcomes, but more often than not they are insufficient for attributing these outcomes to the program or for establishing the health impact of a specific intervention.

Plausibility Evaluations

Plausibility evaluation designs aim to document the health impact of an intervention and to rule out alternative explanations by including a comparison

group and addressing confounding variables. Such evaluations are particularly useful when randomized allocation of the program is not possible for ethical, practical, or political reasons—situations that in real-life evaluations constitute the norm, rather than the exception. Plausibility evaluations are also useful to demonstrate the large-scale effectiveness of programs or interventions whose efficacy has already been demonstrated in smaller-scale studies. These evaluations may avoid some of the artificiality of randomized controlled trials (RCTs), which often address efficacy, by studying health impacts under real-life, less-than-perfect implementation conditions. For this reason, the external validity of plausibility evaluations is likely to be greater than that for tightly controlled trials (Victora, Habicht, & Bryce, 2004).

Several types of design may be used in a plausibility evaluation, some of which are described next.

Before-and-After Study in Program and Comparison Areas

A commonly used type of study design is the before-and-after program–comparison design. In this design, one or more areas—say, districts—that received the program being evaluated are compared with districts that did not receive the program (**FIGURE 18-4**). Because health impact indicators may be affected by factors other than the program, it is essential to document changes on the nonprogram factors depicted in Figure 18-4, in both the program and the comparison districts. If these factors differ at baseline from one set of districts to the other, or if they evolve over time in different ways, they may bias the results of the evaluation.

A detailed discussion of design issues with this type of evaluation is outside the scope of this chapter, but some common recommendations are as follows:

- Select as many districts as possible with and without the programs. Detailed sample size calculations are needed, but ideally aim for 10 or more districts in each group.
- In the data analyses, treat the districts as the units, rather than the individuals living in the districts. Thus, the sample size will be, for example, 10 districts in the program group, rather than the number of individuals surveyed.
- When the number of districts with the program is small, select comparison districts from the larger pool of districts without the program by matching them to the program districts according to key characteristics—the most important of which are baseline levels of the main impact indicator. If data on the impact or outcome indicators of interest are not available at the district level, districts can be matched on health systems, poverty, or education indicators, which are more commonly available at district level (see the later discussion of propensity score matching).
- In some cases, when the number of districts with the program is small, the districts receiving the intervention may be selected based on observable criteria, such as the need for the intervention. For example, districts with the highest levels of child mortality may be selected for the intervention, such that no districts with similar levels of child mortality remain in the pool of possible comparison districts. However, if a sufficient number of districts have child mortality levels that are close to those districts where the intervention is to be implemented, these districts should be selected as the comparison districts. This type of selection, which is known as nonrandom quantitative assignment of treatment, can be assessed using regression discontinuity analysis (West et al., 2008).

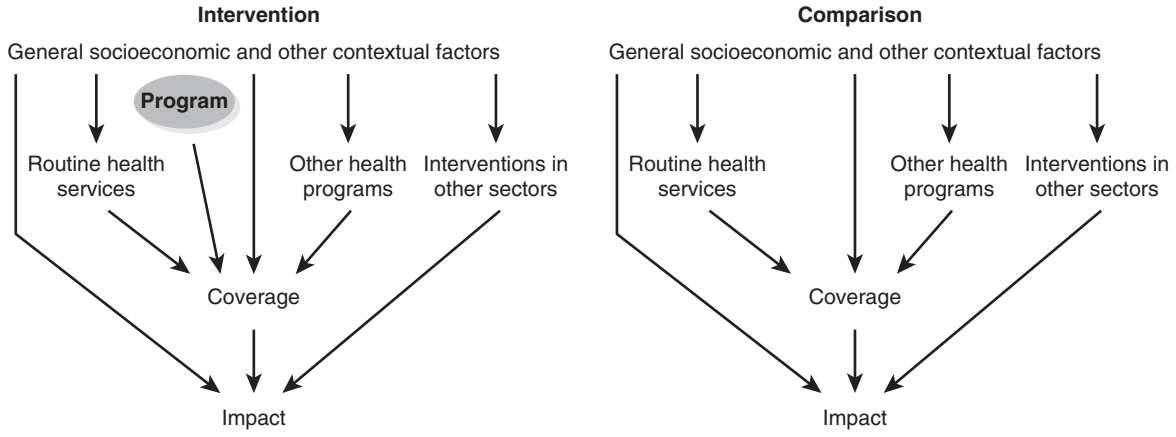


FIGURE 18-4 Schematic program–comparison design.

- Collect detailed information on all nonprogram factors listed in Figures 18-3 and 18-4.
- Take nonprogram factors into account in the data analyses by treating them as potential confounding factors and considering them in the interpretation of results (Cousens et al., 2009).

The before-and-after program–comparison design is well suited to evaluate the introduction of large-scale programs delivering interventions whose efficacy has already been demonstrated under experimental conditions, but that have yet to be scaled up. In this case, the comparison areas are completely devoid of the program's interventions—hence evaluators often refer to them as "virgin" or "untouched" comparison areas.

Many other large-scale evaluations address new strategies for accelerating the delivery of biological or behavioral interventions that are already being implemented, to a greater or lesser extent, in other districts in the same country. An example is the ACSD initiative in West Africa (Exhibit 18-2). This program represented a renewed effort to intensify the delivery of vaccines, ITNs, vitamin A, antenatal care, breastfeeding promotion, and case management of pneumonia, diarrhea, and malaria. An evaluation of ACSD showed that, paradoxically, some of these interventions had higher coverage in districts without ACSD than in districts with ACSD (Bryce et al., 2010). In such cases, the program–comparison design does not make much sense, and other design alternatives are required.

Observational before-and-after designs can be strengthened when sufficient data are available. For example, if outcome or impact data for some years before the start of a program are available, time-series or interrupted time-series designs can help control for confounders that do not change over time (Gertler et al., 2016). Evaluators need to provide an explicit rationale for assuming that important confounders do not change over long periods of time.

In recent years, more attention has been paid to the use of propensity-score matching at the individual level to evaluate large-scale programs using large survey data sets (Axelson, Bales, Minh, Ekman, & Gerdtham, 2009; Jalan & Ravallion, 2003; Mensah, Oppong, & Schmidt, 2010). These methods control for *observed* differences between individuals who received treatment and individuals who did not. Evaluators need to discuss the possibility of unobserved confounders that may lead individuals to receive or not receive treatment, and address how these factors may threaten the validity of the results.

In some cases, different approaches may be combined (Gertler et al., 2016). As mentioned earlier, districts may be matched before the start of implementation, and data across several years may be collected. While these approaches strengthen the overall evaluation design and the plausibility of the results, they are also complex and analytically challenging.

The Ecological Dose–Response Design

When a program includes biological or behavioral interventions that are being promoted throughout a country, a dose–response design may be the most appropriate evaluation approach. For example, in the Malawi iCCM evaluation, every district in the country was implementing iCCM delivered by health surveillance assistants (HSAs), and the evaluation assessed the association between the density of iCCM-trained HSAs and care seeking and under-5 mortality reduction at the district level. **FIGURE 18-5** shows the correlation between density of trained workers and changes in care seeking and mortality; no correlation was observed and no significant associations were found after adjusting for potential confounders.

In light of widespread scaling-up of health interventions in many countries, dose–response designs may become increasingly useful in the future. **EXHIBIT 18-6** proposes an approach to such studies.

Randomized (Probability) Evaluation Designs

Randomized designs are the gold standard for achieving internal validity. Because health programs are delivered to groups rather than individuals, cluster-randomized trials are appropriate for their evaluation (Donner & Klar, 2000). Random allocation of clusters of individuals—for example, districts—to receive or not to receive the program or intervention increases the likelihood that the two study arms will be highly comparable in terms of confounding variables, including those that cannot be measured directly.

This situation is particularly likely to hold when the number of clusters available for randomization is large. If the number is small—as is often the case for large-scale programs delivered at the district level—the program and comparison groups may still differ substantially in terms of baseline confounding factors, in spite of randomization. In these cases, constrained randomization may be used to improve comparability of the two study arms (Hayes & Moulton, 2009).

Even programs that consist of packages of proven interventions may be evaluated through randomized trials. Certain programs, such as provision of vouchers, provision of health insurance, or exemptions from user fees, may be unaffordable if the entire

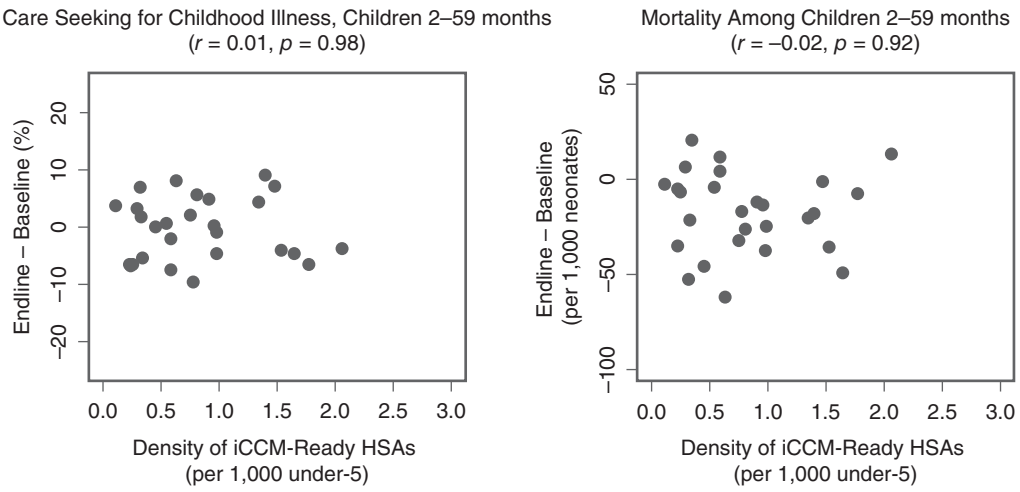


FIGURE 18-5 Correlations between the density of iCCM-ready HSAs and changes between 2010 and 2014 care seeking and mortality for children age 2–59 months in Malawi ($N = 27$ districts).

Abbeviations: HSAs = health surveillance assistants; iCCM = integrated Community Case Management.

Reproduced from Amouzou, A., Kanyuka, M., Hazel, E., Heidkamp, R., Marsh, A., Mleme, T., et al. (2016). Independent evaluation of the integrated Community Case Management of childhood illness strategy in Malawi using a National Evaluation Platform design. *The American Journal of Tropical Medicine and Hygiene, 94*(3), 574–583.

EXHIBIT 18-6 The Evaluation Platform Design

The prevailing evaluation paradigm assumes that programs are implemented in a few districts and not in the rest of the country. Changes in service provision, utilization, coverage, and sometimes health impact are documented over time, and gains in the intervention districts relative to comparison areas are attributed to the program, reflecting the assumption that any improvement is a direct result of program inputs. The counterfactual supposition is that in the absence of the program, outcomes in the target districts would have shown trends similar to those in the comparison areas. Several types of bias may compromise this comparison, including selection biases (e.g., districts chosen for the program may have special characteristics that favor implementation) and confounding factors. This traditional evaluation design remains popular despite its limitations because no feasible alternatives may exist to it; if properly conducted and analyzed, such evaluations often provide valuable information.

Recent experience in evaluating large-scale programs and initiatives suggests that this traditional design has important limitations in the current development context. "Untouched" comparison areas are seldom available because similar biological or behavioral interventions are often being scaled up by other programs or initiatives—with greater or lesser success—in different geographic areas. As a consequence, populations in the comparison area are often exposed to interventions that are similar to those promoted in the program area. For example, in the ACSD evaluation (Exhibit 18-2) in Mali and the iCCM evaluation (Exhibit 18-1) in Malawi, several of the comparison districts were receiving virtually the same interventions as the program districts, with support from other donors. Even in situations where a program is planned to be scaled up in several districts, the reality is that implementation speed tends to vary from district to district, as was the case in the Tanzania ITN voucher study. Therefore, in many circumstances dose–response analyses relating implementation strength to coverage and health impact make more sense than comparisons of program and nonprogram districts.

To further complicate the issue of attribution, health status is influenced by a myriad of factors other than an individual program. At least four categories of contextual factors must be considered (Victora et al., 2005), as suggested in Figure 18-3: (1) preexisting health services, whether public or private; (2) the presence of new health programs other than the one being evaluated; (3) the presence of interventions in other sectors that may affect health (e.g., water, sanitation, or education); and (4) overall socioeconomic and environmental conditions. This broad framework of determinants of health requires that evaluations go well beyond the health sector. In turn, evaluations of all health programs require careful documentation of contextual factors and their incorporation in data analyses and interpretation.

For all these reasons, instead of the current practice of evaluating one program at a time, it makes much more sense to develop a broad evaluation platform to assess all of the multiple programs that are in place within a country. For example, partners can work with national and local governments to support nationwide assessments of programs in different

(continues)

EXHIBIT 18-6 The Evaluation Platform Design *(continued)*

areas: maternal and child health; reproductive health; disease-specific control programs, including those for malaria and HIV/AIDS; and efforts aimed at strengthening health systems (Global Initiative to Strengthen Country Health Systems Surveillance [CHeSS], 2008).

From 2013 to 2017, the National Evaluation Platform (NEP) project implemented NEPs in Malawi, Mali, Mozambique, and Tanzania with funding from Global Affairs Canada and technical assistance from the Institute for International Programs–Johns Hopkins University (IIP-JHU) (Johns Hopkins Bloomberg School of Public Health, n.d.b). The project aimed to implement country-owned and -led NEPs, with a focus on building country capacity to manage, assess, and use existing data to answer evaluation questions as defined by country governments. These country NEPs mapped sources of population and health data along the pathway to impact, as well as other contextual factors; assessed their quality; and worked with governments to identify and respond to priority questions.

In Mali, for example, the Ministry of Health wanted to understand why key maternal, newborn and child health (MNCH) indicators were poorer in one region of the country than in other regions. In response, the Mali NEP partners, including the Center for Research, Study, and Documentation on Child Survival (CREDOS), the National Institute of Public Health (INRSP), the National Institute of Statistics (INSTAT), the National Directorate of Health (DNS), and the MoH Unit for Planning and Statistics (CPS) brought together district-level data from population surveys, from routine health information systems, and on contextual factors to try to understand the factors associated with disparities in district- and regional-level indicator trends over time. IIP-JHU provided technical assistance and capacity building in evaluation question design, data quality assessment, indicator calculation, and multiple approaches to analysis (cross-group comparisons, trend, and regression analyses) to support the work done by the Mali NEP team.

The NEP project was inspired by the evaluation platform approach proposed by Victora and others (2009), which included the following steps:

1. Develop and regularly update a district database that includes existing demographic, epidemiologic, socioeconomic, and health infrastructure variables, derived from sources such as censuses, economic surveys, poverty maps, and service availability censuses. All districts in a country—or in a subnational region—would be included.
2. Conduct an initial survey (or build on existing and future household surveys), to be repeated every three years or so, to measure coverage levels for proven interventions and health status. Ideally this survey would also allow estimation of mortality and prevalence of biomarkers.
3. Establish a continuous monitoring system for documenting provision, utilization, and ideally quality of interventions at the district level, with mechanisms for prompt reporting to local, national, and international audiences.

The experience of the NEP project has shown the potential value of this approach, but also highlighted the challenges in implementing evaluation platforms in a sustainable, country-led way. Development of a comprehensive set of data alone is not sufficient. Statistical and health institutions in many countries do not have the capacity to analyze much of the data that they collect, so a key task for many evaluation platforms will be to build country capacity in evaluation principles; core data concepts; data management, analysis, visualization, and interpretation; and audience-appropriate communication of results. Other challenges include sharing data across institutions and sectors, building ownership and understanding of the platform at senior levels of government (see the *Public Health Infrastructure* chapter for more on data sharing), and ensuring adequate staffing in country institutions to collect, compile, assess, and analyze data to answer evaluation questions.

target population must be covered. In such cases, randomization of individual participants or geographic regions may be possible (Powell-Jackson, Hanson, Whitty, & Ansah, 2014). To perform this type of study, all districts in a country would be listed, possibly stratified into categories of risk (e.g., high, medium, or low frequency of the main study outcome), and then allocated randomly within each stratum either to receive the program under routine conditions of implementation or to not receive it. Politicians and managers—as well as members of the population—need to agree with the process of randomization, as do all involved district authorities and implementation partners.

In practice, a variety of ethical, political, and practical barriers to randomizing districts to receive or not to receive a program may exist (Black, 1996; Victora et al., 2004). For instance, large-scale programs usually encompass biological and behavioral interventions whose efficacy has already been established in RCTs in smaller populations, so that withholding them from population groups for the purpose of evaluation is difficult to justify. Political barriers may arise

from a similar source: Policy makers are understandably reluctant to withhold proven interventions from the population for the purpose of research. Practical barriers include the assumptions within an RCT that the researcher controls the pace and quality of intervention delivery and that the intervention does not change over time—criteria that cannot be met when evaluating strategies or programs that are being scaled up by governments and their partners and often are undergoing fine-tuning in response to early experience or intermediate results from the evaluation.

These barriers to RCTs tend to counterbalance the increased internal validity associated with their use. Thus, it is not surprising that randomized designs are rare in large-scale effectiveness evaluations. In addition, evaluators are often called in when it has already been decided which areas will receive—or are already receiving—the program and, therefore, cannot influence its deployment.

Among the few examples of RCTs used for evaluation programs affecting health outcomes are two Mexican studies, which are summarized in **EXHIBIT 18-7**.

EXHIBIT 18-7 Large-Scale RCTs: The Progresa and Seguro Popular Evaluations in Mexico

Mexican policy makers and researchers have pioneered the use of RCTs for large-scale program evaluation in LMICs. Two large studies have attracted special attention: the evaluation of Progresa, a program that included conditional cash transfers (Rivera, Sotres-Alvarez, Habicht, Shamah, & Villalpando, 2004), and the evaluation of Seguro Popular, a health insurance initiative (King et al., 2009).

Progresa was a large-scale, incentive-based development program with a built-in nutritional intervention. Taking advantage of the lack of resources for covering all eligible families in the country at the same time, the evaluators designed a two-wave, stepped-wedge trial. Of 347 poor rural communities in 6 central Mexican states, 205 were randomly assigned to immediate launch of the program in 1998; another 142 were slated to initiate the program 2 years later (Rivera et al., 2004). Random samples of children in those communities were surveyed at baseline, and again at 1 and 2 years afterward. Participants, all of whom came from low-income households within each community, were selected on the basis of an index of household assets.

Progresa provided micronutrient-fortified foods for women and children as well as health services and cash transfers for the family. The original design sought to provide a 2-year window during which Progresa would be implemented in the intervention communities but not in the comparison communities, but political pressure to accelerate the program led the comparison communities to start receiving the program more than 1 year earlier than originally planned, in late 1999 rather than in early 2001. In addition, nearly half of all young children (younger than 6 months) were lost to follow-up by the time of the 2-year follow-up. Progresa was associated with better growth in height among the poorest and younger infants, but not in the sample as a whole. After 1 year, mean hemoglobin values were higher in the intervention group than in the comparison group, but there were no differences in hemoglobin levels between the two groups at year 2, after both groups were receiving the intervention.

The second evaluation involved an innovative cluster trial of health insurance—Seguro Popular—in Mexico, one of the largest randomized health policy experiments ever carried out. Policy makers in 13 of the 32 Mexican states agreed to join the study. From more than 7,000 clusters (health facility catchment areas) in these 13 states, the evaluators "negotiated access to 74 cluster pairs in seven states, with inclusion based on necessary administrative, political, and other criteria" and then randomly assigned one health cluster from each pair to receive the Seguro Popular program. The main finding was a reduction in catastrophic expenditures in health of 1.9 percentage points. Although program resources reached the poor, contrary to expectations, they did not have any effects on medication spending, utilization of health services, or health outcomes after 10 months of follow-up. The authors of the evaluation (King et al., 2009) recognize that the short implementation period may have precluded documenting an impact on outcomes that are unlikely to change rapidly. They also note that the program varied considerably across areas.

These two Mexican trials make an important contribution to the literature on the evaluation of large-scale programs, but they also reveal that RCTs are not a panacea for all evaluation ills. First, obtaining agreement of policy makers at all levels—from national to local—is not a trivial matter. Moreover, external validity may be affected because of low compliance with randomization at local level. That may have been the case in the Seguro Popular trial, in which 74 cluster pairs were nonrandomly selected from more than 7,000 clusters.

Second, both trials show that getting policy makers and managers to accept randomization is a major achievement, but ensuring their continued support during the trials is also essential. In both studies, the patience of decision makers proved shorter than what would be required for a full assessment of the intervention's impact. Both trials demonstrated that, when taken out of their traditional context—that is, efficacy studies that are completely under the control of researchers—RCTs do not perform well for the evaluation of large-scale, real-life programs.

Stepped Wedge Design

An attractive alternative to the RCT is the stepped wedge design, a variation of randomized cluster trials (Brown & Lilford, 2006). This strategy should not be confused with the stepwise approach to evaluation described in Figure 18-2. A stepped wedge design makes use of implementation timetables that specify introduction of the program or strategy earlier in some geographical areas than in others. This phased rollout allows evaluators to use the later-implementation areas as comparison groups for the earlier-implementation areas.

A hypothetical example of a randomized stepped wedge design follows. Suppose the capacity to train health workers or to procure a commodity is limited, so that only a few districts can be covered in the first year of the program. These areas are selected randomly from all districts in the country, and start to receive the program. In the following year, another wave of districts is selected from those without the program, and so forth, until all districts are covered. At any given time until the last implementation wave is complete, there will be districts with the program and districts without the program; the latter provide a randomized comparison group.

This design has advantages from an ethical standpoint, because it would not have been possible to implement the program in all districts at the same time, and selection of program districts is random. Nevertheless, fewer than 20 such trials have been published in the literature (Brown & Lilford, 2006), and none of these reports have dealt with large study units such as districts. Instead, the study units in most cases were health facilities, vaccination teams, or small communities.

As with standard RCTs, practical constraints to implementing stepped-wedge trials at the district level in LMICs exist:

- It may be difficult to explain the randomization process to district authorities and convince them that there are no hidden allocation biases.
- National governments are often reluctant to leave any district completely devoid of a program. Instead, they may attempt to be "equitable" by allocating some amount of resources to all districts.
- Other bilateral or international organizations delivering health programs in the country already have their favorite districts, and may continue to implement activities that are similar to the program.
- Many infectious diseases demonstrate marked seasonal or cyclical patterns, and it may be difficult to distribute program-implementation waves of districts over time so that temporal patterns do not affect the interpretation of results.
- Implementation timetables are subject to many influences, and often change over time, which may weaken the design.
- This design can be used only for programs or strategies that are unlikely to have effects beyond the border of a district. It would not be appropriate, for example, in evaluations of mass-media campaigns.
- The analysis of this design is particularly complex, and requires strong statistical support during both the design and analysis stages.

These reasons may explain why stepped wedge designs, although available since the 1980s, have never enjoyed great popularity despite their theoretical advantages (Brown & Lilford, 2006).

Occasionally, programs will be implemented in only selected areas of a province or district, due to budgetary or other constraints. In these cases, it may be possible to exploit the employment of a nonrandomized phased implementation (stepped wedge) or arbitrary break in implementation along the borders of districts (Sood et al., 2014). These types of designs, which are called regression discontinuity designs or geographic discontinuity designs, fall into the "plausibility evaluation" category, although they are considered some of the stronger methods for observational studies, and have in some cases closely replicated random designs (Buddelmeyer & Skoufias, 2004).

Defining the Indicators and Obtaining the Data

Documentation of Program Implementation

In terms of the stepwise design described in Figure 18-2, documentation is essential for assessing the first question: Are the interventions and plans for delivery technically sound and appropriate for the epidemiologic and health system context? Documentation will also contribute quantitative information on implementation activities that can help address the second question: Are adequate services being provided at the facility and community level? The ultimate objective of documentation is to go beyond program objectives and goals to understand what actually took place at the population level. Systematic documentation is essential for understanding why a program achieved—or failed to achieve—its intended

effect. Qualitative studies are often a useful adjunct to documentation.

Obtaining detailed data on program implementation is essential because—unlike what happens in tightly controlled, small-scale trials—large-scale programs often fail to deliver what was originally proposed. The ACSD evaluation showed, for example, that whereas vaccines and vitamin A supplements reached high coverage in the target districts, other equally important parts of the ACSD package, such as insecticide-treated nets, were affected by unforeseen stock-outs that resulted in insufficient distribution of these items to the population.

It is often assumed that data on program implementation are readily available from those parties involved in delivering the program. Our experience, however, shows that this is rarely the case. In the Malawi iCCM evaluation, information about the hard-to-reach areas that were supposed to be covered by iCCM-trained CHWs was not possible to obtain. In Burkina Faso, data on drug stock-outs at the CHW and facility level were not available. In the retrospective ACSD evaluation, obtaining data on the early stages of the program proved difficult because records were often incomplete, and some key staff involved in setting up the program were no longer available to be interviewed.

Documentation efforts should strive to cover the following topics, which are drawn from the conceptual framework for the program (Figure 18-1) and reflect the stepwise approach (Figure 18-2):

- Describe the original program proposal—development, adaptation, program strategies and activities, associated policies, and delivery strategies.
- Describe how the original proposal changed over time, and why—for example, whether a result of internal decisions or due to feedback by the evaluation team.
- Quantify the program inputs, processes, and outputs—for example, the number of health workers who were trained, the types of commodities that were procured and distributed, the frequency and duration of stock-outs, the type and frequency of supervision activities, utilization by the target population of the services being provided, and the frequency and nature of community-level activities.
- Describe contextual factors that may affect the program's impact—for example, epidemiologic, health systems, demographic, or sociocultural characteristics.

Methods for documentation include desk reviews of program and policy documents, training materials, workshop and supervision reports, and administrative records related to program inputs and outputs. This step may also require talking to implementers, conducting in-depth interviews, and conducting focus-group discussions.

Where key data on program implementation are not available from program documents, or where interviews with implementers contradict the quantitative program data, primary data collection to describe the state of program implementation may be needed. In the Burkina Faso iCCM evaluation, for example, data provided by health districts showed that the number of iCCM-trained CHWs remained constant over time, following the initial iCCM training. In contrast, interviews with district staff reported substantial turnover among CHWs. To better understand the staffing situation, and to fill gaps in program data with respect to utilization and drug stock-outs, the evaluation team conducted a survey of CHWs in the regions that received the program (Munos et al., 2016). CHWs were asked about their selection, training, and supervision, and data collectors observed their drug stocks and abstracted information from their consultation registers. In addition, in-depth interviews were conducted with a subset of CHWs as well as with caregivers of young children and with the nurses who supervised the CHWs.

Similar assessments of implementation, sometimes referred to as implementation strength assessments or implementation snapshots, are increasingly being conducted during program evaluations to provide high-quality data on program processes and outputs (Hargreaves et al., 2016). These assessments are generally surveys of the front-line providers who are delivering the program interventions. Survey methods are available for this purpose. Typically, a multistage sampling scheme is used to select units within the program areas—for example, health facilities or communities. Within these units, health workers involved in service delivery are then sampled. The content of the assessment will vary somewhat depending on the program activities, interventions, and modes of delivery. For programs delivering health services to populations, key domains may include training, supervision, availability of drugs and commodities, availability of guidelines and other job aids, and utilization of services. In addition to the quantitative assessment, qualitative interviews may be conducted with providers, their supervisors, and program beneficiaries or other community members to better understand challenges in implementing the program or delivering interventions. As discussed in the next section, assessments of the quality of service provision may also be added to the implementation strength assessment.

Measurement of Intervention Quality

Most health programs include interactions between health workers and the target population, with or without the provision of commodities. A behavioral change and communications intervention may include individual or group sessions; a vaccination program will include delivery of the vaccine to the target population; and case management interventions will include interaction between a community- or facility-based provider and a patient, often followed by provision of treatment. While documentation will provide quantitative information on these interactions, it is also important to assess their quality.

Assessments of the quality of service provision are generally conducted in a sample of health workers or health facilities, using the same sampling approach described earlier for implementation strength assessments. For each sampled provider, data collectors typically observe the interactions between the health worker and members of the population. Many survey tools also include a description of the facilities where health workers operate (e.g., the physical space, drug supply, equipment) and incorporate exit interviews to assess user satisfaction. For case management surveys, a second examination of the same patient by a gold-standard examiner is often included, which is then compared with the original conduct of the health worker. The quality of case management for rarely occurring events is sometimes assessed through the use of written or oral scenarios. Further details on how to conduct these surveys, and standard indicators to be used, are available elsewhere (Gouws et al., 2005).

Measuring Coverage

Although population coverage of interventions can sometimes be estimated from routine health system data such as utilization statistics (e.g., dividing the number of births reported by skilled attendants by the estimated number of births in the population), in most instances coverage measurement requires household surveys. This technique is necessary because population movements may occur—for example, mothers from a neighboring district coming to the focus district to give birth—and because the availability of simple data on commodities disbursed (e.g., antibiotics, oral rehydration solution [ORS] packets, or ITNs) does not guarantee that those items reached, and were used by, those in need. In addition, in many LMIC settings, routine health system information on the numerators (services delivered) and the denominators (individuals in need) for coverage indicators can be biased upward or downward due to incomplete records, incentives for over-reporting, or inaccurate census projections that contribute to the denominator. Even with these potential biases, coverage data from routine sources may still be useful for local planning, although their use in an effectiveness evaluation requires a robust assessment of data quality to avoid invalid inferences.

Many LMICs now have regular surveys on maternal and child health, such as Demographic and Health Surveys (DHS; The DHS Program, n.d.b) or UNICEF's (n.d.) Multiple Indicator Cluster Surveys (MICS) that are carried out every three to five years in a given country. In addition, surveys may assess coverage of disease-specific programs, such as malaria indicator surveys (MIS; The DHS Program, n.d.c) and HIV/AIDS indicators surveys (The DHS Program, n.d.a).

These types of surveys include standard indicators, usually defined by multi-institutional or interagency working groups. It is essential that program evaluations comply with agreed-upon indicators, so that the results can be compared with those from other evaluations and studies.

Surveys may be used as part of evaluations in several types of scenarios:

- A national survey (e.g., DHS or MICS) is carried out at a suitable time for the evaluation, and the number of households sampled in the program and comparison areas is sufficiently large to yield precise estimates. This was the case in the Malawi iCCM evaluation, for which DHS data were used as the baseline, and an "MDG survey" conducted by MICS was used as the endline (Amouzou et al., 2016).
- A national survey as described previously is planned, but needs to be oversampled in program and/or comparison areas. This was done in the ACSD evaluation in West Africa (Bryce et al., 2010).
- No suitable surveys are planned, and the evaluation team has to carry out its own survey in program and comparison areas. This was the case in the Burkina Faso iCCM evaluation; while a DHS was available to use as the baseline, the sample size at the district level was insufficient and oversampling in the evaluation districts was not an option (Munos et al., 2016).

Survey samples are usually insufficient for providing precise coverage estimates at the district level. Some exceptions do exist, however, such as the Malawi 2010 DHS and 2014 MDG survey, which sampled 1,000 households per district, and the India National

Family Health Survey, which had a total sample of more than 600,000 households in 2015–2016 (DHS, 2010; National Statistical Office, 2015; India Ministry of Health and Family Welfare, 2017).

Even if the number of sampled households per district is small and few programs are implemented in a single district, pooling across several districts may result in sufficient numbers of individuals to assess coverage. It is often argued that survey results should not be pooled, because few national surveys are designed to provide probability samples at the district level. Nevertheless, these surveys systematically present tables stratified by age, socioeconomic, and ethnic categories, even though the sample was not designed to be strictly representative of such subgroups. In practice, most nationally representative surveys employ implicit stratification within each district by listing enumeration areas in a geographic sequence and systematically sampling these areas; as a result, households included in the sample tend to be spread throughout the districts. By giving due attention to sampling weights, groups of districts where a program was implemented may be separated from a national survey for the purposes of assessing coverage (Rao, 2003; West, Berglund, & Heeringa, 2008).

Surveys can be disaggregated to the district level using either district identifiers in the data sets (which are infrequently available in MICS and DHS and usually provided only in surveys explicitly stratified by district) or geolocation information on clusters. Cluster geocodes are now routinely collected for DHS surveys, with a few exceptions, and are increasingly collected for MICS. To protect the anonymity of respondents, DHS randomly displaces cluster locations by as much as 5 km in rural areas and as much as 2 km in urban areas, with 1% of rural clusters displaced by as much as 10 km (Burgert, Colston, Roy, & Zachary, 2013). In earlier surveys, this displacement made disaggregation by districts problematic; starting in 2009, however, DHS changed the displacement protocol to respect district boundaries, making it possible to disaggregate DHS data to the district level and to regroup districts as needed.

Due to the high cost of surveys, other alternatives for data collection are worth exploring, such as using National Immunization Days (NID) to collect information on coverage or health outcomes related to other programs than those delivered at the NID—for example, breastfeeding promotion or nutritional status (Santos, Paes-Sousa, Silva, & Victora, 2008). As long as coverage of NID attendance is high (say, 90% or so), coverage of other programs can be estimated with little bias.

Measuring or Modeling Impact

Some programs express their main goals in terms of coverage indicators, so impact measurement is not strictly necessary. This is the case, for example, with a vaccination program that seeks to reach high coverage with a vaccine (e.g., rotavirus vaccine) whose efficacy is well established. For such an evaluation, if high coverage is documented, there is no need for expensive and complex assessments of the incidence of diarrhea due to rotavirus.

Even so, many funders wish to go back to their constituencies armed with hard data on the health impact of the intervention—for example, the number of lives saved by their actions. Some initiatives, such as the Catalytic Initiative to Save a Million Lives (UNICEF Canada, n.d.), have actually incorporated their quantitative goals into the program title.

Impact measurement is often included in surveys that also measure coverage. Nevertheless, assessing impact indicators usually makes coverage surveys much more complex. Issues that influence the measurement of impact in program evaluations include the following:

- *How difficult is the measurement?* For example, will impact be assessed through interviews (e.g., full birth histories of women of reproductive age, aimed at assessing child mortality; infant feeding practices), or is there a need for measurements or collection of biological specimens? Measuring child underweight with a weighing scale is easier than measuring stunting, which requires substantial training of interviewers and use of more sophisticated equipment. Drawing blood samples adds another level of logistical complications, including equipment requirements, interviewer safety concerns, and issues related to processing, transportation, storage, and analysis of samples. If data on causes of death are needed—often an important consideration for attributing success to a program—additional requirements will arise in terms of interviewer training and sample size.
- *How rare is the event being measured?* The frequency of the impact measure in the population is key for determining sample sizes. Rare events (e.g., mortality or, even worse, cause-specific mortality) require larger sample sizes than more frequent events (e.g., stunting or malaria parasitemia in an endemic area). Some events (e.g., maternal mortality) are so infrequent that the samples required are prohibitively large, leading evaluators to rely on proxy measures such as coverage

of interventions known to reduce mortality. In the case of cluster surveys, evaluators also need to take into account the "design effect" (Cochran, 1977): Events that are evenly spread throughout the population require samples that are smaller than events that are highly clustered, as is the case for some transmissible diseases. For impact measures such as mortality rates, sample sizes may be reduced by using deaths in the last two or three years before the survey, rather than deaths in the last year. This approach requires full implementation of the program throughout the mortality measurement period, as otherwise its full effect will not be picked up.

- *How large is the effect to be detected?* Picking up large effects requires smaller samples than detecting small effects. For example, detecting a 25% reduction in under-5 mortality over a certain period, relative to baseline levels, will require much smaller samples than detecting a 15% decline in this metric. Even larger sample sizes are needed to detect a 25% difference between the annual rates of mortality decline in program and comparison areas—for example, 10% and 8% per year, respectively.

Given the complexities of estimating impact through household surveys, evaluators or donors may propose the use of routine health systems data or civil registration and vital statistics data to assess impact. In principle, these routine data have a number of advantages over household surveys: They are often available at the district level, on at least an annual basis, and at little to no cost. The greater availability of data can facilitate more complex plausibility evaluation designs and increase the power of an evaluation to detect an effect. Nevertheless, as with routine coverage data, routine data on mortality, morbidity, and nutritional status in LMICs often suffer from important quality limitations that can threaten the validity of an evaluation. Routine mortality or morbidity data can be, and have been, used for high-quality impact evaluations (Rasella, Aquino, Santos, Paes-Sousa, & Barreto, 2013), but a thorough assessment of the quality of these data, including their completeness, internal validity, and external validity, is essential before using them in an evaluation (see, for example, Willey et al., 2011).

Recent advances have led to the development of modeling tools that allow the estimation of under-5 mortality trends based on measured changes in coverage. **EXHIBIT 18-8** describes the Lives Saved Tool

EXHIBIT 18-8 The Lives Saved Tool: An Example of Mortality Modeling Software

Modeling approaches are increasingly recognized as valuable for planning and evaluation in public health (Garnett, Cousens, Hallett, Steketee, & Walker, 2011). The Lives Saved Tool (LiST) is a model that estimates changes in mortality on the basis of intervention coverage data. LiST allows users to estimate the impact of scaling up proven interventions by defining and running multiple country-, state-, or district-specific scenarios (Walker, Tam, & Friberg, 2013). LiST was designed for a range of MNCH stakeholders, including ministry of health personnel, program managers, and NGO partners; the intent was to enable them to combine the best scientific evidence on the effectiveness of interventions with information about the causes of death and current coverage of key interventions to inform their planning and decision making. This evidence-based platform has been widely used to help prioritize investments and evaluate existing programs (Boschi-Pinto, Young, & Black, 2010).

LiST uses data on baseline mortality rates, causes of death, and intervention coverage, together with demographic data and intervention effectiveness, to estimate the number of maternal, under-5, and neonatal deaths that can be averted, both by cause of death and by intervention, as intervention coverage increases. It also estimates levels and changes in mortality rates, disease incidence, and the prevalence of risk factors such as undernutrition (stunting and wasting), by cause and by intervention. LiST is programmed as part of a larger software package called Spectrum, which has been widely used for 25 years and is designed to predict population changes over time by age and sex (Stover, McKinnon, & Winfrey, 2010). Spectrum uses the population projections developed by the United Nations Population Division, which facilitate estimates of population growth through 2050. Through Spectrum, LiST is linked to other modules such as the AIDS Impact Module (AIM), which models UNAIDS data for each country, and FamPlan, which estimates the impact of family planning parameters on outcomes such as the number of births.

LiST and other software can be applied in two ways for the evaluation of large-scale programs. First, they can be used to model or assess whether the interventions being promoted by the program are appropriate given the cause of death profile in the program area. For example, LiST modeling can estimate the potential change in mortality and health status if scale-up efforts are successful. This information helps evaluators answer the first question asked in the stepwise model described in Figure 18-2: Are the interventions relevant to the epidemiologic context? This approach was used at the beginning of the iCCM evaluations conducted in Burkina Faso and Malawi (Bryce et al., 2010).

The second, and more important, use of LiST and similar software is to translate changes in coverage measured through surveys into expected mortality impact. When direct impact measurement is not feasible or too expensive under given conditions, LiST can provide a reasonable estimate of the expected decline in mortality based on the measured coverage gains. This approach was used in the Burkina Faso iCCM evaluation (Munos et al., 2016).

The LiST website (http://livessavedtool.org/) provides access to the most updated version of the software. In addition to links for downloading the tool, other resources and technical documentation are posted to guide users. An embedded manual provides instructions on the software's use, as well as citations for intervention reviews and validation studies.

(LiST), which is becoming widely used for this purpose (Bryce et al., 2010).

Describing Contextual Factors

In impact evaluations, contextual factors are variables external to the program that can confound or modify its observed effect (Victora et al., 2005). Confounding occurs when changes in these external variables differ in program and comparison areas. For example, crop failures, natural disasters, or establishment of new health facilities (or presence of other programs) may occur in one area, but not in others. The ACSD evaluation showed an apparent impact of the intervention on undernutrition prevalence, but documentation of contextual factors identified a famine in some of the comparison districts; once these effects were excluded from the comparison group, there was no longer any evidence of a positive impact of the program (Bryce et al., 2010). Control of confounding factors is essential for improving the internal validity of the evaluation.

Effect modification occurs when certain baseline conditions—for example, the level and causes of mortality, or the strength of health systems—either contribute to or detract from the observed impact of the program. Effect modification is particularly relevant to external validity, or the ability to generalize from the evaluation findings. In the iCCM evaluation in Malawi, high levels of care seeking for sick children from health facilities may have limited the potential impact of the intervention, and partially explained the finding that the program was not associated with a reduction in under-5 mortality (Amouzou et al., 2016). Given that the program was reasonably strongly implemented, we might expect to see a much greater impact on mortality if it were implemented in a setting with lower levels of facility care seeking.

Data on contextual factors may come from a variety of sources. Most LMICs already have a number of databases maintained by governmental, international, or partner institutions, with information disaggregated at the provincial or district level. **TABLE 18-2** identifies sources of information on a variety of contextual factors; this listing was prepared for an evaluation of the Catalytic Initiative to Save a Million Lives in Mozambique.

If quantitative information is available for all geographic areas included in the evaluation, contextual factors may be formally included in the statistical analyses models as covariates. Even if information is incomplete or unquantifiable, contextual factors may help interpret the evaluation findings.

The process of obtaining information on contextual factors is part of the documentation exercise described previously. It can be often done by the same team that is collecting data on program implementation.

Measuring Costs

Cost data can be used in many different ways to inform decisions about public health programming, and are a key component of large-scale program evaluations. They contribute to understanding the budgetary implications of health programs and are needed to perform economic analyses such as those assessing cost-effectiveness. Cost data can help answer questions not only about how much money programs cost, but also about where the costs are incurred (e.g., by patients, at a particular service level, at higher levels), how programs operate, and which factors are hindering program scale-up. Moreover, such data can help answer questions about the relationship between costs and other aspects of health care, such as the quality of services, that are public health goals in their own right (Bishai, Mirchandani, Pariyo, Burnham, & Black, 2008). Finally, cost data can be used in conjunction with data on utilization and household socioeconomic status to understand the equity implications of a program.

FIGURE 18-6 presents a conceptual diagram of the total costs involved in the scale-up of health programs. The *x*-axis represents the time from before the start of the scale-up program until the target levels of coverage (or mortality reduction) have been achieved, or until the end of the evaluation. The *y*-axis represents the total costs of health services, where higher levels of coverage are associated with higher levels of total costs.

TABLE 18-2 Examples of Data Available from Existing Databases at the District and Provincial Levels in Mozambique and Sources of Information

Category	Examples	Source
Socioeconomic factors	Household assets Family income and poverty Parental education and occupation Unemployment Land tenure Economic crises (inflation rates, crop failures, floods)	2007 census Economic censuses and surveys National Institute of Statistics
Demographic factors	Population density Fertility patterns Family size Ethnic groups	2007 census
Environmental characteristics	Water supply Sanitation Urbanization Housing Rainfall Altitude	2007 census National Meteorological Institute
Baseline health conditions	Under-5 mortality Prevalence of malnutrition HIV prevalence Malaria transmission patterns	2007 census 2008 MICS Malaria and HIV surveys
Health services characteristics	Availability of health services (e.g., hospitals, clinics) in public and private sectors Population/facility ratio Health worker staffing patterns Health worker pay Drug supply Baseline utilization rates Availability of referral services Strength of district health management team District health budget (overall and for child health)	Health Metrics Network Ministry of Health Information Systems UNFPA Needs Assessment Survey WHO Service Availability Mapping
Presence of other projects and programs that may affect health status	Micronutrients Indoor residual spraying Immunizations HIV programs Others	UNICEF World Health Organization Network of Organizations Working in Health and HIV/AIDS (NAIMA) Official Development Assistance to Mozambique Database (ODAMOZ)

Five cost categories are incorporated in the framework depicted in Figure 18-6. *Area A*, the rectangular region above the *x*-axis, represents the total amount of resources used for health services at the beginning of the observation period.

Area B represents the change in costs normally associated with secular trends in most health systems. It can also be thought of as the change in area A over time due to changes in the population, inflationary pressures, and the introduction of new

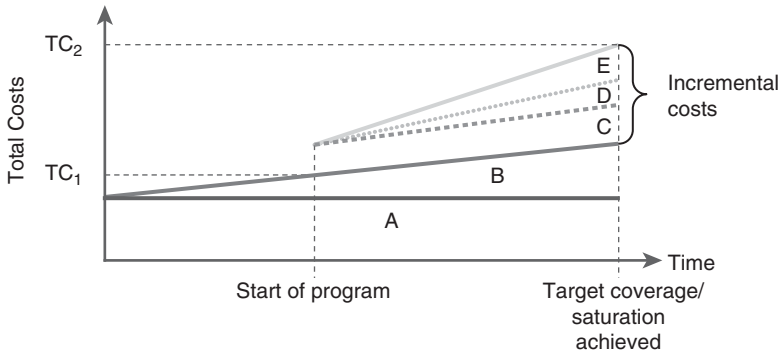


FIGURE 18-6 Conceptual framework for assessing the cost of rapid scale-up.

Reproduced from Johns, B., Opuni-Akuamoa, M., Walker, D. (n.d.). Assessing the costs and cost-effectiveness of rapid-scale up for the maternal, neonatal, and child health: The economic component of the impact evaluation strategy for the Catalytic Initiative. Retrieved from: http://www.jhsph.edu/research/centers-and-institutes/institute-for-international-programs/_documents/rapid_scaleup/CI_costeffectiveness.pdf. With permission of the author.

technologies, among other factors. In Figure 18-6, secular trends are seen as increasing total costs, which is normally—albeit not universally—the case. This relationship does not conceptually imply that the cost per patient or per visit is necessarily increasing, nor is there any particular reason to believe that secular trends will affect costs in a linear way as shown in the diagram; this depiction is provided here for simplicity's sake. Note that secular trends will affect the costs of a health program both before and after its implementation.

The costs in Areas A and B represent multiple kinds of costs:

- **Service delivery level costs:** These costs are incurred in the direct delivery of health services. They include costs incurred at health facilities, but may also include any activity that involves direct contact with patients or the population, such as indoor residual spraying or other "campaign"-type interventions, community delivery of interventions, and so on.
- **Above service delivery level costs:** These costs are not directly linked to the delivery of health services, but rather support the delivery of health services. They may include activities such as training and mentoring, media messages, supervision, demand generation, advocacy, policy development, monitoring, and central laboratory services. Studies suggests that above service delivery level costs can be a substantial proportion of the costs for many programs, even those focused on treatment of diseases (Marseille et al., 2012).
- **Recurrent costs:** Costs that are incurred on a regular basis, defined as at least once a year, are considered recurrent costs. Recurrent costs are typically incurred for items that are consumed after one use (such as medicines, supplies, gasoline, and laboratory reagents) or that need regular payment (such as staff salaries and utilities).
- **Capital costs:** Costs for items that are durable—that can be used multiple times and over multiple years—are considered capital costs. Examples include buildings, medical and office equipment, furniture, and laboratory machines. For most cost assessments that are not directly tied to budgetary expenditures, capital items are *annualized*—that is, the cost of the capital is averaged (in some way) over its expected useful life to determine the equivalent cost of the capital item if it were paid for every year (rather than once).

Both service delivery and above service delivery costs may be recurrent or capital costs. Summing the different kinds of costs gives the **total costs** of the intervention or program (TC_1), encompassing all of the costs incurred in association with a program, before the start of the program's scale-up or the introduction of a new intervention.

Area C in Figure 18-6 represents the **start-up costs** associated with the scale-up; these costs are specific to the scale-up/new program and, therefore, are added to the costs represented in areas A and B. Start-up costs include costs for activities that are incurred once, or infrequently, during a program. Examples include intensive training or policy development. Start-up costs can be thought of as capital items, albeit with start-up costs generally being incurred for activities instead of for tangible items. Like capital costs, start-up costs are typically annualized; thus, they are shown in Figure 18-6 as remaining constant over time, although this again may not necessarily be the case.

Area D represents the above service delivery level costs associated with the scale-up. These costs are specific to the scale-up program and, therefore, are added to the costs represented in areas A and B. Again, there is no predetermined pattern for how these costs will change over time; for convenience, they are depicted in Figure 18-6 as increasing with the expansion of scale-up activities.

Area E represents the additional service delivery level costs due to the increased service utilization associated with the scale-up/new program. Again, if the scale-up/new program increases coverage, total costs should increase. An exception may occur with programs aimed at preventing diseases. In these cases, the decreased need for treatment may decrease service delivery level costs.

The costs associated with areas C and D may be estimated using four different designs:

- **Documentation approach:** Costs are estimated by documenting the costs of activities and then identifying which costs are unique to the scale-up. This process is similar to the process discussed in the section on documentation.
- **First difference approach:** The costs for areas C and D are estimated by first calculating start-up and above service delivery level costs at baseline and then subtracting these baseline costs from the total start-up and above service delivery level costs observed after implementation. This approach ignores secular trends.
- **Aggregate difference approach:** Costs are estimated by subtracting the start-up and above service delivery level costs in comparison districts (if there are any) from these costs in the scale-up districts after scale-up has occurred. This approach assumes that the baseline costs in the comparison and rapid-scale up areas are the same, and is essentially the same as a plausibility design.
- **Before-and-after comparison design:** The two previous methods can be combined to control for secular trends. With this approach, start-up and above service delivery level costs are collected before and after rapid scale-up in both the comparison (if there are any) and implementation districts. The difference in costs in the comparison districts is interpreted as the secular trend in start-up and above service delivery level costs, and subsequently taken out of the difference in costs between the two time points in the scale-up districts (Adam, Bishai, Kahn, & Evans, 2004).

Similar methods are used to estimate the costs that make up area E, although further adjustments may be necessary. Tracking time-series cost data by activity may allow for the calculation of total, incremental, and unit costs in both the scale-up and comparison areas and by different levels of coverage, because data on changes in the utilization, population, and the unit costs will be available.

Incremental costs (the sum of areas C + D + E) are the new costs associated with the program being evaluated. In some cases, separating the incremental costs from the total cost is relatively straightforward. For example, training associated with the new program can be attributed directly to the program. In other cases, separating incremental costs from the total cost is considerably more complex. For example, the new program may increase the utilization of health facilities. In this case, items that existed before the new program (such as health facility staff and infrastructure) are utilized for the delivery of services associated with the new program. Similar to total costs, incremental costs are usually positive, except for some preventive programs, where incremental costs may be negative ("cost-saving").

Unit costs (sometimes referred to as average costs) are variously defined, but can be thought of as the cost per service provided (e.g., cost per outpatient visit) or the cost per beneficiary (e.g., cost per child). Unit costs can be derived from both the total cost and incremental costs. In either case, data on the total or incremental costs and the number of services provided, people reached, or population are needed to calculate the unit costs. While total costs for non-preventive interventions will usually be greater after the implementation of a program than before the program's introduction, unit costs at health facilities may be lower after program implementation than before program launch (even if total costs increase). This will happen if, for example, health facilities existed but were relatively underutilized before a program. In these cases, health staff and infrastructure were serving relatively few patients, and the cost per patient at health facilities would be relatively high. The new program may increase utilization, without necessarily incurring the need for more staff or infrastructure. In such a case, the costs of health staff and infrastructure are now spread over relatively more patients, which lowers the cost per patient. If this lower unit cost is not fully offset by the incremental cost per patient associated with the new program, then the total unit cost will be lower after the program compared to before the program.

Using the full change (comparing the baseline to full implementation—that is, the first difference approach) in unit cost in the intervention areas would assume that all changes in the unit cost of service delivery and the utilization rate are attributable to the intervention, which ignores the influence of secular trends. Secular trends may be controlled for using comparison areas (i.e., via the aggregate or difference-in-difference approaches described earlier).

Thus, costs, like the outcomes or impact of a program, are subject to confounding by secular trends and contextual variables, and similar study designs should be considered for costs as for other aspects of

the evaluation. Sometimes, however, policy makers may demand less rigor in the evaluation of costs than for other aspects of the evaluation; for example, they may accept a plausibility design for evaluating costs, but not accept this approach for assessing impact.

Different research questions and designs will determine which categories of cost data need to be collected. Generally speaking, the incremental unit cost of an intervention is the metric required for a cost-effectiveness analysis, but may not be sufficient for an affordability analysis. On the one hand, when assessing the costs of a new technology, such as a new vaccine or diagnostic test, it may be sufficient and feasible to collect only the incremental costs associated with the technology; it may suffice to use a mathematical model to estimate the cost savings associated with the new technology. On the other hand, interventions with more complicated causal pathways, or that influence demand for health services, may require collection of total costs in intervention and comparison areas.

Given the concerns related to estimating area E in Figure 18-6, it is important to clarify which factors may influence costs apart from the program of interest. Three kinds of variables may potentially affect costs: patient-level variables, service delivery characteristics, and contextual variables (Hussey et al., 2009). In the next subsections, we describe considerations for contextual factors that are specific to costs.

Patient-Level Costs

A great deal of the existing cost-effectiveness literature has been devoted to controlling for differences in costs due to variation in patient-level characteristics; most of this work has been conducted in high-income countries. Variables that have been found to be important include case mix or severity of illness (Hofer et al., 1999; Powell & Hampers, 2003; Tucker, Weiner, Honigfeld, & Parton, 1996), sex, age, living conditions, race/ethnicity, and other socioeconomic variables (Grieve, Nixon, Thompson, & Normand, 2005; Hussey et al., 2009).

Service Delivery Characteristics

Facility-level variables encompass factors that affect costs at the provider level. Quality is one frequently cited example (Amorin et al., 2008; Bishai et al., 2008), although one review found that many authors fail to control for it when examining efficiency measurement (Hussey et al., 2009). The scale of services (e.g., the number of patients seen) and the number and types of outputs for a particular health service provider (i.e., the scope of service provision) may influence the unit costs and total costs of delivering those services (Jacobs, Smith, & Street, 2006). Technical efficiency—the ability to maximize outputs for a given set of inputs or to minimize inputs for a given set of outputs—may also influence unit costs at the facility level (Adam, Ebener, Johns, & Evans, 2008; Adam et al., 2005; Baltussen et al., 2003; Bryce et al., 2005; Jacobs et al., 2006). Other variables that have been used to control for costs include the way in which a health provider is reimbursed (Grieve et al., 2005) and the type of ownership (Hollingsworth, 2003), although the latter consideration really reflects technical efficiency.

The effect of economies of scale on unit costs has been shown to be important for primary health care in LMICs (Berman, Brotowasisto, Nadjib, Sakai, & Gani, 1989). It has also been increasingly studied in disease-specific contexts, where this factor has been shown to influence costs (Guinness et al., 2005; Kumaranayake, 2008; Marseille et al., 2004; Menzies, Berruti, & Blandford, 2012; Valdmanis, Walker, & Fox-Rushby, 2003). Technical efficiency has been less frequently studied, although at least one study from an LMIC indicates that health centers may have different costs due to different levels of technical efficiency (Masiye et al., 2006). Quality may either increase costs or lower costs. All three of these factors—economies of scale, technical efficiency, and quality—may influence the costs of health services.

Contextual Variables

Variables that are outside the control of the health provider, yet reflect the environment in which the provider operates, may also influence the costs of services. In many cases, these variables are similar to those for outcome and impact evaluations. For example, income and other socioeconomic factors may affect utilization of services (Adam et al., 2009) and the costs of staff salaries (Jacobs et al., 2006), while the geographic features of an area may influence the cost of transport, supervision, and training as well as patients' ability to access care due to the need to travel for greater distances and over difficult terrain (Johns & Torres, 2005; Over, 1986). The degree of competition between providers within an area may also influence efficiency and therefore costs, yet be outside the control of an individual provider (Jacobs et al., 2006). Thus, data on factors that may influence costs, as well as data on the costs themselves, need to be collected and analyzed at baseline to establish the extent to which they influence costs, or may result in differences in costs after the program has been implemented.

Best practice is to collect both service delivery and above-service delivery level costs, as well as capital and start-up costs. Of course, collecting comprehensive,

accurate, and precise cost data can itself be a costly activity, and the cost of data collection and analysis should be taken into account when planning the evaluation (Wilson, Mugford, Barton, & Shepstone, 2016). Fully describing the strengths and weaknesses of different study designs, as well as the costs of collecting the data necessary for the different study designs, up front to the funders of the study and other target audiences is a necessary part of deciding the approach to assessing costs.

Previous studies of the costs of large-scale child health interventions in LMICs have, by and large, relied on cross-sectional data collection after the program was implemented, as was the case in three of the four countries participating in the IMCI evaluation (**EXHIBIT 18-9**). Although these studies plausibly established that IMCI did not increase the cost of child health programs by a great deal in comparison with areas that did not implement IMCI, they did not ascertain the costs of the IMCI program.

Data Collection Methods

Three principal methods are used to estimate costs: modeled costing, top-down costing, and bottom-up costing. Within each of these methods, multiple sources or ways of collecting data may be employed. These techniques are complementary methods that can be applied simultaneously to obtain a comprehensive collection of cost data, and for cross-checking the results from each strategy. Modeled cost analysis should be done only before or at the start of the

EXHIBIT 18-9 The Cost of Child Health Services: The Multi-Country IMCI Evaluation

The cost of child health services delivery has been most recently and notably studied in conjunction with the evaluation of the IMCI program in Bangladesh, Brazil, Tanzania, and Uganda (Adam et al., 2005; Adam et al., 2009; Amorim et al., 2008; Bishai et al., 2008). In this assessment, costs were estimated from the societal perspective (providers and households).

Standard MCE cost questionnaires were developed and adapted for use in each country. They consisted of four questionnaires to collect data at the national, district, facility, and community levels, respectively. At national level, data on the following IMCI-related activities were collected: planning and orientation meetings, preparation of IMCI training materials (including translation, adaptation, and printing of guidelines and training materials) and training of national trainers, and administration. Data collection at the district level included training of trainers related to IMCI, supervision for IMCI activities related to under-5 children, and administrative costs of under-5 care. The cost questionnaire at the first-level health facility collected data on personnel (including volunteer labor, if any), drugs, capital items, number of patient visits by type, and staff time allocation based on a time-and-motion study. Data collection of household costs was based on a two-week morbidity module in a household survey. Information on costs incurred during under-5 illness episodes in the two weeks prior to the survey were collected for each child in the household.

In Brazil, Tanzania, and Uganda, no baseline cost data were collected because the evaluations were commissioned after implementation of the program was already underway, when it was no longer possible to collect baseline costs. Therefore, costs were collected after the implementation of IMCI had occurred in the intervention districts, and other matched districts without IMCI were used as comparison districts (Adam et al., 2005; Adam et al., 2009; Amorim et al., 2008; Bishai et al., 2008). Only in Bangladesh, where a fully prospective IMCI evaluation was carried out, were baseline costs collected (S. Arifeen, personal communication). The need to collect baseline costs is another strong reason for undertaking prospective assessment, rather than ambispective or retrospective evaluations.

Given the fact that the evaluations from Brazil, Tanzania, and Uganda were not prospective and, therefore, lacked baseline data, the results proved somewhat difficult to interpret. For example, the study from Tanzania found that the costs at the household and the health center levels (primary healthcare level) were almost the same between the IMCI and comparison areas, while costs at the hospital level and at the district level were higher in the comparison areas than in the IMCI districts. Given the nature of the data collection, the reasons for this difference could not be explained; the authors concluded it could be due to decreased referrals after IMCI training, or to factors not related to IMCI (Adam et al., 2005). Without baseline data, it is difficult to know, from this study, what the costs of IMCI were, and to what extent they increased (or decreased) the costs of providing child health care.

A similar analysis done in Brazil found that costs were higher at the hospital and municipal levels in the IMCI areas than in the comparison areas, but lower at primary healthcare levels, although the authors did not find a statistically significant ($p < 0.05$) difference between the two areas in terms of the total costs (Adam et al., 2009). The study in Brazil tried to control for other differences between IMCI and control areas by including variables for facility location and population in the catchment areas of the health centers. This construction suggested that IMCI is endogenously related to total costs via the number of visits per child. That is, IMCI may improve quality, so that people become more likely to attend the clinic; this greater volume of patients then increases total costs but also alters the unit costs of delivering services (Adam et al., 2009).

scale-up program. Top-down costing is usually done retrospectively—frequently for the last year of the intervention, although it could be carried out on a repeated basis (or even before the start of the program being evaluated, so as to establish baseline costs). Bottom-up costing is usually cross-sectional, so it could also be performed on a repeated basis.

Model-Based Cost Estimates Using Existing Data

Models such as WHO's OneHealth Tool (http://www.who.int/choice/onehealthtool/en/) can be used to derive an ex ante cost-effectiveness ratio to inform policy makers of an expected level of cost-effectiveness for their planned activities. Both the unit costs of delivering services and the program-level costs of operating the scale-up effort need to be estimated to carry out model-based cost or cost-effectiveness analyses. Full provider- and patient-perspective analyses may not be feasible given the available data. Any comparison with later results should be adjusted to account for the perspective used in the initial analysis.

Modeled unit costs can be estimated based on any of the following methods. The preferred methods, which reflect the most context-specific data, are listed first. Nevertheless, different methods can be used to derive unit costs, to the extent possible, so as to perform sensitivity analysis.

1. Unit costs can be collected during a baseline survey of health services providers (e.g., health centers, hospitals, and minimally trained health workers in the community). If such data are collected during the evaluation or are available from another study recently undertaken in the country, they can be used either directly or with adjustments due to anticipated changes in utilization.
2. Unit costs can be based on estimates made in previously undertaken econometric studies. It is recommended that WHO-CHOICE (http://www.who.int/choice/costs/en/) unit costs be used because they are available on a country-by-country basis. WHO-CHOICE unit costs include human resources but exclude drug items; thus, excluded costs must be estimated using method 3.
3. Unit costs can be estimated based on treatment protocols (and estimated staff time if methods 1 and 2 are not used). Although this method suffers from the bias that medical personnel rarely have 100% adherence to treatment protocols, it does allow for a relatively quick assessment of the potential costs of delivery of the program. An alternative, and more desirable, approach is to elicit resource usage from providers using vignettes; this approach requires additional time and resources, however.

Two principal methods of measuring costs are used: top-down costing and bottom-up costing. These methods are seen as complementary methods that can be used to obtain a complete collection of cost data and for cross-checking each other.

Bottom-Up Cost Analysis

Bottom-up costing involves three aspects:

1. An ingredients approach, in which the total quantities of goods and services employed in delivering health care are measured, and multiplied by their respective input prices.
2. Surveys of patients and/or caregivers to collect information on the costs associated with accessing and receiving care. Costs at this level reflect both out-of-pocket payments made to access and receive care and the opportunity costs of patients' and caregivers' time.
3. Detailed review of program accounting and expenditure reports to estimate above service delivery level costs.

Top-Down Cost Analysis

The top-down approach focuses on financial expenditures data, linking these data to other resources such as staff and services and population data. Data are collected by reviewing records and conducting interviews at pertinent governmental levels about the financial flows and costs of health services. The starting point for top-down cost analysis is the total cost (e.g., of a health center, hospital, or health program), which is then parsed and allocated to particular services.

Top-down and bottom-up costing approaches often produce different results due to differences in what is included and different methods of measurement (Cunnama et al., 2016). Some evidence indicates that increased funding, by itself, does not always result in more or better services. Notably, various barriers to the effective usage of funds—such as leakage, corruption, and poor incentive structures—can have a detrimental impact on the effective use of funds (International Institute for Population Sciences [IIPS] and Macro International, 2007). Further, use of some resources, such as electricity, is difficult to directly observe using bottom-up methods. Thus, the costs collected using the bottom-up approach may "miss"

some expenditures and provide an unrealistic estimate of the true costs of services. Conversely, bottom-up approaches provide more details on how services are being delivered. In cases where the costs for the treatment of specific diseases are needed for an evaluation, employing at least some bottom-up cost methods in combination with top-down approaches is necessary (Hendriks et al., 2014). The top-down and bottom-up approaches provide different bases from which to estimate total costs; thus, they represent complementary means of assessing the accuracy of the data collected.

Allocation Methods

Some resources (e.g., staff, facility space, transportation) may be fully dedicated to the program being evaluated, in which case their full costs should be included in the estimation. In contrast, overhead (capital and administrative) costs may need to be separated between the program being evaluated and other health or development activities. For example, structural overhead costs, such as for staff not involved in direct patient contact, may be allocated according to utilization variables (e.g., the percentage of visits that involve children). In contrast, equipment, facility costs, and staff costs that are shared between different types of patients are usually allocated according to the amount of staff time spent on particular activities.

The gold-standard methodology for determining the proportion of time staff spent on particular activities comprises time-motion studies (Bratt et al., 1999). While provider interviews tend to underestimate nonproductive time (Bratt et al., 1999), this bias may not be as important when assessing relative time spent between activities. Time-motion studies were part of the IMCI evaluations in Tanzania, Brazil, and Bangladesh. Higher standards for accurate allocation of costs, which generally involve more time and effort to collect, should be applied to the largest of cost categories, while less precise methods can be used for more minor costs (Tan et al., 2009).

Incorporating Equity in the Evaluation

There is growing concern that programs and interventions may fail to reach those who need them most, and that overall progress in health indicators can hide important gender, socioeconomic, or ethnic group differences. Assessing the impact of a program on equity, therefore, is becoming increasingly more important.

The focus on equity can start at an early stage in the evaluation, by assessing whether programs are being deployed in those districts where they are most needed (i.e., in those areas with the highest mortality and the lowest standard of living). These analyses require combining data on baseline contextual factors with implementation strength—an approach that is further discussed in the later section on data analysis.

Another step to address equity concerns is to ensure that—within program areas—the evaluation team can document the degree to which different population subgroups (stratified by gender, socioeconomic position, urban versus rural residence, and ethnic group, for example) are benefiting from the interventions being promoted. Similar information should be collected for nonprogram areas. This effort requires ensuring that the data collection tools used for the evaluation incorporate the information required to classify families or individuals according to the stratification variables to be used in the analyses. Simple and reasonably accurate measurement of socioeconomic position is now possible in household surveys, using principal component analyses of household assets to classify families into five equal groups, or wealth quintiles (O'Donnell, Van Doorslaer, Wagstaff, & Lindelow, 2007). Because it is possible to calculate summary indices of inequality that use information from the whole population rather than a single subgroup (see the later section on data analysis), in most cases there is no need to expand the original sample sizes to perform such analyses. The evaluations described in Exhibits 18-2 and 18-3 incorporated equity analyses using these tools (Bryce et al., 2010; Hanson et al., 2009).

Incorporation of the equity dimension in program evaluations is no longer an option, given its importance and the wide availability and simplicity of existing tools. Rather, it is a requirement for public health interventions.

Carrying Out the Evaluation

In prospective evaluations, there is a need for continued interaction between the evaluation and implementation teams.

Starting the Evaluation Clock

In a prospective evaluation, documentation of program implementation will help assess how rapidly the interventions are being rolled out. This information should be used to plan the timing of post-baseline data collection on coverage and impact indicators. If outcomes are measured too early, there is a risk that the full effect of the program may not be captured. This was the case, for example, in a randomized evaluation

of the Mexican Seguro Popular program, in which—for reasons beyond the control of the investigators—the endline survey took place only 10 months after the program was launched (King et al., 2009).

In the IMCI evaluation in Tanzania, formal meetings between the evaluation and implementation teams were held to decide whether the "evaluation clock" should be started—that is, whether data collection for assessing outcomes could be carried out. In the retrospective ACSD evaluation, similar discussions led to the identification of three periods: baseline (prior to the program), phase-in, and endline (program fully implemented); each period lasted 18 to 24 months. The main analyses of program impact included comparisons between the endline and baseline phases.

Feedback to Implementers and Midstream Corrections

In classical efficacy trials, the interventions are defined a priori and remain unchanged during the study. In contrast, for large-scale programs including proven interventions, midstream corrections are commonly undertaken to increase coverage and improve quality. Such corrections may result from the implementation team's own assessment of how the program is performing, based on internal monitoring. Changes may also arise in response to feedback from the evaluation team, as indicated in the earlier discussion of documentation. In prospective program evaluations, it would be unethical for the evaluators not to provide feedback on obvious implementation shortcomings.

Feedback from evaluators may also improve the program. This type of change may affect the external validity of the study findings, but in light of the ethical imperative this difficulty is unavoidable. To minimize this problem, it is essential that feedback should be provided in a structured manner—for example, in one session conducted every three to six months. All feedback and any resulting actions should be carefully documented and reported as part of the final evaluation results. Plans for feedback should be developed and agreed upon with program implementers in advance. In the iCCM evaluations, the process for disseminating feedback varied somewhat across countries, but feedback was provided first at the national level via either annual stakeholder meetings or "validation" workshops conducted for specific data collection activities. These national meetings also provided a venue for discussing the implications of interim results for both the program and the evaluation. Dissemination of this feedback at the district level was led by the Ministry of Health, in some cases with support from the evaluation team.

Ongoing feedback is not relevant to retrospective evaluations. Nevertheless, whatever the design of the evaluation, it is important to provide feedback to implementers and policy makers at the end of the study. This point is discussed further in the section on dissemination of evaluation results.

Linking the Independent Evaluation to Routine Monitoring and Evaluation

All programs must include internal monitoring and evaluation systems. Monitoring typically includes documenting program inputs and outputs—for example, the number of health workers trained, the number of ITN vouchers distributed, or the number of mothers attending breastfeeding promotion activities.

As far as possible, the independent evaluation data collection activities should be integrated with routine monitoring, such as by including the same data sources and indicator definitions. Nevertheless, because the data quality needs for a rigorous evaluation often exceed those required for routine monitoring, it may be necessary to establish a parallel data collection or data quality checks system. This structure should not be viewed as a replacement for routine monitoring activities, which must continue during and after the independent evaluation. It is also important to provide feedback to implementers when shortcomings in the program or the monitoring system are identified through the use of routine monitoring data. For example, if the monitoring system is not documenting key inputs such as stockouts of commodities at the health facility level, evaluators should discuss with implementers why these data are not being collected and explore how the data could be collected and used to improve program performance.

Similarly, the presence of an independent evaluation does not replace the need for focused internal evaluations aimed at fine-tuning the program. As far as possible, these two exercises should be coordinated so that they complement each other.

Data Analyses

Data analyses should be driven by the evaluation questions, and should make full use of the data collected on implementation strength, quality, and outcomes—complemented by data on costs and contextual factors. Integrating data from different sources in a comprehensive and logical plan of analysis is not an easy task, and its success requires careful forward planning from the time the evaluation is being designed. Data analysis plans should be developed alongside the program impact model, objectives, and measurement plan at the start of the evaluation.

In real-world evaluations, evaluators frequently need to change evaluation plans in response to changes made in the program, to account for new programs or interventions that are rolled out in comparison areas, or to address the unavailability of certain data. For example, as happened in the Malawi iCCM evaluation, programs similar to the one being evaluated may be rolled out in comparison districts by other partners, resulting in comparison areas that are no longer valid approximations of the counterfactual. In turn, evaluators may change a quasi-experimental evaluation design into a dose–response design, or may update the types of inferences that can be drawn from the quasi-experimental design. In any case, it is essential to document the initial design and analysis plan and any changes made to them, and to make note of possible biases resulting from these changes.

Issues in the analyses of nonrandomized evaluations are reviewed elsewhere (Cousens et al., 2009), but the following basic concepts should certainly drive the planned analyses:

- The units of analysis should be the same as the geographic units where the program is being implemented (e.g., districts).
- The sample size is the number of districts, not the total number of individuals surveyed in all study districts.
- Detailed analyses of baseline comparability are essential.
- Standard stratification and regression techniques will often be appropriate, but attention must be given to the number of degrees of freedom, which will often be small as a result of the number of districts available for analyses.
- Propensity scores may be useful for summarizing several confounders (Sturmer et al., 2006), particularly when the number of degrees of freedom is small.
- Comparisons between program and nonprogram areas, when baseline data exist, may rely on the difference-in-differences approach. This method focuses on the net change in each set of districts.
- Multilevel analyses may be appropriate if the data structure is hierarchical—for example, districts within the program area, health facilities within each district, providers within each facility, and patients seen by each provider.
- The plan of analysis should ideally examine all assumptions included in the conceptual model (i.e., all the arrows in the model). An example is provided in **TABLE 18-3**.
- The plan of analyses should be flexible, as real-world implementation may follows paths that are not planned ahead of time.

Table 18-3 shows possible research questions for evaluating a hypothetical MNCH program in a country in sub-Saharan Africa. The program includes several MNCH interventions, which will be initially rolled out in a limited number of districts. Although a stated objective of the program is to eventually scale up the program to all districts in the country, this step will depend on further partner and governmental buy-in and additional funding. This somewhat vague implementation plan is typical of many health programs in LMICs. A good example is the ACSD program (Exhibit 18-2), in which the full package was implemented in selected high-impact districts but a smaller set of interventions were scaled up nationally.

The analytical approaches suggested in Table 18-3 encompass the traditional program versus nonprogram districts comparison as well as dose–response analyses, such as those carried out in the context of an evaluation platform (Exhibit 18-6). Of the primary objectives, questions 1–4 may be addressed through a traditional comparison of two groups of districts (with or without the program). As pointed out earlier, such comparisons are often favored by funders, but may make little sense if the same or similar interventions are also present in the comparison areas. Answering questions 5–7 requires dose–response analyses, which go beyond the simple stratification of districts in two groups. If a stepped wedge design is used for the evaluation, these questions can be tailored to this type of analytical approach (Hussey & Hughes, 2007).

Question 8 is fundamental for both approaches: It requires control of confounding factors to address the issue of attribution of an observed impact to a given program. In a context in which multiple programs are being studied simultaneously through an evaluation platform approach, the presence of other programs may also be included as a confounding factor. For example, the Tanzania ITN voucher evaluation showed that the program's effects persisted even after taking into account the large number of free ITNs that were distributed in some districts in 2005 (Hanson et al., 2009).

Questions 9–12 represent secondary objectives of the hypothetical program evaluation. Question 9 assesses whether programs are being deployed where they are most needed. Although implementers often state that they are prioritizing the poorest or highest-mortality districts, studies have shown that this is not always true (Victora et al., 2006). Identification of indicators of baseline mortality, poverty levels, or

TABLE 18-3 Proposed Approaches for Data Analyses, According to Each Objective of a Hypothetical Evaluation of a Maternal, Newborn, and Child Health Program

Research Question/Objective	Analytical Approach
Primary Objectives	
1. Will the program-focus districts implement maternal, newborn, and child health (MNCH) interventions more strongly and at higher levels of quality than other districts in the country?	Comparison of implementation strength in program and non-program districts
2. Will the program-focus districts achieve greater increases in coverage of MNCH interventions than other districts in the country?	Comparison of changes in coverage over time in program and non-program districts
3. Will the program-focus districts achieve greater health impact (i.e., reductions in mortality and undernutrition) than other districts in the country?	Comparison of changes in impact indicators over time in program and non-program districts
4. How much did the program cost compared to other approaches, and how does their cost-effectiveness compare?	Comparison of costs and cost-effectiveness in program and non-program districts
5. Among all districts in the country, will those with stronger implementation of MNCH interventions achieve higher coverage levels than those with weaker implementation?	Dose–response analyses: Regression of changes in coverage over time on implementation score
6. Among all districts in the country, will those with stronger implementation of MNCH interventions achieve greater health impact (i.e., reductions in mortality and undernutrition) than those with weaker implementation?	Dose–response analyses: Regression of changes in impact indicators over time on implementation score
7. How much did the implementation cost, and is it cost effective?	Dose–response analyses relating implementation costs to impact
8. If there is evidence of an impact, can external explanations for these associations be ruled out?	Incorporation of relevant confounding factors in the preceding analyses
Secondary Objectives	
9. Was the program focused on the districts where it was most needed, in terms of baseline health and socioeconomic characteristics?	Stratification of all districts according to baseline characteristics, and comparison of strata in terms of implementation strength
10. Did implementation of the program contribute to reducing socioeconomic and other inequities in MNCH indicators?	Analyses of time trends in inequalities in coverage and impact indicators, according to strength of implementation of the program
11. Are there positive or negative effects of the program in terms of health systems, other health programs, or related indicators?	Comparison of health systems indicators and coverage of other interventions, according to strength of implementation of the program
12. Are changes in program implementation and in intervention coverage compatible with the observed impact on mortality and nutrition?	Comparison of actual changes in impact indicators with those predicted by modeling changes in intervention coverage (e.g., LiST)

health systems strength available in the district database (Exhibit 18-6) will help answer this question.

A related question is number 10, which assesses changes in health inequalities between population subgroups. Earlier in this chapter, we described the importance of collecting information on gender, socioeconomic position, urban versus rural residence, and ethnic group during surveys commissioned by the evaluation. Several approaches may be used to express the magnitude of inequality, and to describe how this level evolves over time in program versus non-program areas. In particular, concentration indices and slope indices of inequality have become increasingly popular means of examining these issues (O'Donnell et al., 2007; PAHO, 1999; Pamuk, 1985; Wagstaff, Paci, & van Doorslaer, 1991).

Question 11 relates to side effects of the program, which are often unintended and overlooked. Evaluators can address this question by collecting relevant information on the implementation and coverage of other health programs, such as through an evaluation platform approach.

Finally, question 12 relies on statistical modeling to check whether the observed impact on health status is compatible with the magnitude of changes in coverage of the interventions being promoted. In the case of maternal, newborn, and child mortality, the LiST software (Exhibit 18-8; Johns Hopkins Bloomberg School of Public Health, n.d.a) may be used to conduct these simulation exercises. Demonstration that the degree of mortality decline can be explained by changes in coverage is a strong argument that reinforces the impact model and increases the ability to attribute the health impact to the program interventions.

For evaluations that do not go as far as measuring mortality impact—which include most evaluations of maternal health and HIV/AIDS programs—use of LiST or similar simulation tools can provide evidence to funders and national policy makers that observed improvements in coverage are most likely contributing to reducing mortality rates.

Analyzing Costs and Cost-Effectiveness

A number of process, intermediate, and outcome indicators are available to be compared with cost; these indicators are summarized in **TABLE 18-4**. To the extent possible, all of these indicators should be used, because each provides useful information. However, the minimum set of recommended indicators includes the following items:

- Cost per beneficiary
- For curative services, the cost per person treated
- For preventive services, the cost per preventive item delivered
- Cost per death averted

Costs should be presented in the currency most relevant for the target audience. In most cases, this means that results should be presented in local currency units (for national policy makers), in U.S. dollars (for donors and program managers in other countries), and in purchasing power parity adjusted dollars (for comparison with other studies and for international policy makers). WHO website (http://www.who.int/choice/costs/ppp/en/index.html) offers more details about purchasing power parity adjusted dollars.

Once the incremental cost has been established, it can be compared with the incremental benefits to determine the *incremental cost-effectiveness ratio* (ICER). At the very least, cost-effectiveness should be calculated using deaths averted as the denominator. Other metrics (e.g., life-years saved, disability-adjusted life-years [DALYs] averted) can be applied if relevant for a particular country, although these metrics also imply different analytic horizons or time frames.

The *ceiling ratio* is the relative cost-effectiveness value against which the acceptability of ICERs is judged. If the value of an ICER is below the ceiling ratio, an intervention is deemed acceptable on grounds of cost-effectiveness. Previous recommendations, based on work by the Commission on Macroeconomics and Health (2001), classified interventions as "highly cost-effective" for a given country if they avert a DALY for less than the per capita national gross national income (GNI) or gross domestic product (GDP), and as "cost-effective" if this value is less than three times the per capita national GNI. The choice of ceiling ratios is affected by the decision maker's valuation of a unit of health gain; it is a particularly crucial and politically sensitive element of economic evaluation. Recently, the utility of the GDP per capita threshold has been questioned (Marseille, Larson, Kazi, Kahn, & Rosen, 2014), and the need for country-specific thresholds and information related to affordability, feasibility, and equity has been highlighted when making decisions about whether to adopt a program (Bertram et al., 2016).

We suggest placing findings in a broader context by comparing them to other economic evaluations that have been undertaken in the same or neighboring countries after adjustments have been made for inflation. Placing results in the context of a ceiling ratio provides a general picture of the program's value for money, but does not address the issue of whether the intervention represents the best

TABLE 18-4 Types of Process, Intermediate, and Outcome Indicators and Data Needed

Type of Indicator	Indicator	What the Indicator Measures	Additional Data Needed for the Indicator
Process	Ex ante cost-effectiveness	Expected costs and value for money	Budget projections, work plans, coverage targets
Process	Incremental/total cost per person treated	Services provided	Utilization rates
Process	Incremental/total cost per preventive item delivered (e.g., bed net or vitamin A capsule)	Services provided	Utilization rates
Process	Incremental/total cost per capita	Services provided, program effort, and sustainability	Population
Intermediate	Incremental/total cost per patient correctly treated	Treatment leading to health gains	Utilization rates adjusted by quality
Intermediate/outcome	[Incremental] cost of quality improvement	Improvements in services due to program	Quantitative measure of quality improvements
Outcome	Incremental cost per death averted	Mortality reduction	Mortality rates
Outcome	Incremental cost per life year gained	Mortality reduction	Mortality rates and age of death (and life expectancy)
Outcome	Incremental cost per disability-adjusted life-year (DALY) averted	Mortality and morbidity reduction	Mortality rates, age of death (and life expectancy), average morbidity associated with diseases, correct treatment by disease

Modified from Johns, B., Opuni-Akuamoa, M., Walker, D. (n.d.). Assessing the costs and cost-effectiveness of rapid-scale up for the maternal, neonatal, and child health: The economic component of the impact evaluation strategy for the Catalytic Initiative. Retrieved from http://www.jhsph.edu/research/centers-and-institutes/institute-for-international-programs/_documents/rapid_scaleup/CI_costeffectiveness.pdf. With permission of the author.

possible use of additional resources. Comparing the results to other studies will partly alleviate this problem. It also informs the decision maker about other options to the program besides "doing nothing," "the status quo," or what is happening in the comparison areas. Note, however, that the methods and assumptions in other studies may be different from those employed here or in comparison with each other, which makes these comparisons problematic. Results need to be interpreted in the full light of these shortcomings.

Interpretation and Attribution

Adequacy evaluations were discussed earlier in this chapter, in the design section. These assessments typically include a before-and-after comparison of outcomes within the program areas, without an external comparison. The conclusion may be that outcome indicators improved, remained unchanged, or worsened. Such evaluations may be useful for assessing changes in coverage and attributing them to a given program, but are rarely sufficient for attributing

impact. As a consequence, adequacy evaluations per se have a limited role in attribution.

Nevertheless, assessment of the adequacy of impact indicators—whether or not they improved over time—is essential, even in more complex plausibility evaluations that also include a comparison group. Because every plausibility evaluation provides the opportunity to assess adequacy, it is possible to interpret these findings jointly. **TABLE 18-5** shows how combining adequacy and plausibility may help the evaluators interpret the results and attribute them to a given program. (Table 18-5 is based on a program- versus non-program areas design, but can be adapted to dose–response and stepped wedge designs.) The plausibility findings are assumed to be not confounded—that is, they are expected to be unaffected by changes in contextual factors.

Our experience with large-scale evaluations is that implementers often concentrate on adequacy findings (e.g., did mortality rates fall in the program area?). Although this understanding is important, the result does not necessarily mean that the program can be held responsible for the observed success, as in some cases the comparison areas may have done just as well. An example is the ACSD evaluation (Bryce et al., 2010).

Table 18-5 also shows situations in which no improvement occurred in the program areas but non-program areas fared worse. This outcome would support a safety net effect in which the program helped offset the deterioration in external factors.

Finally, there is the possibility that the program displaced a more efficient program or strategy. This outcome may have occurred in the ACSD evaluation in Mali (Bryce et al., 2010), where documentation of program implementation and contextual factors showed that other funding agencies were supporting the scale-up of MNCH interventions similar to ACSD in most districts from the comparison area.

Ensuring the internal validity of an evaluation requires several steps that were described earlier in this chapter—namely, taking into account baseline comparability, documenting and adjusting for important confounding factors, ensuring valid measurement of implementation and outcomes, and using appropriate analytical strategies. An equally important consideration is to document changes in the intermediate steps laid out in the impact model, and to show that these outcomes are consistent with the impact findings.

Plausibility that the impact findings can (or cannot) be attributed to the intervention does not depend on a single finding or analysis, but rather on the progressive accumulation of evidence supporting the

TABLE 18-5 Joint Interpretation of Findings from Adequacy and Plausibility Analyses

How Did Program Areas Fare Relative to Non-Program Areas? (Plausibility Assessment)	How Did Impact Indicators Change Over Time in the Program Areas? (Adequacy Assessment)		
	Improved	**No Change**	**Worsened**
Better	Both areas improved, but program led to faster improvement	Program provided a safety net	Program provided a partial safety net
Same	Both areas improved to a similar extent; no evidence of an additional program impact	No change in either area; no evidence of program impact	Indicators worsened in both areas; no evidence of a safety net
Worse	Both areas improved, particularly non-program areas; presence of the program may have precluded the deployment of more effective strategies	Program precluded progress; presence of the program may have hindered the deployment of more effective strategies	Program was detrimental; presence of the program may have hindered the deployment of more effective strategies

hypothesis that the observed impact is due to the program. This type of buildup was noted in the Tanzania ITN voucher evaluation (Hanson et al., 2009).

Attribution is a central issue in evaluation and should be considered prior to deciding on an evaluation design. When multiple agencies or programs are active in the same country promoting similar interventions, as was the case in the Malawi iCCM evaluation, it is difficult, if not impossible, to attribute any observed improvements to a single partner. A more sensible approach in such situations is to adopt a "contributorship" model of attribution, in which all those who contribute to a program are acknowledged and successes or failures are shared among the partners. Doing a good job on the documentation component of the evaluation should make this type of joint attribution possible. If the number of districts is reasonably large (say, 30 or more), a dose–response design may be used to account for differential levels of implementation by different partners and the resulting variability in the effect size.

Evaluators are sometimes requested not only to provide evidence that overall impact can be attributed to a program, but also to assess which of several program components or interventions played the most important role. As mentioned in the description of the ACSD program, it is usually impossible to disentangle these different components, unless there are few interventions, the various interventions address different causes of morbidity and mortality, and little comorbidity exists between these causes. A similar issue arises when different funders support separate interventions that are packaged in the same program, in which case separate attribution of results is seldom possible. Funders and implementers must be made aware early on that these types of attribution are not feasible.

Disseminating Evaluation Findings and Promoting their Uptake

Evaluators have an ethical responsibility to make their findings available so that they can be used to guide program implementation and future evaluations. The provision of feedback to program implementers was addressed earlier in this chapter; the same principles apply to the final results addressing the evaluation questions. Dissemination activities should be planned and carried out with several audiences in mind:

- *Policy makers and program implementers at the country level*. These audiences should be carefully defined, and are likely to include district managers and even communities as well as national-level decision makers. Dissemination channels can take various forms, including face-to-face presentations and print materials. For example, in the Burkina Faso iCCM evaluation, two-page "district profiles," which used basic charts to visualize the evaluation data collected for that district and its performance relative to other districts, were developed for each district by the evaluation team.
- *Global scientific public health communities*. Evaluators have a responsibility to publish their findings in peer-reviewed scientific journals, and especially those likely to reach readers who can make use of the findings. Presentations at professional conferences usually reach a smaller audience but can also be useful. The latter forums offer an opportunity for younger members of the evaluation team to gain visibility and experience in presenting to colleagues.

Our experience indicates that generating interest in and uptake of the results of large-scale program evaluations is an ongoing process, and often requires several attempts using multiple channels. One strategy to enhance dissemination of the results is to ensure that those who can make use of the findings are involved from the start in defining the evaluation questions and are kept informed of the progress of the evaluation and its intermediate results.

Working in Large-Scale Evaluations

Conducting large-scale evaluations is not for the faint-hearted. This chapter has focused on the technical aspects of designing and conducting an evaluation, mentioning only in passing some of the political and personal challenges involved. We close the presentation here by highlighting a selection of these issues, all of which are crucial for designing an evaluation that can produce sound evidence that is widely used to improve public health practice and programs.

First, good evaluations require effective communication. **EXHIBIT 18-10** lists some common misperceptions about evaluation and evaluators from the perspective of those implementing or funding programs, and vice versa.

Second, good evaluations require a broad range of skills and techniques, as well as an interdisciplinary approach. Evaluators should have sufficient experience and a multifaceted "toolbox" to respond to the specific questions posed by stakeholders, rather than

EXHIBIT 18-10 Building Stronger Bridges Between Implementers and Evaluators

TABLE 18-6 lists some of the things implementers say about evaluators, and evaluators say about implementers. In a perfect world, these perceptions would be brought into alignment and both groups would recognize the value of sound evaluation as a contribution to strong programs. Making this synchrony happen is a challenge that must be faced every day, in every evaluation.

TABLE 18- 6 What Are They Saying?

Implementers About Evaluators and Evaluations	Evaluators About Program Implementers and Implementation
■ Not tied to the realities of the field; too academic ■ Interested in publishing rather than saving lives ■ Too slow in producing results ■ Cost too much; resources would be better spent in doing more implementation ■ Can bring only bad news ■ Too focused on impact, rather than on process (black box) ■ Insensitive to the policy process—produce results that are too late, too negative, insufficiently action oriented ■ We could do a better job ourselves, from within the program	■ Don't understand the importance of external, independent evaluations ■ Underestimate the need for data checking and time for analysis before sharing results ■ Need to recognize and act on intermediate results more rapidly ■ Want to "cherry pick" the findings to support their own program ■ Don't understand that evaluation is a technical discipline and requires special training to be done well ■ Need to do more internal monitoring and documentation of their own programs and to cover such costs from their budgets, rather than expect evaluators to do these tasks

just those questions they have been trained to answer. This goal is often best accomplished by establishing an evaluation team, as was done in all of the large-scale evaluations used as examples in this chapter.

Third, good evaluations require sufficient time and funding. Unfortunately, evaluation is often an afterthought, and setting aside funds for evaluation is perceived as "taking money away from beneficiaries." Underfunded and rushed evaluations often lead to poor-quality data, significant design limitations, and invalid results. In situations where sufficient funds truly cannot be set aside for an impact evaluation (say, at least 5% of program costs, and more if the evaluation needs to conduct its own household surveys), it may be preferable to focus instead on conducting a process evaluation, or using evaluation funds to conduct high-quality monitoring. However, if a measure of the program's effectiveness is required, it is essential to allocate adequate funds for the evaluation, and to begin evaluation planning and data collection well before the program is rolled out.

Fourth, good evaluations require patience and flexibility. Large-scale evaluations can move only as fast as implementation, and real-world programs are vulnerable to frequent delays, setbacks, and false starts. Timetables for program implementation often slip or are changed for reasons outside the control of the evaluation team. In-country research partners have many demands on their time, and capacity-building exercises are often needed as a precursor to full analysis of the data. Evaluation results may sit on a shelf until political circumstances or changes in leadership provide an opportunity for them to be incorporated into program plans. Ideal designs (based on the content in texts like this one) must often be modified to reflect what is possible and affordable in specific country contexts.

The authors of this chapter extend an invitation to readers to become engaged in the science of large-scale evaluations, either as evaluators themselves or as informed consumers and supporters of evaluation. Few formal training programs for public health evaluators exist, although the number has increased in recent years. Making a commitment to evaluation often requires seeking out and using resources from a variety of disciplines, and bringing them together into an effective mix of skills. Recent efforts to establish a community of practice—for example, the 3ie initiative—should be continued and expanded to support this emerging culture and to ensure that more and better evaluations of large-scale programs are supported in the future.

Acknowledgments

We thank Jennifer Bryce and Damian Walker for their substantive contributions to the earlier version of this chapter that was published in the third edition of this text.

Discussion Questions

1. Which arguments would you use to convince (1) a donor agency or (2) a ministry of health to allocate funds to conduct impact evaluations of its health programs?
2. What are the major challenges in conducting impact evaluations of health programs in LMICs?
3. Imagine that you are advising an NGO that wants to evaluate one of its health programs implemented in a LMIC. How would you go about determining whether the organization should use an adequacy, plausibility, or probability design for the evaluation? What are the tradeoffs for these designs in terms of feasibility, complexity, and level of inference?
4. How are data about costs obtained and used in program evaluations? Which kinds of questions can these data answer?

References

Adam, T., Bishai, D., Kahn, M., & Evans, D. (2004). *Methods for the costing component of the multi-country evaluation of IMCI.* Geneva, Switzerland: World Health Organization.

Adam, T., Ebener, S., Johns, B., & Evans, D. B. (2008). Capacity utilization and the cost of primary care visits: Implications for the costs of scaling up health interventions. *Cost Effectiveness and Resource Allocation*, *6*, 22.

Adam, T., Edwards, S. J., Amorim, D. G., Amaral, J., Victora, C. G., & Evans D. B. (2009). Cost implications of improving the quality of child care using integrated clinical algorithms: Evidence from Northeast Brazil. *Health Policy 89*, 97–106.

Adam, T., Manzi, F., Schellenberg, J. A., Mgalula, L., de Savigny, D., & Evans, D. B. (2005). Does the integrated management of childhood illness cost more than routine care? Results from the United Republic of Tanzania. *Bulletin of the World Health Organization*, *83*, 369–377.

Amorim, D. G., Adam, T., Amaral, J. J., Gouws, E., Bryce, J., & Victora, C. G. (2008). Integrated management of childhood illness: Efficiency of primary health in Northeast Brazil. *Revista de Saude Publica, 42*, 183–190.

Amouzou, A., Hazel, E., Shaw, B., Miller, N. P., Tafesse, M., Mekonnen, Y., & Black, R. E. (2016). Effects of the Integrated Community Case Management of Childhood Illness strategy on child mortality in Ethiopia: A cluster randomized trial. *American Journal of Tropical Medicine and Hygiene, 94*(3), 596–604.

Amouzou, A., Kanyuka, M., Hazel, E., Heidkamp, R., Marsh, A., Mleme, T., & Bryce, J. (2016). Independent evaluation of the Integrated Community Case Management of Childhood Illness strategy in Malawi using a national evaluation platform design. *American Journal of Tropical Medicine and Hygiene, 94*(3), 574–583.

Axelson, H., Bales, S., Minh, P. D., Ekman, B., & Gerdtham, U.-G. (2009). Health financing for the poor produces promising short-term effects on utilization and out-of-pocket expenditure: Evidence from Vietnam. *International Journal for Equity in Health, 8*(1), 20.

Baltussen, R., Adam, T., Tan-Torres, E. T., Hutubessy, R., Acharya, A., Evans, D. B., & Murray, C. J. L. (2003). Methods for generalized cost-effectiveness analysis. (Eds.) In *WHO guide to cost-effectiveness analysis*, *2*, 106. Geneva, Switzerland: World Health Organization.

Berman, P., Brotowasisto, Nadjib, M., Sakai, S., & Gani, A. (1989). The costs of public primary health care services in rural Indonesia. *Bulletin of the World Health Organization*, *67*, 685–694.

Bertram, M. Y., Lauer, J. A., De Joncheere, K., Edejer, T., Hutubessy, R., Kieny, M.P., & Hill, S. R. (2016). Cost-effectiveness thresholds: Pros and cons. *Bulletin of the World Health Organization, 94*(12), 925–930.

Bishai, D., Mirchandani, G., Pariyo, G., Burnham, G., & Black, R. (2008). The cost of quality improvements due to Integrated Management of Childhood Illness (IMCI) in Uganda. *Health Economics*, *17*, 5–19.

Black, N. (1996). Why we need observational studies to evaluate the effectiveness of health care. *British Medical Journal*, *312*(7040), 1215–1218.

Boschi-Pinto, C., Young, M., & Black, R. (2010). The Child Health Epidemiology Reference Group reviews of the effectiveness of interventions to reduce maternal, neonatal and child mortality. *International Journal of Epidemiology*, *39*(1 suppl), i3–i6.

Bratt, J. H., Foreit, J., Chen, P. L., West, C., Janowitz, B., & De Vargas, T. (1999). A comparison of four approaches for measuring clinician time use. *Health Policy and Planning*, *14*, 374–381.

Brown, C. A., & Lilford, R. J. (2006). The stepped wedge trial design: A systematic review. *BMC Medical Research Methodology*, *6*, 54.

Bryce, J., Friberg, K., Kraushaar, D., Nsona, H., Afenyadu, G., Nare, N., & Walker, N. (2010). LiST as a catalyst in program planning: Experiences from Burkina Faso, Ghana and Malawi. *International Journal of Epidemiology*, *39*(1 suppl), i40–i47.

Bryce, J., Gilroy, K., Jones, G., Hazel, E., Black, R. E., & Victora, C. G. (2010). The Accelerated Child Survival and Development programme in West Africa: A retrospective evaluation. *The Lancet*, *375*(9714), 572–582.

Bryce, J., Victora, C. G., Boerma, T., Peters, D. H., & Black, R. E. (2011). Evaluating the scale-up for maternal and child survival: A common framework. *International Health, 3*(3), 139–146.

Bryce, J., Victora, C. G., Habicht, J. P., Black, R. E., & Scherpbier, R. W. (2005). Programmatic pathways to child survival: Results

of a multi-country evaluation of Integrated Management of Childhood Illness. *Health Policy and Planning*, *20*(suppl 1), i5–i17.

Bryce, J., Victora, C. G., Habicht, J. P., Vaughan, J. P., & Black, R. E. (2004). The multi-country evaluation of the Integrated Management of Childhood Illness strategy: Lessons for the evaluation of public health interventions. *American Journal of Public Health*, *94*(3), 406–415.

Buddelmeyer, H., & Skoufias, E. (2004). *An evaluation of the performance of regression discontinuity design on Progresa* (Working Paper 3386). Washington, DC: World Bank Group.

Burgert, C. R., Colston, J., Roy, T., & Zachary, B. (2013). *Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys.* DHS Spatial Analysis Report No. 7. Calverton, MD: ICF International.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York, NY: John Wiley & Sons.

Commission on Macroeconomics and Health. (2001). *Macroeconomics and health: Investing in health for economic development.* Geneva, Switzerland: World Health Organization.

Cousens, S., Hargreaves, J., Bonell, C., Armstrong, B., Thomas, J., Kirkwood, B. R., & Hayes, R. (2009). Alternatives to randomisation in the evaluation of public-health interventions: Statistical analysis and causal inference. *Journal of Epidemiology and Community Health*, *6*, 6.

Cunnama, L., Sinanovic, E., Ramma, L., Foster, N., Berrie, L., Stevens, W., & Vassall, A.(2016). Using top-down and bottom-up costing approaches in LMICs: The case for using both to assess the incremental costs of new technologies at scale. *Health Economics, 25*(suppl 1), 53–66.

Donner, A., & Klar, N. (2000). *Design and analysis of cluster randomization trials in health research*. London, UK: Arnold.

Evaluation: The top priority for global health. (2011). *The Lancet*, *377*, 85–95.

Evaluation Gap Working Group. (2006). *When will we ever learn? Improving lives through impact evaluation*. Washington, DC: Center for Global Development. Retrieved from http://www.cgdev.org/content/publications/detail/7973

Every Woman Every Child. (2015a). The global strategy for women's, children's and adolescents' health (2016–2030): Survive, thrive, transform. Retrieved from http://www.who.int/life-course/partners/global-strategy/ewec-globalstrategyreport-200915.pdf?ua=1

Every Woman Every Child. (2015b). The unified accountability framework: Supporting country-led efforts with the Global strategy for women's children's and adolescents' health. Retrieved from http://www.who.int/pmnch/activities/accountability/framework.pdf?ua=1

Garnett, G. P., Cousens, S., Hallett, T. B., Steketee, R., & Walker, N. (2011). Mathematical models in the evaluation of health programmes. *The Lancet*, *378*(9790), 515–525.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2016). *Impact evaluation in practice* (2nd ed.). World Bank Publications.

Global Initiative to Strengthen Country Health Systems Surveillance (CHeSS). (2008). *Summary report of a technical meeting and action plan*. Bellagio, Italy: Bellagio Rockefeller Center.

Gouws, E., Bryce, J., Pariyo, G., Armstrong Schellenberg, J., Amaral, J., & Habicht, J. P. (2005). Measuring the quality of child health care at first-level facilities. *Social Science & Medicine*, *61*(3), 613–625.

Grieve, R., Nixon, R., Thompson, S. G., & Normand, C. (2005). Using multilevel models for assessing the variability of multinational resource use and cost data. *Health Economics*, *14*, 185–196.

Guinness, L., Kumaranayake, L., Rajaraman, B., Sankaranarayanan, G., Vannela, G., Raghupathi, P., & George, A. (2005). Does scale matter? The costs of HIV-prevention interventions for commercial sex workers in India. *Bulletin of the World Health Organization*, *83*, 747–755.

Habicht, J. P., Victora, C. G., & Vaughan, J. P. (1999). Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International Journal of Epidemiology*, *28*(1), 10–18.

Hanson, K., Marchant, T., Nathan, R., Mponda, H., Jones, C., Bruce, J., & Schellenberg, J. A. (2009). Household ownership and use of insecticide treated nets among target groups after implementation of a national voucher programme in the United Republic of Tanzania: Plausibility study using three annual cross sectional household surveys. *British Medical Journal*, *2*, 239.

Hanson, K., Nathan, R., Marchant, T., Mponda, H., Jones, C., Bruce, J., & Schellenberg, J. A. (2008). Vouchers for scaling up insecticide-treated nets in Tanzania: Methods for monitoring and evaluation of a national health system intervention. *BMC Public Health*, *8*, 205.

Hargreaves, J. R., Goodman, C., Davey, C., Willey, B. A., Avan, B. I., & Schellenberg, J. R. (2016). Measuring implementation strength: Lessons from the evaluation of public health strategies in low- and middle-income settings. *Health Policy and Planning, 31*(7), 860–867.

Hayes, R. J., & Moulton, L. H. (2009). *Cluster randomised trials*. Abingdon, UK: Taylor & Francis.

Hendriks, M. E., Kundu, P., Boers, A. C., Bolarinwa, O. A., Te Pas, M. J., Akande, T. M., & Swan Tan, S. (2014). Step-by-step guideline for disease-specific costing studies in low- and middle-income countries: A mixed methodology. *Global Health Action, 7*, 23573.

Hofer, T. P., Hayward, R. A., Greenfield, S., Wagner, E. H., Kaplan, S. H., & Manning, W. G. (1999). The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. *Journal of the American Medical Association*, *281*, 2098–2105.

Hollingsworth, B. (2003). Non-parametric and parametric applications measuring efficiency in health care. *Health Care Management Science*, *6*, 203–218.

Hussey, M. A., & Hughes, J. P. (2007). Design and analysis of stepped wedge cluster randomized trials. *Contemporary Clinical Trials*, *28*(2), 182–191.

Hussey, P. S., de Vries, H., Romley, J., Wang, M. C., Chen, S. S., Shekelle, P. G., & McGlynn, E. A. (2009). A systematic review of health care efficiency measures. *Health Services Research*, *44*, 784–805.

India Ministry of Health and Family Welfare. (2017). *National Family Health Survey–4 2015–2016: India fact sheet.* Mumbai, India: Author. Retrieved from http://rchiips.org/NFHS/pdf/NFHS4/India.pdf

International Institute for Population Sciences (IIPS) and Macro International. (2007). *National Family Health Survey (NFHS-3), 2005–06: India: Volume I*. Mumbai: IIPS. Retrieved from https://dhsprogram.com/pubs/pdf/frind3/frind3-vol1andvol2.pdf

Jacobs, R., Smith, P., & Street, A. (2006). *Measuring efficiency in health care: Analytic techniques and health policy.* Cambridge, UK: Cambridge University Press.

Jalan, J., & Ravallion, M. (2003). Estimating the benefit incidence of an antipoverty program by propensity-score matching. *Journal of Business & Economic Statistics, 21*(1), 19–30.

Johns, B., & Torres, T. T. (2005). Costs of scaling up health interventions: A systematic review. *Health Policy and Planning*, *20*, 1–13.

Johns Hopkins Bloomberg School of Public Health. (n.d.a). LiST: An evidence-based decision-making tool for estimating intervention impact. Retrieved from http://www.livessavedtool.org/

Johns Hopkins Bloomberg School of Public Health. (n.d.b). National Evaluation Platform. Retrieved from http://www.jhsph.edu/research/centers-and-institutes/institute-for-international-programs/current-projects/national-evaluation-platform/

King, G., Gakidou, E., Imai, K., Lakin, J., Moore, R. T., Nall, C., & Llamas, H. H. (2009). Public policy for the poor? A randomised assessment of the Mexican universal health insurance programme. *The Lancet*, *373*(9673), 1447–1454.

Kleinbaum, D. G., Kupper, L. L., & Morgenstern, H. (1982). *Epidemiologic research: Principles and quantitative methods*. Hoboken, NJ: John Wiley & Sons.

Kumaranayake, L. (2008). The economics of scaling up: Cost estimation for HIV/AIDS interventions. *AIDS*, *22*(1 suppl), S23–S33.

Mangham, L. J., & Hanson, K. (2010). Scaling up in international health: What are the key issues? *Health Policy and Planning*, *25*(2), 85–96.

Marseille, E., Dandona, L., Saba, J., McConnel, C., Rollins, B., Gaist, P., & Kahn, J. G. (2004). Assessing the efficiency of HIV prevention around the world: Methods of the PANCEA project. *Health Services Research*, *39*, 1993–2012.

Marseille, E., Giganti, M. J., Mwango, A., Chisembele-Taylor, A., Mulenga, L., Over, M., & Stringer, J. S. A. (2012). Taking ART to scale: Determinants of the cost and cost-effectiveness of antiretroviral therapy in 45 clinical sites in Zambia. *PLoS One, 7*(12), e51993.

Marseille, E., Larson, B., Kazi, D. S., Kahn, J. G., & Rosen, S. (2014). Thresholds for the cost-effectiveness of interventions: Alternative approaches. *Bulletin of the World Health Organization, 93*(2), 118–124.

Masiye, F., Kirigia, J. M., Emrouznejad, A., Sambo, L. G., Mounkaila, A., Chimfwembe, D., & Okello, D. (2006). Efficient management of health centres human resources in Zambia. *Journal of Medical Systems*, *30*, 473–481.

Mensah, J., Oppong, J. R., & Schmidt, C. M. (2010). Ghana's national health insurance scheme in the context of the health MDGs: An empirical evaluation using propensity score matching. *Health Economics, 19*(suppl), 95–106.

Menzies, N. A., Berruti, A. A., & Blandford, J. M. (2012). The determinants of HIV treatment costs in resource limited settings. *PLoS One, 7*(11), e48726.

Miller, N. P., Amouzou, A., Tafesse, M., Hazel, E., Legesse, H., Degefie, T., & Bryce, J. (2014). Integrated Community Case Management of Childhood Illness in Ethiopia: Implementation strength and quality of care. *American Journal of Tropical Medicine and Hygiene, 91*(2), 424–434.

Mulligan, J. A., Yukich, J., & Hanson, K. (2008). Costs and effects of the Tanzanian national voucher scheme for insecticide-treated nets. *Malaria Journal*, *7*, 32.

Munos, M., Guiella, G., Roberton, T., Maïga, A., Tiendrebeogo, A., Tam, Y., & Baya, B. (2016). Independent evaluation of the rapid scale-up program to reduce under-five mortality in Burkina Faso. *American Journal of Tropical Medicine and Hygiene, 94*(3), 584–595.

National Statistical Office. (2015). *Malawi MDG Endline Survey 2014.* Retrieved from http://www.nsomalawi.mw/images/stories/data_on_line/demography/MDG%20Endline/MES%202014%20Report.pdf

O'Donnell, O., Van Doorslaer, E., Wagstaff, A., & Lindelow, M. (2007). *Analyzing health equity using household survey data: A guide to techniques and their implementation*. Washington, DC: World Bank Publications.

Over, M. (1986). The effect of scale on cost projections for a primary health care program in a developing country. *Social Science & Medicine, 22*(3), 351–360.

Oxman, A. D., Bjorndal, A., Becerra-Posada, F., Gibson, M., Block, M. A., Haines, A., & Wibulpolprasert, S. (2010). A framework for mandatory impact evaluation to ensure well informed public policy decisions. *The Lancet*, *375*(9712), 427–431.

Pamuk, E. R. (1985). Social class inequalities in mortality from 1921 to 1972 in England and Wales. *Population Studies*, *39*, 17–31.

Pan American Health Organization (PAHO). (1996). Impact of Uruguay's introduction of the *Haemophilus influenzae* type b (Hib) vaccine. *EPI Newsletter*, *18*, 6.

Pan American Health Organization (PAHO). (1999). Health analysis: Risks of dying and income inequalities. *Epidemiology Bulletin*, *20*(4), 7–10.

Powell, E. C., & Hampers, L. C. (2003). Physician variation in test ordering in the management of gastroenteritis in children. *Archives of Pediatric and Adolescent Medicine, 157*(10), 978–983.

Powell-Jackson, T., Hanson, K., Whitty, C. J. M., & Ansah, E. K. (2014). Who benefits from free healthcare? Evidence from a randomized experiment in Ghana. *Journal of Development Economics, 107*, 305–319.

Rao, J. K. N. (2003). *Small area estimation*. Hoboken, NJ: Wiley Interscience.

Rasella, D., Aquino, R., Santos, C. A., Paes-Sousa, R., & Barreto, M. L. (2013). Effect of a conditional cash transfer programme on childhood mortality: A nationwide analysis of Brazilian municipalities. *The Lancet, 382*(9886), 57–64.

Rivera, J. A., Sotres-Alvarez, D., Habicht, J. P., Shamah, T., & Villalpando, S. (2004). Impact of the Mexican program for education, health, and nutrition (Progresa) on rates of growth and anemia in infants and young children: A randomized effectiveness study. *Journal of the American Medical Association*, *291*(21), 2563–2570.

Santos, L. M., Paes-Sousa, R., Silva, J. B. Jr., & Victora, C. G. (2008). National Immunization Day: A strategy to monitor health and nutrition indicators. *Bulletin of the World Health Organization*, *86*(6), 474–479.

Sood, N., Bendavid, E., Mukherji, A., Wagner, Z., Nagpal, S., & Mullen, P. (2014). Government health insurance for people below poverty line in India: Quasi-experimental evaluation of insurance and health outcomes. *British Medical Journal, 349*, g5114.

Stover, J., McKinnon, R., & Winfrey, B. (2010). Spectrum: A model platform for linking maternal and child survival interventions with AIDS, family planning and demographic projections. *International Journal of Epidemiology*, *39*(1 suppl), i7–i10.

Sturmer, T., Joshi, M., Glynn, R. J., Avorn, J., Rothman, K. J., & Schneeweiss, S. (2006). A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates

compared with conventional multivariable methods. *Journal of Clinical Epidemiology*, *59*(5), 437–447.

Tan, S. S., Rutten, F. F., Van Ineveld, B. M., Redekop, W. K., & Hakkaart-Van Roijen, L. (2009). Comparing methodologies for the cost estimation of hospital services. *European Journal of Health Economics, 10*(1), 39–45.

The DHS Program. (2010). Malawi: Standard DHS, 2010. Retrieved from https://dhsprogram.com/what-we-do/survey/survey-display-333.cfm

The DHS Program. (n.d.a). AIDS indicators surveys: AIS overview. Retrieved from https://dhsprogram.com/What-We-Do/Survey-Types/AIS.cfm

The DHS Program. (n.d.b). Demographic and health surveys: DHS overview. Retrieved from https://dhsprogram.com/What-We-Do/Survey-Types/DHS.cfm

The DHS Program. (n.d.c). Malaria indicator surveys. Retrieved from https://dhsprogram.com/What-We-Do/Survey-Types/MIS.cfm

Tucker, A. M., Weiner, J. P., Honigfeld, S., & Parton, R. A. (1996). Profiling primary care physician resource use: Examining the application of case mix adjustment. *Journal of Ambulatory Care Management*, *19*, 60–80.

UNICEF. (2005). *Accelerating child survival and development: A results-based approach in high under-five mortality areas: Final report to CIDA*. New York, NY: Author.

UNICEF. (n.d.). Childinfo: Multiple Indicator Cluster Surveys (MICS). Retrieved from http://www.childinfo.org/mics.html

UNICEF Canada. (n.d.). The Catalytic Initiative to Save a Million Lives. Retrieved from https://www.unicef.ca/sites/default/files/legacy/imce_uploads/unicef_and_canada_lr_0.pdf

USAID. (2016, March 3). Child Survival and Health Grants Program. Retrieved from https://www.mcsprogram.org/our-work/community-health-civil-society-engagement/child-survival-health-grants-program/

Valdmanis, V., Walker, D., & Fox-Rushby, J. (2003). Are vaccination sites in Bangladesh scale efficient? *International Journal of Technology Assessment in Health Care*, *19*, 692–697.

Victora, C. G., Black, R. E., & Bryce, J. (2009). Evaluating child survival programmes. *Bulletin of the World Health Organization*, *87*(2), 83.

Victora, C. G., Habicht, J. P., & Bryce, J. (2004). Evidence-based public health: Moving beyond randomized trials. *American Journal of Public Health*, *94*(3), 400–405.

Victora, C. G., Huicho, L., Amaral, J. J., Armstrong-Schellenberg, J., Manzi, F., Mason, E., & Scherpbier, R. (2006). Are health interventions implemented where they are most needed? District uptake of the integrated management of childhood illness strategy in Brazil, Peru and the United Republic of Tanzania. *Bulletin of the World Health Organization*, *84*(10), 792–801.

Victora, C. G., Schellenberg, J. A., Huicho, L., Amaral, J., El Arifeen, S., Pariyo, G., & Habicht, J. P. (2005). Context matters: Interpreting impact findings in child survival evaluations. *Health Policy Plan 2005, 20,* i18–i31.

Wagstaff, A., Paci, P., & van Doorslaer, E. (1991). On the measurement of inequalities in health. *Social Science & Medicine*, *33*(5), 545–557.

Walker, N., Tam, Y., & Friberg, I. K. (2013). Overview of the Lives Saved Tool (LiST). *BMC Public Health, 13*(3), S1.

West, B. T., Berglund, P., & Heeringa, S. G. (2008). A closer examination of subpopulation analysis of complex-sample survey data. *Stata Journal*, *8*(4), 520–531.

West, S. G., Duan, N., Pequegnat, W., Gaist, P., Des Jarlais, D. C., Holtgrave, D., & Mullen, P. D. (2008). Alternatives to the randomized controlled trial. *American Journal of Public Health, 98*(8), 1359–1366.

Willey, B. A., Armstrong Schellenberg, J. R., Maokola, W., Shirima, K., Chemba, M., Mshinda, H., & Schellenberg, D. (2011). Evaluating the effectiveness of IPTi on malaria using routine health information from sentinel health centres in southern Tanzania. *Malaria Journal, 10,* 41.

Wilson, E. C., Mugford, M., Barton, G., & Shepstone, L. (2016). Efficient research design: Using value-of-information analysis to estimate the optimal mix of top-down and bottom-up costing approaches in an economic evaluation alongside a clinical trial. *Medical Decision Making, 36*(3), 335–348.

World Health Organization (WHO). (2014). Accountability for women's and children's health 2014 progress report. Retrieved from http://www.everywomaneverychild.org/wp-content/uploads/2016/11/COIA_report_2014.pdf

World Health Organization (WHO) & UNICEF. (2012). WHO/UNICEF joint statement: Integrated Community Case Management (iCCM): An equity-focused strategy to improve access to essential treatment services for children. Retrieved from https://www.unicef.org/health/files/iCCM_Joint_Statement_2012.pdf

CHAPTER 19

Health and "The Economy"

Richard Smith, Jennifer Prah Ruger, Dean T. Jamison, David E. Bloom, and David Canning

Introduction

At the most basic level, being poor is bad for one's health and being ill is bad for one's wealth. Of course, beyond that simple statement, the discussion becomes more complicated. Clearly, though, health and economics have a close, symbiotic relationship. That the health of an individual determines his or her productivity and ability to work, and that having more resources enables better nutrition and better health care, are well known. Although this dynamic is often examined at the individual level, it is less often investigated at the macroeconomic level. However, at the broadest level, economic growth is clearly associated with changes in disease profiles. This consideration is especially critical for countries undergoing major economic growth. Notably, transitioning from being a low- and middle-income country (LMIC) to being a high-income country is associated with a lag in health impacts, such that the country experiences a period of time with high levels of both communicable *and* noncommunicable diseases—a phenomenon known as the "double burden" (Boutayeb, 2006).

Macroeconomic policies concerning public expenditures targets, inflation control, tax policy, and exchange rates, among others, affect the provision of health care and a population's health status. For instance, national income and fiscal targets constrain how much a government can spend on health care; the exchange rate affects the costs of vaccines and drugs; and tax policies relating to tobacco, alcohol, and "fast food" influence people's demand for these products and ultimately their health. The health sector itself is also a major economic sector—in terms of employment, for instance—and is linked to other economic sectors via the consumption of other goods and services (e.g., electricity, construction, and pharmaceuticals). It is also integrated with the global market for goods, services, and people, which is subject to considerable international trade (see the *International Trade and Health* chapter).

Whether the final objective of development policy is economic growth, poverty alleviation, or improved quality of life, including health, these domains are inextricably linked. Given these interactions, it is critical that health professionals are equipped to understand the association between population health and the economy if they are to devise health and development policies that will improve people's quality of life worldwide. This chapter provides an overview of these interactions and relationships. It outlines the critical features, and terminology, of the macroeconomic perspective, and considers the relationships through the lenses of the household (looking especially at the interaction of economic growth and health), risk factors for disease (communicable, noncommunicable, and outbreaks), and the economy and its relationship with the healthcare sector.

A Macroeconomic Focus

Economics can be broadly split into microeconomics and macroeconomic. Microeconomics essentially concerns choices and activities at the individual or firm level. It considers which goods firms decide to produce and which goods households decide to consume. Households and firms interact within a market, where price movements seek to equate demand and supply. Typically, these markets combine to form "sectors," such as agriculture, manufacturing, or health care. In turn, the interactions of these sectors constitute "the economy." Macroeconomics then concerns choices and activities across several of these markets and sectors and, therefore, "the economy" as a whole.

Generally, health economic analysis addresses microeconomics, which focuses on the primary, secondary, and tertiary care services specifically designed to affect health status. The *Design of Health Systems* chapter, which focuses on the "economic costs" of specific diseases, takes this approach. Because the current chapter takes a much broader approach, it uses some specific terms that require definition (**EXHIBIT 19-1**).

The most central term is "gross national income" (GNI), which is used internationally to classify countries into stages of development. Each year, the World Bank revises the per capita GNI benchmarks that are used to classify countries. **TABLE 19-1** displays the World Bank's 2016 classification of countries, based on annual 2015 data (http://blogs.worldbank.org/opendata/new-country-classifications-2016).

Growth in GNI over time is termed "economic growth." This growth causes countries to transition between these classifications and is the core concept of economic development. GDP and GNI are also used to measure trends in economic activity. A *recession* is generally defined as a downturn in economic activity, with real GDP falling in two or more successive quarters (i.e., 3-month periods). Generally, economies face periods of economic growth and recession; these ups and downs are known as the *business cycle*. A *depression* is an economic downturn that is more severe than a recession, usually defined as a period when real GDP has decreased by more than 10%.

Economic performance and growth, expressed through GNI, is a major political goal and a key determinant, and result, of changes in population health.

EXHIBIT 19-1 Key Macroeconomic Terms

Appreciate: When a currency rises relative to other currencies, it is appreciating in value.
Balance of payments (BOP): A measure of the currency flows between countries.
Constant dollars: Constant dollars, or currency, correspond to values that have been adjusted for inflation and so reflect the "real" or actual purchasing power.
Current dollars: Current dollars, or currency, refer to the actual dollars spent.
Depreciate: When a currency falls relative to other currencies, it is depreciating in value.
Depression: A severe downturn in economic growth—often judged as a 10% decrease in real gross domestic product (GDP).
Devaluation: When a country's currency value declines in relation to other currencies, it is devaluating.
Fiscal policy: Policies introduced by the government to influence the economy through taxes and government spending.
Gross domestic income (GDI): The income from all economic activities that take place within a country, including wages, profits, rents, and interest.
Gross domestic product (GDP): The size or output of an economy. GDP comprises the total value of goods and services produced in one year in a country. It concerns the output produced in a specific geographic location, regardless of who produces it.
Gross national income (GNI): The income from the economic activities that citizens and firms of that country undertake, regardless of where they occur. GNI is GDP plus income that the country's citizens earn from abroad, minus income that foreign citizens earn in that country.
Hyperinflation: Inflation that exceeds 50% per month.
Inflation: The general rise in prices through time, which means that money loses its value over time.
Macroeconomics: The performance and functioning of the economy as a whole; the relationships among economic growth, output, employment, and inflation.
Monetary policy: Policies the government uses to reach its objectives by adjusting interest rates and the amount of money in circulation.
Purchasing power parity (PPP): An adjustment of prices that accounts for the currency's ability to purchase goods.
Recession: A downturn in economic activity wherein real GDP falls for two successive quarters (i.e., 3-month periods).

TABLE 19-1 World Bank Country Classification, 2015

World Bank Country Classification	2015 GNI per Capita	Number of Countries
Low-income countries	$1,025 or less	53
Lower-middle-income countries	$1,026–$4,035	55
Upper-middle-income countries	$4,036–$12,475	41
High-income countries	$12,476 or more	60

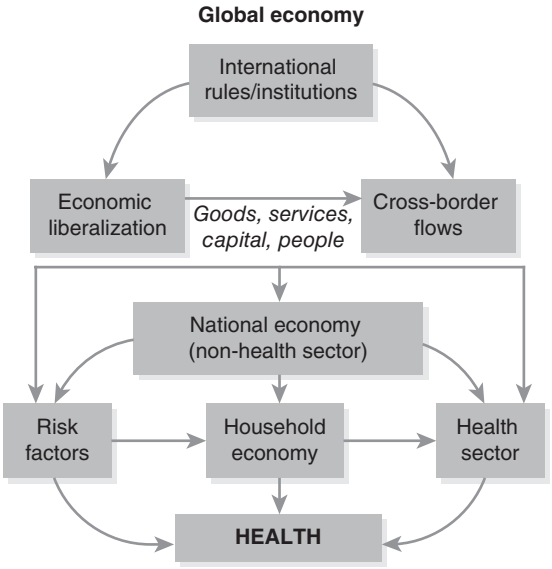


FIGURE 19-1 A simplified view of the relationship of the economy to health.

Modified from Smith, R.D., & Hanson K. (2012). *Health Systems in low-middle-income countries: An economic and policy perspective*. Oxford, UK:Oxford University Press.

FIGURE 19-1 illustrates the core mechanisms by which the economy relates to health. The lower half of the figure represents the national economy of interest, and the upper half represents the rest of the world and, critically, the trading relationships. In the lower half—the national economy—health looms large as the core concern for those reading this text. Surrounding it are four categories of factors that influence health status: (1) risk factors, representing genetic predisposition to disease, environmental influences on health, infectious disease, and other factors; (2) the household economy, representing factors associated with how individual and household behaviors and decisions impact health; (3) the health sector, representing the impact of goods and services consumed principally to improve health status; and (4) the non-healthcare sectors of the national economy, representing the meta-influences of non-health sectors and factors on these other three categories. How these factors relate is also illustrated in Figure 19-1. Those concerned with the health sector focus on one small part of this jigsaw puzzle—the links between the health sector and health.

The upper half of Figure 19-1 illustrates the influences of factors outside the national context—that is, the "rest of the world"—and highlights the main links to the domestic context. For example, various international influences directly affect *risk factors* for health, such as cross-border flows that increase exposure to infectious diseases through rapid cross-border transmission of communicable diseases. International factors also affect health by influencing the *national economy*. Economic liberalization—reducing barriers to make trade easier—influences economic growth and the distribution of positive and negative aspects of this growth, which in turn influence health (e.g., causing changes in consumption of various goods that are either beneficial or detrimental to health). Finally, trade directly affects the *health sector* by providing and distributing health-related goods, services, and people, such as pharmaceutical products, health-related knowledge and technology (e.g., new genomic

developments), and the movement of patients and professionals. The *International Trade and Health* chapter considers these aspects in more detail.

Also of note in the upper half of Figure 19-1 is the importance of international structures for governance, both of health and other aspects of life, which the *Global Health Governance and Diplomacy* chapter discusses. Here, we focus on the main impacts and relationships depicted in the lower half of Figure 19-1, using the structure of the three broad areas of the household economy, risk factors, and the health sector.

The Household as the Nexus of the Health–Wealth Relationship

Much macroeconomic policy concerns economic growth, as higher GNI leads to more opportunities to consume—which, in our specific context, will result, all else being equal, in better health. Central to this link between macroeconomic policy and health is the household. Health affects household income through its effects on labor productivity, savings rates, investments in physical and human capital (education, for example), and age structure. The other direction of causality derives from the impact that household income has on health through improved capacity to purchase food; adequate sanitation, housing, and education; and incentives to limit fertility.

In general, analyses suggest that "wealthier countries are healthier countries" (Pritchett & Summers, 1996). Although the direction of causation is debatable, a "virtuous circle" is engendered between increasing wealth and health such that both directions of causation are valid (Suhrcke, McKee, Arce, Tsolova, & Mortensen, 2006). This section looks at these links between health and wealth at the macroeconomic level, through the lens of the household, beginning with the role of health on wealth, and then turning to the role of wealth on health.

The Role of Health in Economic Growth

A considerable body of literature addresses the role of health in economic growth, both in general (Acemoglu & Johnson, 2007) and related to specific diseases, especially human immunodeficiency virus (HIV) (Lovász & Schipp, 2009; McDonald & Roberts, 2006) and malaria (Gallup & Sachs, 2001). One of the main empirical tools used is cross-country regression of economic growth (usually measured in terms of the growth rate of per capita GDP) on the variables believed to explain that growth, which include some proxy for "health," such as life expectancy, mortality rates, or disease prevalence (Bloom & Canning, 2000; Solow, 1956).

The evidence suggests that health is closely linked with economic growth and development. **FIGURE 19-2** illustrates the multiple pathways through which illness can influence economic development. The top half of the figure shows the age-structure effects of demographic transition as seen by a change in the dependency ratio, which has been a significant determinant of growth in East Asia, for example. High levels of fertility and child mortality (both, in part, a result of child illness), along with reductions in the labor force brought on by mortality and early retirement, can cause an increase in the dependency ratio that ultimately reduces income. Conversely, reversing these effects can decrease the dependency ratio, which increases per capita income. In addition, not only does the health of the mother before and during pregnancy affect the child's health (Kinra et al., 2008), but childhood health also affects adult health and productivity (Fogel, 2005; Park, Falconer, Viner, & Kinra, 2012). The lower half of Figure 19-2 illustrates the effects of illness and malnutrition in reducing labor productivity, operating through other factors such as reduced investments in human and physical capital. Finally, reduced labor productivity has a direct impact on reducing per capita income.

Life expectancy, in particular, has been shown to be a powerful predictor of household income levels and consequent economic growth: Lower levels of mortality and higher levels of life expectancy have a statistically and quantitatively significant effect on income levels and growth rates (Bloom, Canning, & Sevilla, 2004). When Jamison and colleagues (2005) combined their empirical estimates of the effects of adult survival rates on national income with country-specific estimates of improvements in survival rates across 53 countries, they found that, on average, 11% of total growth rate in per capita income is due to health improvements, although substantial variation across countries exists. Historical studies have reached similar conclusions using different types of data. For example, in a study of the determinants of economic progress in Great Britain between 1780 and 1979, Fogel (1997) estimated that 30% of per capita growth could be explained by health and nutritional improvements. Bloom, Canning, and Jamison (2004) reviewed the results of regression analyses that use life expectancy or other measures of mortality rates to explain economic growth. Their report indicates that findings in the literature of a large and significant impact of

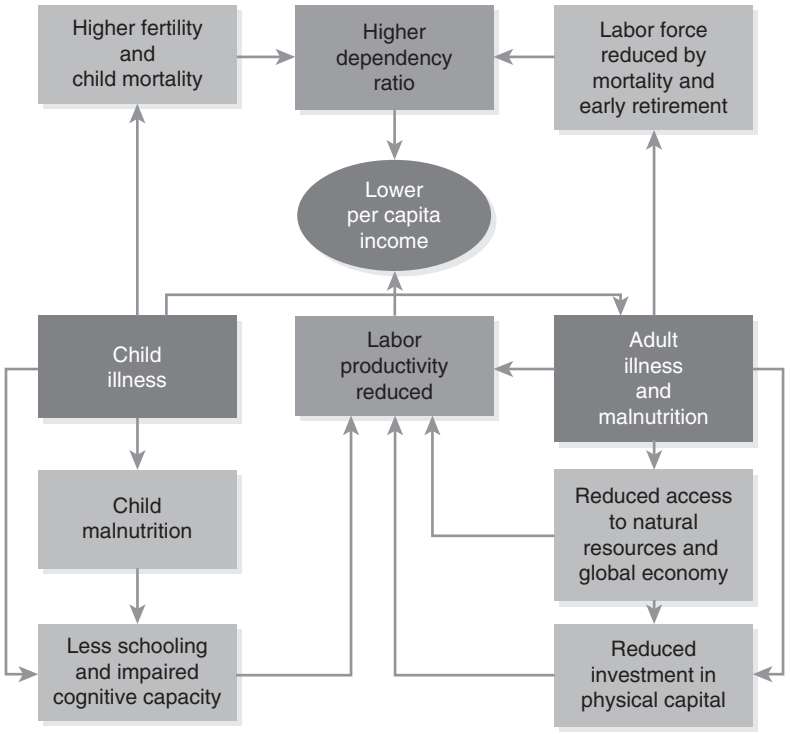


FIGURE 19-2 Channels through which illness reduces income.

Adapted and reprinted with permission from World Health Organization (WHO). *(1999). World health report 1999: Making a difference.* Geneva, Switzerland: World Health Organization.

initial health status on subsequent economic growth rates are robust.

Although solid evidence supports a large effect of health on economic growth, the issue of how to interpret this result remains a thorny one. Bhargava and colleagues (2001) concluded that the mechanisms underlying these effects are threefold: (1) the improvements in productivity that arise from a healthier workforce and less morbidity-related absenteeism, (2) the increased incentive that higher life expectancy gives individuals and firms to invest in physical and human capital, and (3) the increase in savings rates as working-age individuals save for their retirement years.

Productivity may be especially sensitive to health effects in low-income settings, as much of the work in these environments requires physical strength and endurance. For example, in one study, Indonesian men with anemia were found to have 20% lower productivity than men without it (Shastry & Weil, 2002). In an experiment, men were randomly assigned to receive either an iron supplement or placebo. Anemic men who received the iron treatment increased their productivity to nearly the levels of non-anemic workers; these productivity gains were relatively large compared to the costs of treatment. Further, the profile of disease in low-income settings, especially the predominance of infectious diseases, affects the full age range, whereas in high-income countries the predominance of non-infectious disease means that those most affected are at older ages. Thus, the effects of health on economic growth rates are likely to be strongest at the low levels of income. It has further been suggested that differences in health, physical capital, and education are roughly equal in their importance in producing cross-country differences in income levels, supporting the view that health has a large macroeconomic effect (Shastry & Weil, 2002).

The other potential explanations for life expectancy's importance relative to economic growth focus on the incentives to invest in education and to save for retirement. For instance, Bloom and colleagues (2003) have reported empirical evidence that life expectancy is linked to national savings rates. However, the incentive to invest more—both individually and nationally—in education as life expectancy increases is of particular significance because of the close links between health and education, and education and economic growth. Health and nutritional problems can have significant consequences for the educational success of many school-age children (also see the *Nutrition* chapter), which affects lifetime productivity and earnings, and therefore, national economic growth. A panel of experts taking part in the Copenhagen Consensus of 2008 (https://www.scribd.com/document/78190555/Copenhagen-Consensus-2008-Results) ranked a wide range of potential interventions in terms of their ability to advance global welfare, and particularly the welfare of LMICs Recognizing that limited funding

is available for such interventions, they examined the evidence base to determine the most cost-effective manner in which to spend development money. Six of the top 10 interventions involved disease control or nutrition, and their top-ranked intervention was micronutrient supplements for children.

The weight of the research evidence suggests, then, that improved health increases household income and, in turn, national GNI per capita, although the precise nature of this causation can be debated. The effect can also be modulated through effects on demography. Acemoglu and Johnson (2007) have argued that, even though health improvements may increase individual productivity, the reduced mortality rates will subsequently lead to a surge in population numbers that can depress income levels through a Malthusian effect of crowding out scarce resources such as land. This effect can be avoided, however, if fertility declines. Health advances in most high-income countries are associated with rapid declines in infant and child mortality rates. These declines, and increases in the proportion of surviving children, can substantially reduce desired fertility (Kalemli-Ozcan, 2002).

For example, in the 1940s, health improvements that led to changes in mortality and fertility rates in East Asia provided the impetus for a demographic transition. In the first phase of the demographic transition, an initial decline in infant and child mortality prompted a subsequent decline in fertility rates. These changes in mortality and fertility altered Asia's age distribution: The working-age population began growing much more rapidly than the youth dependent population. This resulted in a disproportionately high percentage of working-age adults, and thus a lower dependency-ratio, and led to increased rates of economic growth (Bloom, Canning, & Sevilla, 2002). Note, however, that while this "demographic dividend" provides an opportunity for increasing wealth, it does not guarantee such results. East Asia's growth rates were achieved because both the government and private sectors mobilized the growing workforce, successfully managed economic opportunities, adopted new industrial technologies, invested in education and human capital, and exploited global markets. These factors helped East Asia realize the economic growth potential created by the demographic transition.

The Role of Economic Growth in Health

We have seen how health can contribute to greater economic growth and reduced poverty. Here we consider the converse relationship—that is, how greater economic development can influence health. An important question in assessing the historical determinants of health is how increasing income, in relation to other factors, influenced health improvements, especially during the 20th century. Preston's (1980, 2007) seminal work provides a broad framework for assessing income's impact, and a classic study from the World Health Organization (WHO, 1999) provides a clear examination of these relationships.

FIGURE 19-3 shows that from 1952 to 1992, mean income on a global scale increased by approximately 66%, from roughly $1,500 to $2,500. The curves further indicate that the infant mortality rate (IMR) declined more than predicted (to 55 deaths per 1,000 births rather than to 116 deaths per 1,000 births) based on the 1952 relationship between income and IMR, suggesting that factors more important than income influence health.

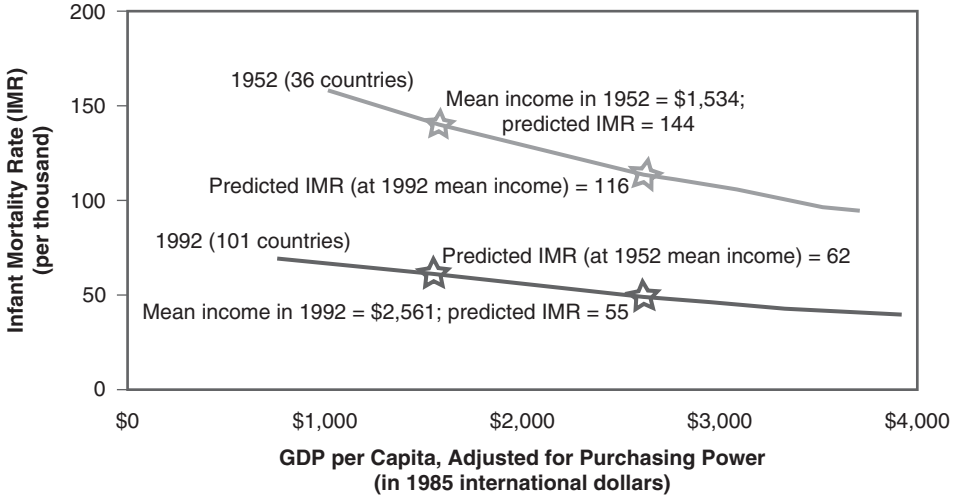


FIGURE 19-3 The role of improvements in income in reducing infant mortality rates.

At its heart, the positive impact of increased income on health is due to a variety of intervening determinants becoming more affordable (other than being directly made available through health care, which is considered later), such as housing, education, water, sanitation, and nutrition. There are also links from policies in other areas. Industrial safety programs and environmental regulations, for example, have specific health objectives and are targeted at specific groups—industrial workers and the general public, respectively. Housing policies aimed at subsidizing or providing housing for lower-income individuals have broader social and health objectives, but may also have concrete health impacts. Empowering women through better education, among other interventions, can enhance their capacity to improve their own health and that of their families (Thompson, 2007). The *Understanding and Acting on Social Determinants of Health and Health Equity* chapter provides an extensive overview of the determinants of health that are outside the health sector and their attractiveness from a health policy perspective.

It is important to remember that macroeconomic concerns *aggregate* populations, and not individuals. In many cases, changes in macroeconomic policy will have different impacts on different groups; some will benefit and some not. For example, employment creation in one area may be accompanied by job destruction in another area as labor moves from one sector or industry to another. This process, known as "churning," requires social safety nets and smooth employment transition mechanisms to avoid negative effects on the workforce. Not only might economic insecurity mean that nonpoor households become poor, affecting material conditions (nutrition, housing), but it can also itself be stressful, thereby directly impacting psychological health (Wilkinson & Marmot, 2003).

A growing body of research describes the impact of the distribution of economic growth on population health inequalities, and the experience of relative deprivation as a cause of ill health. In the United Kingdom, for example, the seminal Whitehall Study—a longitudinal study of behaviors and health status of more than 10,000 British civil servants—found that over a 10-year period males age 40 to 64 in the lower grades (clerical and manual) of the civil service had an age-adjusted mortality rate that was 3.5 times higher than that for males in senior administrative grades (Marmot et al., 1991; Marmot & Theorell, 1988; Wilkinson, 1986). A gradient in mortality stretched across all levels of the civil service: Each status group had a higher mortality rate than the group immediately above it, which the provision of medical care or other primary goods, such as absolute levels of income, shelter, food, or education, could not explain.

Other research, aimed at confirming the Whitehall Study findings cross-nationally, has revealed a similar correlation between hierarchy and health. In a study of 11 Organization for Economic Cooperation and Development (OECD) countries, conducted from 1975 to 1985, Wilkinson (1990) used two measures of income inequality—the Gini coefficient and the ratio between the percentage of income flowing to the top and bottom 30% of a country's population. The analysis identified an inverse relationship between income inequality in a country and the average national life expectancy. While nations with absolute levels of higher average income per capita had higher average life expectancy than those with a lower level of average national income (e.g., higher-income countries versus lower-income countries), the study suggested that greater income inequality *within* a country might also have a negative influence on life expectancy. Nations with a relatively flat income gradient, such as Japan and Sweden, had higher life expectancies than countries such as the United States and West Germany, which had steeper gradients (Preston & Elo, 1995).

More recent work supports the findings of an income–health gradient. A Canadian study found that people in lower-income deciles had a lower health-adjusted life expectancy compared with people in higher-income deciles (McIntosh, Finès, Wilkins, & Wolfson, 2009). Likewise, a study in South Korea found that both men and women in lower socioeconomic positions had lower levels of self-reported health status and higher levels of morbidity, even after controlling for behavioral risk factors (Kim & Ruger, 2010).

Other work has interpreted the apparent effects of income inequality on health as resulting from effects of ranking within a reference group that is much smaller than a country (Deaton, 2003; Karlsson, Nilsson, Lyttkens, & Lesson, 2010). If, for example, an academic had a low salary relative to other members within her discipline, this disparity might have a deleterious effect on health even if her absolute income was relatively high and she lived in a country where income was relatively evenly distributed. Even if individuals' reference groups are unobserved, individual income levels will provide partial information about the individuals' standing in their reference group (with the amount of information depending on the relative levels of within-group and across-group inequality). The widely documented adverse effects of inequality on average health status can occur even if either within-group or between-group inequality has no real

effect on health. Deaton and Paxson's (2001) analyses using U.S. data support the conclusion of no effect. Zhang and Wang (2010), in contrast, found differences in development among children in eastern and western China that were related to regional economic status and living standards. The jury remains out on how reference-group rankings and socioeconomic inequalities influence health.

Just as health conditions at any time seem to improve with income level, so, too, might adverse income shocks have detrimental effects on health. Economic recessions—with their concomitant declines in employment and income, as well as rising uncertainty regarding employment and financial well-being—have been linked to increased morbidity and mortality, and to poor nutrition and mental health (Catalano, 2009; Uutela, 2010).

An economic crisis may affect human health through myriad channels. First, loss of employment and reduction in wages can increase stress and lead to physical and mental health problems, including substance abuse, depression, and suicide (Uutela, 2010). An important finding in this line of research is that the health effects of economic fluctuations and resulting unemployment can be ameliorated by social protection mechanisms that offset the loss of income (Stuckler, Basu, Suhrcke, Coutts, & McKee, 2009).

Second, with less real income, individuals and households tend to spend less money on the promotion and preservation of health. This trend could mean reduced attention to primary care, inadequate adherence to medication regimens, and impaired access to high-quality food or adequate quantities of food. The Food and Agriculture Organization (FAO) of the United Nations reports that household attempts to cope with economic crises can put both child and adult nutrition and health status at risk. Such strategies include replacing more nutritious food with less nutritious and less expensive food options, reducing overall food consumption, and reducing the frequency of healthcare visits (Brinkman, de Pee, Sanogo, Subran, & Bloem, 2009; FAO, 2009).

Third, fewer financial resources may be available for public health programs during and after economic crises. During the 2008 economic crisis, for example, high-income countries spent large sums of public money on "bailout" initiatives to generate higher levels of market liquidity. These policies may put global population health at risk, to the extent that governments finance these macroeconomic policies through cuts in foreign assistance for health and cuts in their own future social spending.

Finally, individuals and households may spend down their savings and assets during economic crises. This practice then renders them more vulnerable to future economic shocks.

Past experience shows that the health of the poor is especially vulnerable to the negative effects of economic downturns. FAO (2009) has reviewed the impact of past crises on nutrition and health among the poor. For example, Cameroon's crisis of the 1990s was associated with an increase in the prevalence of underweight children, especially among the poorest half of the population. Zimbabwe's drought of the mid-1990s caused losses in real income, leading to stunting among young children in the poorest households—a phenomenon that has been linked to cognitive impairments and reduced earnings during adulthood (FAO, 2009). If past crises are any indication of what can be expected from the 2008 global economic crisis, the full effects are likely to be serious, widespread, and long-lasting, with significant implications for other realms of well-being (Chang, Stuckler, Yip, & Gunnel, 2013; Kaplan, 2012) (**EXHIBIT 19-2**).

EXHIBIT 19-2 Global Economic Crisis of 2008

The global economic crisis that began in 2008 has raised important and ongoing issues for population health. In discussing these issues, it is important to remember that the crisis came on the heels of the food and fuel crises of 2006–2008, which had already begun to erode people's health and the purchasing power of their income in a number of countries. In addition, the crisis began in high-income countries— unlike the financial crises of the 1980s and 1990s, which began in the low- and middle-income world. In LMICs, the 2008 global recession has meant declines in exports, foreign direct investment, remittances, and donor aid, and rising rates of unemployment. Government budgets were also negatively affected, which in turn has had detrimental effects on health-sector funding.

The true magnitude of the health effects of the 2008 economic crisis remains largely unknown, and the lack of systematized monitoring of the situation has been noted (Richards, 2009). In addition, little is known about the distribution of impacts across the income spectrum. For some groups within the larger population, reductions in income might potentially lead to improved health behavior and reductions in risky behavior, such as increased walking or cycling in place of driving (Catalano, 2009). For the most part, however, available data indicate an overall negative impact on population health from such economic downturns (Parmar, Stavropoulou, & Ioannidis, 2016). FAO (2009), for example,

reported on problematic consequences in Armenia, Bangladesh, Ghana, Nicaragua, and Zambia. In Nicaragua, the crisis led to a decline in the consumption of meat and dairy products, which provide macronutrients and micronutrients that are particularly vital for young children and women who are pregnant or lactating. FAO's Zambia case documents severe impacts of the crisis on healthcare access and utilization: Workforce reductions in the mining sector resulted in the loss of both income and benefits provided by mining companies. This, in turn, translated into a loss of access to the hospitals and clinics established by the companies, which provided not only antiretroviral treatment for HIV-positive individuals, but also nutritious diets.

Implications of the economic downturn for health systems have also been explicated. For example, U.S. hospitals reported increases in uncompensated care and reductions in revenues as a result of the crisis. In response, many hospitals postponed capital investments, such as infrastructure renovations, purchasing of equipment, or upgrades to information technology systems (Dranove, Garthwaite, & Ody, 2015).

Although a full picture of the 2008 crisis's impact on health is lacking, we can look to past crises to shed some light on this issue. For example, Frankenberg et al. (1999) examined data contained in the Indonesia Family Life Surveys to study the impact of the 1998 economic crisis on household expenditures, employment, education, utilization of healthcare and family planning services, and health outcomes. Although their study found little evidence of a negative impact of the crisis on health outcomes, perhaps because of the narrow time frame, it revealed that the economic stress had negative impacts on health-seeking behavior and health system operations. In particular, the study noted decreased use of public health services by both adults and children, declines in vitamin A supplementation for children, medication and supply stock-outs at public facilities, and price increases in both public and private health facilities.

Similarly, an Australian Agency for International Development (1999) report explored the magnitude of the health impacts of the 1997–1998 Asian economic crisis. The report cited a sharp increase in the price of pharmaceuticals, declines in vitamin A intake, and declines in the frequency of child healthcare visits during this period. In another study of the link between macroeconomic crises and health, Brainerd (1998) examined the effects of market reform on mortality in Russia and other transition economies in the late 1980s and early 1990s. Standardized death rates in these countries increased from 1989 to 1994, with Russia and Latvia experiencing the greatest increases—more than 50%. Brainerd provided strong evidence of a significant positive association between unemployment and the standardized death rate for adult males across 15 countries in the former Soviet Union.

Finally, it is important to recognize that most, if not all, economies are not self-contained, but rather are integrated into the global system, as indicated in Figure 19-1. Integration is generally associated with more rapid growth rates, which then accelerate the health effects from development, although the full picture for any one country will inevitably be more nuanced. We do not consider international trade further here, as the *International Trade and Health* chapter specifically focuses on aspects of trade and health.

Macroeconomics and Risk Factors for Disease

The relationship between the economy and health is mediated through risk of disease and injury, which is related in part to household behavior and in part to far wider considerations. Economic behavior partly determines this risk, and the cost incurred from this risk partly determines economic performance. Often a distinction is made between the direct costs—meaning the impact of prevention and treatment of disease and injury on health systems—and the indirect costs—meaning the non-health-system effects on labor productivity or other sectors. A distinction is also made between the risks associated with communicable disease (an important subset of which are major outbreaks or pandemics), which are often the most visible in a global context, and the risks associated with noncommunicable disease (NCD). For instance, most of the health areas subject to intensive global concern and funding relate to infectious diseases—especially malaria, HIV/AIDS, and pandemic preparedness—but increasing attention is now being paid to the (economic and health) impacts associated with diet, obesity, tobacco consumption, unhealthy consumption of alcohol, and aging. In this section, we look at the relationship between the macroeconomy and the risk factors associated with the three areas of communicable disease, pandemics, and noncommunicable disease.

Communicable Disease

Communicable diseases are transmitted by some type of vector—human, animal, or other. The relationship between macroeconomics and communicable diseases is twofold. First, the overall environment in which people live (e.g., air quality, sanitation, housing) is determined, in large part, by their income and wealth. Second, the increased movement of people, animals, and goods associated with an increasingly

active economy—especially internationally—affects the movement of disease. Although several other communicable diseases have emerged or re-emerged in recent years, the major international focus remains on tuberculosis (TB), malaria, and HIV/AIDS; the latter two considered in more detail here.

Malaria remains one of the most significant public health problems in lower-income countries and generates significant economic costs (as discussed in the *Infectious Diseases* chapter). Chima and colleagues (2003) and Malaney and colleagues (2004) categorize the direct costs of malaria as comprising the direct costs of prevention (e.g., mosquito coils, aerosol sprays, bed nets, residual spraying, and mosquito repellents) and treatment (e.g., drugs, treatment fees, transport, and costs of subsistence at a health center), and the indirect costs of labor time lost because of illness. Given these two components, household and public-sector expenditures are increased and labor inputs are decreased due to malaria. Reduced school attendance and performance, and increased cognitive impairment, can likewise reduce human capital. For example, using the malaria eradication programs in the United States in the 1920s and in Brazil, Colombia, and Mexico in the 1950s as natural experiments, Bleakley (2010) found that children in malarial areas in these countries born just after the eradication programs were undertaken had significantly higher earnings than those born just before the programs were initiated.

Although studies of malaria tend to evaluate the costs of disease within a specific economic environment, potentially more significant is the effect of disease on broader economic opportunities. If a factory goes unbuilt in a region because of malaria, for example, the consequences for income may be far more substantial than malaria's consequence for productivity in traditional agriculture. Gallup and Sachs (2001) advance this perspective, arguing for a macroeconomic approach that uses cross-country time-series data to capture effects that would go unobserved with collection of only household data. In particular, their seminal work uncovered a significant negative relationship between malaria and GDP growth, with the effects having a far greater estimated magnitude than those reported in household-level studies, and suggesting that the most important ways in which malaria affects long-run economic growth are through its effects on a country's ability to attract foreign direct investment and to create an environment suitable for modern economic growth. Thus, allowing countries to undertake new enterprises, rather than just improving productivity in traditional enterprise, may be the most economically significant consequence of malaria control programs.

The reverse is also true: Economic development can have a significant impact on risks of malaria. Tusting et al. (2013) undertook a systematic review and meta-analysis of studies that measured socioeconomic status and parasitologically confirmed malaria or clinical malaria in children age 0 to 15 years. The odds of malaria infection were significantly higher in the poorest children than in the least poor children, and were robust to subgroup analyses. Lower socioeconomic status was also associated with environments more conducive to malaria transmission, such as poor housing quality, inadequate diets, and lack of bed nets. The authors concluded that increased investment in interventions to support socioeconomic development is clearly warranted, as such initiatives could prove highly effective and, more importantly, be sustainable against malaria in the long term.

In contrast, and perhaps counter-intuitively, studies have found the economic effects of HIV/AIDS on the economy can be relatively small. For example, a seminal study of the impact of HIV/AIDS on India's national economy (Anand, Pandav, & Nath, 1999) found that the estimated total annual costs of HIV/AIDS in India under low (1.5 million infected), medium (2.5 million), and high (4.5 million) prevalence scenarios were 7 billion, 20 billion, and 59 billion rupees ($200 million, $571 million, and $1.7 billion), respectively. AIDS treatment and productivity loss were the two major components of these costs. From this, the estimated annual cost of HIV/AIDS (assuming a prevalence of 4.5 million persons) was roughly 1% of India's GDP. Likewise, other studies have found that, especially early in its course, the AIDS epidemic had an insignificant effect on the growth of per capita income (Haacker, 2004).

But this is not the entire picture. One reason for the small financial effects associated with HIV/AIDS is that economic consequences are usually measured in terms of income per capita. A high death rate reduces both income and the number of people in the economy, causing a sharp reduction in total output, but producing a more modest effect on income per capita. Thus, using as a metric the income level of the survivors, as is done with income per capita, is inappropriate. When Jamison and colleagues (2001) placed a value on decreases in a country's mortality rates and added this factor to the change in per capita income, they identified dramatic negative economic consequences of HIV/AIDS in five countries of eastern Africa in the period after 1990. The importance of the direct welfare effects of good health have also been emphasized by Nordhaus (2003), who argues that the welfare gains from improved health and longevity in the United States in the 20th century were greater than

the welfare gains due to rising income levels over the period.

According to Bell and colleagues (2003), the large number of orphans in countries with high AIDS mortality rates will lead to a higher short-term dependency ratio, lack of resources for education, and the production of successive cohorts with lower levels of schooling and income. In particular, this trend may plunge sub-Saharan Africa into even greater poverty. Moreover, the economic impact of HIV/AIDS could take years to realize and may be much greater and more widespread than forecasts estimate (Veenstra & Whiteside, 2005).

Outbreaks and Pandemics

There is a growing concern about the potential economic impact of infectious disease outbreaks and pandemics, following the outbreaks of severe acute respiratory syndrome (SARS) in 2003, H5N1 avian influenza in 2006, H1N1 swine influenza in 2009, and most recently Ebola in 2013 to 2016. Brahmbhatt (2005) highlights two major types of economic costs that arise from such outbreaks: (1) the costs of increased illness and death of affected people and animals, and (2) the costs of control and coping strategies to mitigate effects of the pandemic. For instance, the economic impact of the H1N1 outbreak in Mexico was estimated to be caused more by collapsing trade flows for pork products, and tourism, than by the health effects of the disease (Rassy & Smith, 2013).

The outbreak of SARS, in particular, caused a large decline in international travel due to quarantines, fear of infection, and worries about severe economic disruption (Wilder-Smith, 2006). According to Brahmbhatt (2005), the most immediate economic losses during the SARS outbreak were due to absenteeism, disruption of production, and severe decreases in tourism, mass transportation, retail sales, and hotel and restaurant use. Keogh-Brown and Smith (2008), however, found that, in retrospect, the effect of SARS on income levels was relatively short-lived and modest (**EXHIBIT 19-3**).

Pandemic influenza raises more concern, as it potentially combines the lethality of SARS with far greater pathogenicity. Studies have attempted to analyze the direct illness effects of influenza pandemics, typically through modeling the production effects of changes in labor supply. For example, Keogh-Brown and colleagues (2010) estimated the impact on the U.K. economy using epidemiologic data from pandemics that occurred in 1957 and 1968. Their analysis suggested that the economic impact of a repeat of the 1957 or 1968 pandemics would be short-lived, constituting a loss of between 3.35% and 21% of GDP in the first pandemic quarter and between 0.58% and 4.5% of GDP in the first year for various scenarios. Critically, the economic shockwave would be gravest when absenteeism (through school closures) increased beyond a few weeks, creating policy repercussions for influenza pandemic planning—the most severe economic impact is due to policies to contain the pandemic rather than the pandemic itself.

In another article by Smith and Keogh-Brown (2013), estimates are presented of the macroeconomic impact of the H1N1 (swine flu) pandemic of 2009 on the economies of South Africa, Thailand, and Uganda. Estimates suggest that the direct illness effects of such a pandemic are less than 0.5% loss of GDP, although the relative impacts on different economies will vary. In the 2009 pandemic, the impacts in Uganda were relatively larger than those in South Africa, and both

EXHIBIT 19-3 The Economics of SARS

Severe acute respiratory syndrome (SARS) is an infectious disease that spreads between humans in a similar way to the cold virus. Emerging in the Guandong province of China in late 2002, the disease was transmitted with great rapidity to Australia, Brazil, Canada, China, Hong Kong, South Africa, Spain, and the United States, creating serious public health concerns. The outbreak peaked during the second quarter of 2003 and was declared over by July 2003.

Although approximately 10,000 individuals were infected, of whom 10% died, the overall impact on health was far less devastating than initially feared. However, the possible economic impact of SARS was also a focus of concern. During the outbreak, travel and tourism income took a noticeable downturn in many infected countries. It was anticipated that fear of disease would negatively affect those industries that gather people in public places, such as restaurants, cinemas, and retail establishments.

The Canadian Public Health Agency (2003) suggested that SARS cost the country $350 million in lost tourism revenues, $220 million from reduced airport activity, and $380 million in lost nontourism retail sales. Estimates of the costs to Canada from SARS ranged as high as $2 billion. Keogh-Brown and Smith (2008) suggest that SARS caused losses to Hong Kong's GDP of approximately $3.7 billion and losses of 3% in China's GDP, even though less than 0.03% of Hong Kong's population and less than 0.0004% of China's population were actually infected with SARS. The overall estimate of the global macroeconomic burden was between $30 billion and $100 billion (approximately $3 million to $10 million per case).

of those countries' estimated GDP losses were 50% to 100% larger than the loss for Thailand. Interestingly, these studies found that the potential indirect effects on the economies concerned were more significant than the direct effects; a similar pattern found with studies of the Ebola outbreak (World Bank, 2015).

Based on the experience of the SARS outbreak, Smith (2006) argues that in the early stages of a new infectious disease, the key issue is risk, including the unknown dimensions of the threat, and the sense of panic that this uncertainty instills in the population. Thus, in the early stage of a pandemic, obtaining accurate assessments of the disease and appropriate responses is vital. Kleinman and associates (2008) also argue for a biosocial approach to new infectious diseases, combining the speedy acquisition of a biological understanding of the disease with policies that get the relevant information, and policy prescriptions, into the public domain.

A study by Smith and colleagues (2009) supports this general conclusion that the behavioral effects of a pandemic or significant outbreak are at least as important, if not more important, for an economy than the direct health impacts. These researchers examined the potential economic costs of a modern pandemic based on scenarios for pandemic influenza in the United Kingdom. While they found that moderate GDP losses would ensue (ranging from 0.5% in mild pandemics to 2% in severe pandemics), behavioral changes to avoid infection could potentially double these costs. For example, the cost difference between closing schools for 4 weeks at a pandemic's peak and closing schools for all 13 weeks of a pandemic was estimated to be £27 billion ($44 billion). Similarly, analysis of prophylactic absence from work showed that if a behavioral response that prompts individuals to refrain from work can be avoided, then the potential value of interventions to prevent a harmful economic response might be greater than the value of preventing negative health effects alone. In addition, if fear is the driver of behavioral change, the mortality rate of an outbreak might have a more significant effect than the number of people infected. The validity of these conclusions clearly depend on the assumptions made in scenario design, but the model results capture interaction effects that cannot be captured from a microeconomic approach focused on health alone. A further study by Smith and colleagues (2011) found similar behavioral change effects, such that prophylactic absence and school closure effects outweigh the economic impacts of the disease itself. An emphasis on preparedness, emergency procedures, risk assessment, and preventive actions at an outbreak's onset is potentially important to reduce the macroeconomic effects from such epidemics.

Nevertheless, in addition to the costs associated with an epidemic, or any other health concern, some financial benefits may accrue from such an event. Macroeconomics especially is concerned with flows of money around the economy. Thus, a financial cost to one sector, organization, or individual typically represents a financial benefit (income) to another sector, organization, or individual. For instance, a SARS outbreak or pandemic influenza can be good news if a company is involved in the manufacture of face masks or alcohol gel, or even drugs such as Tamiflu. It is even possible that a health emergency could engender positive economic outcomes overall on GNI, if the increases in incomes outweigh the losses. At the extreme, this is the effect associated with "post-war boom" involved in reconstruction. Of course, this does not mean that war is desirable, anymore than a pandemic would be. Rather, it illustrates the importance of recognizing the limitations associated with assessing macroeconomic impacts based on financial measures, and of critically understanding the basis for them to aid appropriate interpretation.

Noncommunicable Disease

Perhaps less obvious is the relationship between macroeconomic activity and risks concerning noncommunicable disease. Although macroeconomic growth can be beneficial when it leads to an expansion in the consumption of the goods that *improve* the probability of better health, such as clean water, safe food, and education, it also facilitates the increased consumption of goods that may be harmful or hazardous to health, such as tobacco, alcohol, and "fast food." As highlighted in the *Chronic Diseases and Risks* chapter, NCDs can have a significant impact on national economies, albeit often through a different profile of effects than communicable disease, and with greater impact on long-term reductions in productivity of the workforce (through morbidity) as well as impacts on the size of that workforce (through mortality) (Leeder, Raymond, Greenberg, Liu, & Esson, 2004).

In recent years, a greater effort has been made to document the economic impact of NCDs. For example, Tsai and colleagues (2011) estimated that 5% to 10% of overall U.S. healthcare expenditures is attributable to the NCD-related health effects of being obese or overweight. Other studies (Butcher & Park, 2008; Lakdawalla, Bhattacharya, & Goldman, 2004) suggest that some increases in unemployment rates in the United States, in the face of overall improvements in population health, may be attributed to obesity. Similarly, while their results may reflect associations rather than causal effects, Lindeboom and colleagues

(2009) present evidence of the existence of differences in employment probabilities between obese and nonobese men and women that do not disappear when controlling for many potentially important factors. Further, Carter et al. (2013) show that health shocks, such as cancer diagnosis, result in a significantly increased risk of subsequent nonparticipation in the labor force.

In general, perhaps the most significant relationship between the economy and NCDs is the risk posed through poor nutrition, and vice versa, especially in children. Research suggests that links exist between childhood health and adult productivity. For example, adult height as a marker for general health has been found to be inversely related to mortality, implying that a shorter height can ultimately result in fewer productive working years (Schultz, 2002). Other studies have assessed the effects of nutritional status on productivity in lower-income countries (Martorell, Melgar, Maluccio, Stein, & Rivera, 2010) and the cumulative impact of parental nutrition, childhood nourishment, and health over the course of the life cycle on adult height, which ultimately affects individuals' wage and earning potential.

In addition to the long-term effects of early childhood health and nutrition, adult health and nutrition have direct effects on productivity. For example, a number of studies over many years (Alderman & Linnemayr, 2009; Basta, Soekirman, Karyadi, & Scrimshaw, 1979; Gardner, Edgerton, Senewiratne, Barnard, & Ohira, 1975; Spurr, 1983) have demonstrated the influence of iron deficiency on oxygen consumption and the positive impact of iron supplementation on labor productivity. In Mexico, Knaul (1999) studied the productivity gains from investments in health and nutrition for women. The author analyzed the impact of age at menarche, or first menstruation (an indicator of cumulative health and nutritional status), on female labor market productivity, finding that investments in health and nutrition had significant effects on productivity. A 1-year decline in age of menarche was associated with a 25% increase in wages. More recently, Martorell and colleagues (2010) found substantial improvements in future human capital and economic productivity in a review of the long-term impacts of a nutritional intervention. Overall, the effects of nutritional status on work performance are of interest to policy makers in designing food and economic programs to improve energy and nutrient intake.

In a different approach, Lock and colleagues (2010) assessed the macroeconomic impacts of healthier eating as a result of reductions in consumption of meat and dairy products in the United Kingdom and Brazil using a linked trade model. These authors suggest that, while reducing consumption of unhealthy goods can generate positive health-related macroeconomic effects, the substitution effects from reduced consumption are key drivers of both the health and economic impacts. For example, if reducing consumption of meat and dairy products leads to increased consumption of other unhealthy food goods, it will reduce the health benefits. In addition, the ease with which other sectors can expand in response to reduced consumption of unhealthy products drives the economic impacts. The authors highlight that countries that are dependent on international trade in meat products—as is true for Brazil—may suffer economically if their trade partners reduce consumption of meat. Such countries will also suffer in terms of adverse health effects if reductions in trade lead to increased availability of unhealthy goods on the domestic market.

Perhaps the most studied NCD is cardiovascular disease, with a major report from the U.S. Institute of Medicine (2010) indicating that this disease has significant economic consequences for health care and beyond. Leeder and colleagues' (2004) overview indicated major impacts across a range of countries from cardiovascular—for instance, payroll losses to India of more than $198 million per year and social security costs due to disability payments of $64 million in 2000. The same authors also demonstrated how NCDs further affect economies by dampening productivity and economic growth.

Similarly, there are significant economic costs from mental health disorders, as indicated in the *Global Mental Health* chapter, again due to reduced productivity and premature mortality. Although the evidence supporting this relationship is growing, it comes mainly from high-income countries, with relatively few studies focusing on LMICs (Lund et al., 2011). It is also the case that economic conditions, especially unemployment, can significantly determine mental health, with the most extreme association being between economic distress and suicide (Chang, Gunnell, Sterne, Lu, & Cheng, 2009).

The Macroeconomy and the Healthcare Sector

Considerable research has addressed the internal aspects of finance and delivery of health systems, as covered in the *Design of Health Systems* chapter. Nevertheless, health systems ("sectors" in macroeconomic language) warrant attention here as well, because they are greatly influenced by the economy, are significant and growing economic sectors, and in many cases are one of the largest government (public) sectors.

Perhaps the most visible link between macroeconomics and the health sector is at the overall level of healthcare spending. Most nations, rich or poor, face the problem of per capita healthcare costs rising faster than per capita GDP. For instance, total global expenditures on health care grew from 3% of GDP in 1948 to 9.9% by 2014. Moreover, this sector takes an increasing proportion of GDP for countries as they develop, rising from 4.5% in lower-middle-income countries to 12.3% in high-income countries (World Bank, 2017).

An important piece of work conducted by Schieber and Maeda (1997) found that the income elasticity of demand for health services—defined as the percentage change in health expenditures as a result of a 1% change in income—in 1994 was roughly 1.13 worldwide, with the range being 1.47 for high-income countries, 1.19 for middle-income countries, and 1.00 for low-income countries. Thus, for every 1% increase in per capita income, health expenditures increase by 1.47% in middle- and high-income countries, by 1.19% in middle-income countries, and by 1.00% in lower-income countries.[1] As these data indicate, the increase in health expenditures in low-income countries is less responsive to increases in income than it is in middle- and high-income countries. Schieber and Maeda also found that private health spending (income elasticity of 1.02) is less responsive to increases in per capita income compared to public health spending (income elasticity of 1.21).

On a global scale, healthcare outlays are absorbing an increasing share of government, employer, and household incomes. This constant fiscal pressure forces nations to confront two basic questions: How to finance this rising burden and how to contain the pressures for health expenditures growth. Thus, an essentially macro issue drives the most fundamental of all health system issues. Most countries spend less than 10% of GDP on health care, which is seen as perhaps the stable upper limit. Of course, for low-income countries, and for some middle-income countries, their GDP levels are so low that even a 10% allocation produces very little actual health care. Aid assistance is therefore required if they are to provide essential health care. Importantly, for low-income countries, increasing the share of GDP devoted to health care means that less is available for spending on other vital areas that would encourage broader economic growth, which itself will enable increased nominal (actual monetary amount) spending on health care (**EXHIBIT 19-4**).

EXHIBIT 19-4 Development Assistance for Health

For many of the lowest-income countries, the absolute level of spending they are able to devote to health care is deemed to be well below the minimum required; instead, these countries rely on international finance to fund their healthcare services. International finance can take many forms. Official development assistance (ODA) is defined as funding flows to the Organization of Economic Cooperation and Development, Development Assistance Committee's (OECD-DAC) list of recipients, which are administered for the purposes of promoting the economic development and welfare of developing countries, and which are concessional in character. ODA includes loans with an equivalent grant element of 25% or more. Development assistance for health (DAH) is slightly broader than ODA and includes non-concessional loans and funds from private foundations and nongovernmental organizations (NGOs) that contribute directly to the promotion of development and welfare in the health sector in developing countries.

The provision of DAH is complex, with no dominant group of actors, standardized set of aims, or consistent approach. A wide variety of public and private institutions provide assistance, from both OECD and non-OECD countries. Moreover, development funds often do not flow directly from sources of finance (government treasuries, private contributions to NGOs) to the recipient, but rather may flow through a variety of channeling institutions such as bilateral aid agencies, multilateral institutions, and development banks, or through intermediary funds such as the Global Fund for Aids, Tuberculosis and Malaria.

Funds can be provided with the aim of improving health services directly or indirectly (e.g., improving governance) or to support economic well-being that may impact the determinants of health. Conversely, funding to health sectors may be used to support and complement broader development objectives such as good governance or ensuring security. Finally, implementation can be done directly by governments, by donors, or through third parties, such as private companies (Vassall & Martínez Álvarez, 2012).

[1] Many economists view the evidence regarding the income elasticity of expenditures on health as indicative of the nature of health as a normal consumption good, or even a "luxury good" (with rising proportions of income spent on it as income increases). However, in recent years there has been an upsurge of interest in the nature of health care as an investment good. Expenditures on health care can lead to improvements in health status that, in turn, promote income growth, and the efficiency and equity of health systems are important in both improving health and affecting income, productivity, and the overall economy.

TABLE 19-2 Health Expenditures by Income Group and Region, 2014

Region or Income Group	Total Health Expenditures per Capita	Total Health Expenditures (% of GDP)	Public Health Expenditures (% of GDP)	Public Health Expenditures (% of Total Health Expenditures)
World	1,059	9.9	6.0	60.1
Low-income countries	37	5.7	2.4	42.4
Middle-income countries	290	5.8	3.0	52.0
High-income countries	5,266	12.3	7.7	62.3
Low- and middle-income countries	265	5.8	3.0	51.7
East Asia and Pacific	329	5.3	4.6	55.1
Europe and Central Asia	575	9.5	7.2	75.5
Latin America and Caribbean	694	7.3	3.8	51.0
Middle East and North Africa	253	5.3	3.1	49.5
South Asia	67	4.4	1.4	31.2
Sub-Saharan Africa	98	5.5	2.3	42.6

World Bank, 2017.

As **TABLE 19-2** illustrates, in 2014 LMICs, on average, spent 5.7% of GDP on health, whereas high-income countries spent 12.3% of GDP on health. Regionally, South Asia spent the least, 4.4% of GDP, whereas Latin America and the Caribbean spent the most, approximately 7.3% of GDP. In cash terms, average spending on health care by low-income countries was $37 per capita in 2014, compared with $290 for middle-income countries and $5,266 for high-income countries. In the lowest-income countries, annual health spending can be as low as $2 or $3 per capita. As well as health spending generally, the public share of health spending increases with increases in national and per capita income. As shown in Table 19-2, the public share of health expenditures (as a percentage of total health spending) was 42% in low-income countries, 52% in middle-income countries, and 62% in high-income countries in 2007 (World Bank, 2017).

Two elements of the economy also determine overall healthcare finance, which are linked to wider trade and, therefore, covered in more detail in the *International Trade and Health* chapter. Briefly, government income available to spend on health care (and other public services) is generated primarily through taxes. Tax income is broadly dichotomized into taxes that are "easy to collect" (such as import tariffs) to those that are "hard to collect" (such as consumption taxes, income tax, and value-added tax [VAT]) (Aizenman & Jinjarak, 2009). Tariff revenues are a important source of public revenues in many developing countries; their share ranges from less than 1% of public revenues within the OECD to approximately 80% in Guinea, with typical examples of Cameroon at 28% and India at 18%. As economies develop and join more fully with the global trading environment, these tariffs tend to fall, reducing the proportion of government income from "easy to collect" sources. Although theoretically governments should be able to shift tax bases from tariffs to domestic taxes, such as sales or income taxes, in practice LMICs find this

difficult, especially because of the informal nature of their economies and the large subsistence sectors (Aizenman & Jinjarak, 2009). Whereas high-income countries are usually able to replace 100% of the lost tariff revenues, middle-income countries replace only 40% to 60% and low-income countries at best manage just 30% replacement of these tariffs (Baunsgaard & Keen, 2005).

The most common source of tax revenues as countries increase their economic development is income tax, or more precisely payroll tax. This is basically any tax that is taken off the wage, through the payroll, before the employee receives it—the difference between the gross wage and the net wage is reflected in the payroll tax. It includes elements such as wage tax, national insurance contributions, the employed person's insurance contributions, and so forth. As countries grow and expand their spending on health care (as indicated earlier), they increasingly rely on payroll taxes to fund their expenditures. This shifts an increasing amount of the burden of healthcare costs to the formal paid workforce, which may affect the relative efficiency of the workforce in the global economy (i.e., wages may be higher to reflect these costs, which are then passed through to increased costs for exports and a reduction in the country's competitive position).

A second macroeconomic determinant of health sector finance, also linked to trade, is the exchange rate. This factor is a key determinant of the relative prices of imported and domestically produced goods and services. Importantly for the health sector, many of the goods (principally pharmaceuticals, but also various elements of other technologies, such as computer equipment, surgical tools, and even lightbulbs) used to provide health care are imported. Changes in the exchange rate brought about by macroeconomic developments may, therefore, prompt the price, and hence the cost, of health care to increase or decrease. Conversely, changes in demand for domestically produced goods from overseas importers may trigger changes in the price of those goods domestically in response (e.g., increased foreign demand may push up local prices). Thus, increased linkage between economies at the macro-level generates greater levels of exogenous (i.e., beyond the domestic health sector control) influences over prices, and hence the costs of health care.

The Macro-Efficiency of Healthcare Sectors

Irrespective of funding, countries differ in how efficiently their health systems convert resources into population health status, a relationship that is often referred to as the "macro-efficiency" of health systems. The macro-efficiency reflects how aggregate healthcare expenditures are linked to health outcomes. Considerable variation exists by country in health status as well as healthcare expenditures. For example, in 2015, life expectancy at birth ranged from 51 years in the Central African Republic to more than 80 years in France, Hong Kong, Iceland, Italy, Switzerland, and Japan (World Bank, 2017). While the global average life expectancy was 72 years, for low-income countries it was 62 years, middle-income countries 71 years, and high-income countries 81 years (World Bank, 2017). Of course, as indicated earlier, many factors influence health outcomes, and expenditures on health care often correlate poorly with outcomes, as illustrated in **TABLE 19-3**.

There is little, if any, evidence of the opposite relationship—that is, the impact on the economy of greater spending on health care. It is clear that a sector accounting for more than 10% of GDP is a significant component of that economy, but little research has addressed how changes to the overall size of the health sector affect, for instance, overall GDP or employment. For instance, higher healthcare expenditures might reduce a country's relative productivity in the global economy due to the large costs of health benefits for the labor force, and due to the resultant reduced investments in other sectors of the economy. Even so, it is not clear that overall productivity and growth are severely hindered by high healthcare expenditures.

Overall, it is not possible to determine the most efficient level or percentage of healthcare spending for any given economy. Given the wide range of spending on health care, the challenge is to increase spending among lower-income countries and for all countries to enhance the efficient production of health within health systems. This topic is explored in more detail in the *Design of Health Systems* chapter.

Conclusion

Evidence shows a solid positive bidirectional relationship between a country's wealth and its population's health. Whatever the precise mechanisms, broadly wealthier populations are healthier, and healthier populations are wealthier. Yet, despite the long-established relationship between health and the macroeconomy, application of a macroeconomic lens in health policy debate is comparatively rare. Moreover, health effects are disregarded in many macroeconomic policy applications. Much remains to be done to enable holistic policy analysis that correctly values the health effects and health implications of

TABLE 19-3 Healthcare Expenditure (HCE), Proportion of Expenditure That Is Government-Funded, Infant Mortality Rate, and Life Expectancy at Birth for Selected Countries in 2014, Ranked by Overall HCE

Country	HCE (% of GDP)	% of HCE That Is Publicly Funded	Infant Mortality Rate (per 1,000)	Life Expectancy (years)
India	4.7	—	38	68
Mauritius	4.8	49.2	12	74
China	5.5	55.8	9	76
United Kingdom	9.1	83.1	4	82
Japan	10.2	83.6	2	84
Germany	11.3	77.0	3	81
France	11.5	78.2	4	83
United States	17.1	48.3	6	79

World Bank, 2017.

economic policies and—perhaps more important in leveraging health policies—the macroeconomic effects of health policies. Although detailed, sector-specific health economic analysis has an important role to play in capturing individual-level costs and health-sector implications of health policies and disease burdens, the macroeconomic implications and methods for macroeconomic analyses of health are currently under-researched and underutilized. This chapter has provided an overview of this relationship, examined through the lenses of the household (looking especially at the interaction of economic growth and health), risk factors for disease (communicable, noncommunicable, and outbreaks), and the economy and its relationship with the healthcare sector.

Much macroeconomic policy concerns economic growth, as higher GNI leads to more opportunities to consume, which is in turn expected to result in better health. The chapter presented evidence that suggests health is closely linked with economic growth and development, through multiple pathways, including the demographic transition, improved educational attainment, and labor productivity. Indeed, as much as 10% to 30% of economic growth may be due to improved health of a country's population. Similarly, greater economic development can influence health, especially though increased access to more-affordable health determinants, such as housing, education, water, sanitation, and nutrition, as well as through improved health care (although it is important to remember the "income–health gradient").

A clear relationship also exists between the economy and health directed through the risk of disease and injury. Economic behavior partly determines this risk, and the cost incurred from this risk partly determines economic performance. This is true for communicable diseases, including malaria, HIV/AIDS, and pandemics. The overall environment in which people live (e.g., air quality, sanitation, housing) is determined, in large part, by their income and wealth, and the increased movement of people, animals, and goods associated with an increasingly active and international economy will affect the movement of disease. Communicable diseases, then, affect aspects of the economy related to, for instance, educational attainment or labor productivity. Outbreaks provide specific challenges for economies—not just through the immediate costs of increased illness and death of affected people and animals, but also due to the costs of control and coping strategies to mitigate effects of the pandemic such as quarantines, and the reductions in economic activity through fear of infection. In terms of noncommunicable disease, economic growth increases the ability of populations to consume both goods that *improve* the probability of better health, such as clean water, safe food, education, and goods

that may be harmful or hazardous to health, such as tobacco, alcohol, and fast food. Diet, and the impact of under- and over-nutrition, were highlighted in this chapter as significant impacts on economies across countries at all levels of development.

Finally, health systems are greatly influenced by the economy, and are themselves significant and growing economic sectors. The most visible link is the level of healthcare spending that an economy can "afford," with clear relationships being established between the level of economic development and the level of spending, both in absolute terms and in terms of per capita GDP.

Policies that have positive health effects can have significant wider economic impacts, which are often positive. As observed in *The Lancet* (2013), "It is sadly true that most health ministers lack domestic political muscle. They might talk tough amongst themselves, but back home they have to get in line behind colleagues in finance, defence, trade, and even education." Finance often speaks more to those in power than does health. While health may have a powerful economic case to make, it must embrace greater engagement with macroeconomics. As outlined in this chapter, simultaneous capture of the macroeconomic implications of health policies on other sectors and the wider economy is essential for health policies to be considered and weighed fairly against environmental, trade, and agricultural policies. Similarly, the potentially significant health effects and externalities of non-health-focused policies should be captured and considered, as their contributions may be relevant to policy optimization decisions. Macroeconomic assessments and modeling provide an even playing field for policy assessments from both health and non-health perspectives and a medium through which health and non-health policy implications can be jointly assessed using the same framework. This also facilitates dialogue across sectors to enable optimal policy decisions that weigh both the health benefits and the overall economic implications of policies, while accounting for important demographic, welfare, environmental, trade, and agricultural implications (Smith, 2012).

The consequences of health and health policy for economic development are crucial to policy makers in the health and development fields. Development institutions, nongovernmental organizations, and academic communities should seek to investigate strengthen the links between health and economic development. Primarily, they can encourage continued efforts to understand health and development links by helping LMICs develop their own capacity to study these connections; by encouraging further generation of relevant primary data; and by ensuring that the appropriate household-level and cross-national databases are maintained, updated, and accessible to all.

Acknowledgments

Much of the original work used in the 2001 version of this chapter was prepared for the 1998 transition team of the then-incoming World Health Organization Director-General Dr. G. H. Brundtland. Preparation of that version was supported by grants from the Royal Norwegian Ministry of Foreign Affairs to WHO and the World Bank. The WHO Commission on Macroeconomics and Health supported extension and revision, and discussions with the commission's chair, Jeffrey Sachs, provided valuable ideas and feedback during preparation of the 2006 version of this chapter. Conclusions and opinions are those of the authors and not necessarily those of any of the sponsoring agencies. Felicia Knaul, Lawrence Lau, Michael Merson, Pia Malaney, and Anne Mills also made valuable suggestions and comments on earlier versions. Previous versions also benefited from assistance by Leslie Evans to prepare the tables and graphics, Christina Lazar for research assistance, and Elizabeth Cafiero and Marija Ozolins for research and editorial assistance. This version was supported by detailed comments from Anne Mills.

Discussion Questions

1. In what ways and to what extent do major infectious diseases, such as malaria and HIV/AIDS, affect the income and well-being of people living in LMICs?
2. What can policy makers and practitioners do in non-health-service–related settings, such as schools, to improve the health and well-being of girls and boys in LMICs?
3. How do changes in fertility and mortality patterns influence the level and capacity of those persons contributing to economic production in a given society?
4. How have recent pandemics directly and indirectly affected national and global economies?

References

Acemoglu, D., & Johnson, S. (2007). Disease and development: The effect of life expectancy on economic growth. *Journal of Political Economy, 115*(6), 925–985.

Aizenman, J., & Jinjarak, Y. (2009) Globalization and developing countries: A shrinking tax base? *Journal of Development Studies, 45*(5), 653–671.

Alderman, H., & Linnemayr, S. (2009). Anemia in low-income countries is unlikely to be addressed by economic development without additional programs. *Food Nutrition Bulletin, 30*(3), 265–269.

Anand, K., Pandav, C. S., & Nath, L. M. (1999). Impact of HIV/AIDS on the national economy of India. *Health Policy, 47,* 195–205.

Australian Agency for International Development (AusAID). (1999). Impact of the Asia crisis on children: Issues for social safety nets. Retrieved from https://trove.nla.gov.au/work/5720260?q&versionId=46257166

Basta, S. S., Soekirman, M. S., Karyadi, D., & Scrimshaw, N. S. (1979). Iron deficiency anemia and the productivity of adult males in Indonesia. *American Journal of Clinical Nutrition, 32,* 916–925.

Baunsgaard, T., & Keen, M. (2005). Tax revenue and (or?) trade liberalization. International Monetary Fund WP/05/112. Retrieved from http://www.imf.org/external/pubs/ft/wp/2005/wp05112.pdf

Bell, C., Devarajan, S., & Gersbach, H. (2003). *The long-run economic costs of AIDS: Theory and an application to South Africa* (Policy Research Working Paper No. WPS 3152). Washington, DC: World Bank.

Bhargava, A., Jamison, D., Lau, L., & Murray, C. (2001). Modeling the effects of health on economic growth. *Journal of Health Economics, 20*(3), 423–440.

Bleakley, H. (2010). Malaria eradication in the Americas: A retrospective analysis of childhood exposure. *American Economic Journal: Applied Economics, 2*(2), 1–4.

Bloom, D. E., & Canning, D. (2000). The health and wealth of nations. *Science, 287,* 1207–1209.

Bloom, D. E., Canning, D., & Graham, B. (2003). Longevity and life cycle savings. *Scandinavian Journal of Economics, 105,* 319–338.

Bloom, D. E., Canning, D., & Jamison, D. T. (2004). Health, wealth and welfare. *Finance and Development, 41,* 10–15.

Bloom, D. E., Canning, D., & Sevilla, J. (2002). *The demographic dividend: A new perspective on the economic consequences of population change* (MR-1274). Santa Monica, CA: RAND.

Bloom, D. E., Canning, D., & Sevilla, J. (2004). The effect of health on economic growth: A production function approach. *World Development, 32*(1), 1–13.

Boutayeb, A. (2006). The double burden of communicable and non-communicable diseases in developing countries. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 100*(3), 191–199.

Brahmbhatt, M. (2005, September 23). *Avian influenza: Economic and social impacts.* Washington, DC: World Bank.

Brainerd, E. (1998). Market reform and mortality in transition economies. *World Development, 16*(11), 2013–2027.

Brinkman, H. J., de Pee, S., Sanogo, I., Subran, L., & Bloem, M. W. (2009). High food prices and the global financial crisis have reduced access to nutrition food and worsened nutritional status and health. *Journal of Nutrition, Supplement: The Impact of Climate Change, the Economic Crisis, and the Increase in Food Prices on Malnutrition,* 153S–161S.

Butcher, K. F., & Park, K.-H. (2008). Obesity, disability, and the labor force. *Economic Perspectives, XXXII*(1), 1–15. Retrieved from https://ssrn.com/abstract=1096565

Canadian Public Health Agency. (2003). *Learning from SARS: Renewal of public health in Canada. A report of the National Advisory Committee on SARS and Public Health*. Ottawa, Canada: Canadian Public Health Agency.

Carter, K. N., Gunasekara, F. I., Blakely, T., & Richardson, K. (2013). Health shocks adversely impact participation in the labour force in a working age population: A longitudinal analysis. *Australian and New Zealand Journal of Public Health, 37,* 257–263.

Catalano, R. (2009). Health, medical care, and economic crisis. *New England Journal of Medicine, 360*(8), 749–751.

Chang, S., Gunnell, D., Sterne, J., Lu, T., & Cheng, A. (2009). Was the economic crisis 1997–1998 responsible for rising suicide rates in East/Southeast Asia? A time-trend analysis for Japan, Hong Kong, South Korea, Taiwan, Singapore and Thailand. *Social Science and Medicine, 68,* 1322–1331.

Chang, S. C., Stuckler, D., Yip, P., & Gunnel, D. (2013). Impact of 2008 global economic crisis on suicide: Time trend study in 54 countries. *British Medical Journal, 347*, f5239. doi:10.1136/bmj.f523.

Chima, R., Goodman, C., & Mills, A. (2003). The economic impact of malaria in Africa: A critical review of the evidence. *Health Policy, 63,* 17–36.

Deaton, A. (2003). Health, inequality, and economic development. *Journal of Economic Literature, 41*(1), 113–158.

Deaton, A., & Paxson, C. (2001). Mortality, education, income and inequality among American cohorts. In D. Wise (Ed.), *Themes in the economics of aging* (pp. 129–170). Chicago, IL: University of Chicago Press.

Dranove, D., Garthwaite, C., & Ody, C. (2015). The economic downturn and its lingering effects reduced Medicare spending growth by $4 billion in 2009–12. *Health Affairs, 34,* 1368–1375.

Fogel, R. W. (1997). New findings on secular trends in nutrition and mortality: Some implications for population theory. In M. R. Rosenzweig & O. Stark (Eds.), *Handbook of population and family economics* (Vol. 1A, pp. 433–481). Amsterdam, Netherlands: Elsevier Science.

Fogel, R. W. (2005). Changes in the disparities in chronic diseases during the course of the 20th century. *Perspectives in Biology and Medicine, 48*(1 suppl), S150–S165.

Food and Agriculture Organization (FAO). (2009). *The state of food insecurity in the world 2009: Economic crises impacts and lessons learned.* Rome, Italy: Author.

Frankenberg, E., Thomas, D., & Beegle, K. (1999). *The real costs of Indonesia's economic crisis: Preliminary findings from the Indonesia Family Life Surveys.* RAND Labor and Population Program Working Paper Series 99-04. Los Angeles: RAND & University of California, Los Angeles.

Gallup, J. L., & Sachs, J. (2001). The economic burden of malaria. *American Journal of Tropical Medicine and Hygiene, 64*(1–2 suppl), 85–96.

Gardner, G. W., Edgerton, V. R., Senewiratne, B., Barnard, R. J., & Ohira, Y. (1975). Physical work capacity and metabolic stress in subjects with iron deficiency anemia. *American Journal of Clinical Nutrition, 30,* 910–917.

Haacker, M. (2004). *The macroeconomics of HIV/AIDS.* Washington, DC: International Monetary Fund.

Institute of Medicine. (2010). *Preventing the global epidemic of cardiovascular disease: Meeting the challenges in developing countries.* Washington, DC: National Academies Press.

Jamison, D. T., Lau, L. J., & Wang, J. (2005). Health's contribution to economic growth in an environment of partially endogenous technical progress. In G. Lopez-Casasnovas, B. Rivera, & L. Currais (Eds.), *Health and economic growth* (pp. 67–91). Cambridge, MA: MIT Press.

Jamison, D. T., Sachs, J., & Wang, J. (2001). *The effect of the AIDS epidemic on economic welfare in sub-Saharan Africa* (Working Paper WG 1). Geneva, Switzerland: World Health Organization, Commission on Macroeconomics and Health.

Kalemli-Ozcan, S. (2002). Does the mortality decline promote economic growth? *Journal of Economic Growth, 7,* 411–439.

Kaplan, G. A. (2012). Economic crises: Some thoughts on why, when and where they (might) matter for health: A tale of three countries. *Social Science and Medicine, 74,* 643–646.

Karlsson, M., Nilsson, T., Lyttkens, C. H., & Lesson, G. (2010). Income inequality and health: Importance of a cross-country perspective. *Social Science and Medicine, 70,* 875–885.

Keogh-Brown, M., & Smith, R. (2008). The economic impact of SARS: How does the reality match the predictions? *Health Policy*, 88(1), 110–120.

Keogh-Brown, M., Wren-Lewis, S., Edmunds, W. J., Beutels, P., & Smith, R. D. (2010). The possible macroeconomic impact on the UK of an influenza pandemic. *Health Economics, 19,* 1345–1360.

Kim, H.-J., & Ruger, J. (2010). Socioeconomic disparities in behavioral risk factors and health outcomes by gender in the Republic of Korea. *BMC Public Health, 10*(1), 195.

Kinra, S., Rameshwar Sarma, K. V., Ghafoorunissa, Mendu, V. V., Ravikumar, R. Mohan, V., . . . Ben-Shlomo, Y. (2008). Effect of integration of supplemental nutrition with public health programmes in pregnancy and early childhood on cardiovascular risk in rural Indian adolescents: Long term follow-up of Hyderabad nutrition trial. *British Medical Journal, 337,* 1468–5833. doi:10.1136/bmj.a605

Kleinman, A. M., Bloom, B. R., Saich, A., Mason, K. A., & Aulino, F. (2008). Introduction: Avian and pandemic influenza: A biosocial approach. *Journal of Infectious Diseases, 197,* S1–S3.

Knaul, F. M. (1999). *Linking health, nutrition and wages: The evolution of age at menarche and labor earnings among adult Mexican women* (Working Paper Series R-355). Washington, DC: Inter-American Development Bank.

Lakdawalla, D. N., Bhattacharya, J., & Goldman, D. P. (2004). Are the young becoming more disabled? *Health Affairs, 23,* 168–176.

Leeder, S., Raymond, S., Greenberg, H., Liu, H., & Esson, K. (2004). *A race against time: The challenge of cardiovascular disease in developing countries.* New York, NY: Trustees of Columbia University.

Lindeboom, M., Lundborg, P., & Van Der Klaauw, B. (2009). *Obesity and labor market outcomes: Evidence from the British NCDS.* IZA Discussion Paper No. 4099. Bonn, Germany: Institute of Labor Economics.

Lock, K., Smith, R. D., Dangour, A., Keogh-Brown, M., Pigatto, G., Hawkes, C., Fisjberg, R., & Chalabi, Z. (2010). Health, agricultural and economic effects of adopting healthy diet recommendations. *The Lancet, 376,* 1699–1709.

Lovász, E., & Schipp, B. (2009). The impact of HIV/AIDS on economic growth in sub-Saharan Africa. *South African Journal of Economics, 77,* 245–256.

Lund, C., De Silva, M., Plagerson, S., Cooper, S., Chisholm, D., Das, J., . . . & Patel, V. (2011). Poverty and mental disorders: Breaking the cycle in low-income and middle-income countries. *The Lancet, 378,* 1502–1514.

Malaney, P., Spielman, A., & Sachs, J. (2004). The malaria gap. *American Journal of Tropical Medicine and Hygiene, 71*(suppl 2), 141–146.

Marmot, M. G., Stansfeld, S., Patel, C., North, F., Head, J., White, I., . . . Davey Smith, G. (1991). Health inequalities among British civil servants: The Whitehall II study. *The Lancet, 337,* 1387–1393.

Marmot, M. G., & Theorell, T. (1988). Social class and cardiovascular disease: The contribution of work. *International Journal of Health Services, 18,* 659–674.

Martorell, R., Melgar, P., Maluccio, J. A., Stein, A. D., & Rivera, J. A. (2010). The nutrition intervention improved adult human capital and economic productivity. *Journal of Nutrition, 140,* 411–414.

McDonald, S. & Roberts, J. (2006). AIDS and economic growth: A human capital approach. *Journal of Development Economics, 80,* 228–250.

McIntosh, C. N., Finès, P., Wilkins, R., & Wolfson, M. C. (2009). Income disparities in health-adjusted life expectancy for Canadian adults, 1991 to 2001. *Health Reports, 20,* 55–64.

Nordhaus, W. (2003). The health of nations: The contribution of improved health to living standards. In K. H. Murphy & R. H. Topel (Eds.), *Measuring the gains from medical research: An economic approach* (pp. 9–40). Chicago, IL: University of Chicago Press.

Park, M. H., Falconer, C., Viner, R. M., & Kinra, S. (2012). The impact of childhood obesity on morbidity and mortality in adulthood: A systematic review. *Obesity Reviews, 13*(11), 985–1000. doi: 10.1111/j.1467-789X.2012.01015.x

Parmar, D., Stavropoulou, C., & Ioannidis, J. P. A. (2016). Health outcomes during the 2008 financial crisis in Europe: Systematic literature review. *British Medical Journal, 354,* i4588. http://dx.doi.org/10.1136/bmj. i4588

Preston, S. H. (1980). Causes and consequences of mortality decline in less developed countries during the twentieth century. In R. A. Easterlin (Ed.), *Population and economic change in developing countries* (pp. 289–360). Chicago, IL: University of Chicago Press.

Preston, S. H. (2007). The changing relation between mortality and level of economic development. *International Journal of Epidemiology, 36,* 484–490.

Preston, S. H., & Elo, I. T. (1995). Are educational differentials in adult mortality increasing in the United States? *Journal of Aging and Health, 7,* 476–496.

Pritchett, L., & Summers, L. (1996). Wealthier is healthier. *Journal of Human Resources, 31*(4), 841–868.

Rassy, D., & Smith, R. D. (2013). The economic impact of H1N1 on Mexico's tourist and pork sectors. *Health Economics, 22,* 824–834.

Richards, T. (2009). Governments must act now to prevent slide into poverty and ill health after recession. *British Medical Journal, 339,* b4087.

Schieber, G., & Maeda, A. (1997). A curmudgeon's guide to financing health care in developing countries. In G. J. Schieber (Ed.), *Innovations in health care financing: Proceedings of a World Bank conference, March 10–11, 1997* (World Bank Discussion Paper No. 365) (pp. 1–38). Washington, DC: World Bank.

Schultz, T. P. (2002). Wage gains associated with height as a form of health human capital. *American Economic Review, 92,* 349–353.

Shastry, G. K., & Weil, D. N. (2002). How much of cross-country income variation is explained by health? *Journal of the European Economic Association, 1,* 387–396.

Smith, R. D. (2006). Responding to global infectious disease outbreaks: Lessons from SARS on the role of risk perception, communication and management. *Social Science & Medicine, 63*(12), 3113–3123.

Smith, R. D. (2012). Why a macro-economic perspective is critical to the prevention of non-communicable disease. *Science, 337,* 1501–1503.

Smith, R. D., & Keogh-Brown, M. R. (2013). Macroeconomic impact of a mild influenza pandemic and associated policies in Thailand, South Africa and Uganda: A computable general equilibrium analysis. *Influenza and Other Respiratory Viruses, 7,* 1400–1408.

Smith, R. D., Keogh-Brown, M., & Barnett, A. (2011). Estimating the economic impact of pandemic influenza: An application of the computable general equilibrium model to the UK. *Social Science and Medicine, 73,* 235–244.

Smith, R. D., Keogh-Brown, M. R., Barnett, T., & Tait, J. (2009). The economy-wide impact of epidemic influenza on the UK: A computable general equilibrium modeling experiment. *British Medical Journal, 339*. doi:10.1136/bmj.b4571

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics, LXX,* 65–94.

Spurr, G. B. (1983). Nutritional status and physical work capacity. *Yearbook of Physical Anthropology, 26,* 1–35.

Stuckler, D., Basu, S., Suhrcke, M., Coutts, A., & McKee, M. (2009). The public health effect of economic crises and alternative policy responses in Europe: An empirical analysis. *The Lancet, 374,* 315–323.

Suhrcke, M., McKee, M., Arce, R. S., Tsolova, S., & Mortensen, J. (2006) Investment in health could be good for Europe's economies. *British Medical Journal, 333,* 1017–1019.

The Lancet. (2013). What can the UN General Assembly do for global health? *The Lancet, 382,* 1000. doi:10.1016/S0140-6736(13)61959-5

Thompson, J. E. (2007). Poverty, development, and women: Why should we care? *Journal of Obstetric, Gynecologic and Neonatal Nursing, 36*(6), 523–530.

Tsai, A. G., Williamson, D. F., & Glick, H. A. (2011). Direct medical cost of overweight and obesity in the USA: A quantitative systematic review. *Obesity Reviews, 12,* 50–61.

Tusting, L. S., Willey, B., Lucas, H., Thompson, J., Kafy, H. T., Smith, R. D., & Lindsay, S. W. (2010). Socioeconomic development as an intervention against malaria: A systematic review and meta-analysis. *The Lancet, 382,* 963–972.

Uutela, A. (2010). Economic crisis and mental health. *Current Opinion in Psychiatry, 23,* 127–130.

Vassall, A., & Martínez Álvarez, M. (2012). The health system and external financing. In R. D. Smith & K. Hanson (Eds.), *Health systems in low- and middle-income countries: An economic and policy perspective* (Chapter 9). Oxford, UK: Oxford University Press.

Veenstra, N., & Whiteside, A. (2005). Economic impact of HIV. *Best Practice and Research Clinical Obstetrics and Gynaecology, 19*(2), 197–210.

Wilder-Smith, A. (2006). The severe acute respiratory syndrome: Impact on travel and tourism. *Travel Medicine and Infectious Disease, 4*(2), 53–60.

Wilkinson, R. G. (Ed.). (1986). *Class and health: Research and longitudinal data.* New York, NY: Tavistock.

Wilkinson, R. G. (1990). Income distribution and mortality: A natural experiment. *Health and Illness, 12,* 391–412.

Wilkinson, R. G. (1996). *Unhealthy societies: The afflictions of inequality.* London, UK: Routledge.

Wilkinson, R. G., & Marmot, M. (2003). *Social determinants of health: The solid facts* (2nd ed.). Copenhagen, Denmark: World Health Organization.

World Bank. (2015). The socio-economic impacts of Ebola in Liberia. World Bank Group. Retrieved from http://documents.worldbank.org/curated/en/162381468179333776/pdf/96196-WP-P152547-PUBLIC-Box391443B-Liberia-Apr-13.pdf

World Bank. (2017). Health expenditures by income group and region, 2014. World Bank Group. Retrieved from https://datacatalog.worldbank.org

World Health Organization (WHO). (1999). *World health report 1999: Making a difference.* Geneva, Switzerland: Author.

Zhang, Y. X., & Wang, S. R. (2010). Differences in development among children and adolescents in eastern and western China. *Annals of Human Biology, 37,* 658–667.

CHAPTER 20

International Trade and Health

Johanna Hanefeld, Richard Smith, and Rupa Chanda

Introduction

Trade—the movement of goods, services, people, and capital between countries—has become established as a critical determinant of public health, both positively and negatively, that "directly and indirectly affects the health of the global population with an unrivalled reach and depth" (MacDonald & Horton, 2009). The last decade has seen the development of an overwhelming evidence base highlighting the connections between trade, trade agreements, and population health. Trade itself is seen as a defining feature of the current era of globalization, making trade and health a core aspect of the field of global health. For instance, increased trade in food, facilitated especially by refrigerated transport, has diversified many diets and improved access to fruit and vegetables. At the same time, dependency on global food markets, intensive farming methods, and cash crops have created concerns about food security and environmental degradation. Closely associated aspects of trade liberalization—policies to reduce the barriers to trade—and trade agreements—the legal basis for the rules that govern trade—have also come under increasing scrutiny. Those concerns have ranged from access to medicines to treat human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) in the 1990s, to the focus on the World Trade Organization (WTO) and increasing trade in services in the 2000s, to the emphasis on food security and alcohol availability in recent years.

As the raison d'être of trade is to increase both wealth and the availability of goods and services, changing trade patterns—whether through an increase or a reduction in trade—will inevitably impact many of the known determinants of health, including employment, nutrition, environmental factors, social capital, and education. For instance, increased trade in alcohol and tobacco products is likely to be harmful to health, and intensified competition for foreign investment may create pressures to reduce health and safety measures in the workplace. Trade will also affect the health sector itself, most clearly through direct trade in health-related goods and services (such as pharmaceuticals, health workers, foreign direct investment in health services, and patient mobility), but more broadly as well. For instance, trade liberalization can affect overall public expenditures by allowing governments to purchase less expensive foreign goods and services, which will in turn affect the amount of money available to fund public health care. Alternatively, it may reduce tariffs, and hence government income, and lead to a reduction in public spending on health care. Moreover, trade—especially rapid and widespread movement of people, animals, and goods—may facilitate the rapid and extensive spread of disease, a factor associated with outbreaks such as severe acute respiratory syndrome (SARS) and Ebola. In addition, trade has triggered changes in the structure of companies, with the 20th century especially seeing a huge increase in the number of transnational corporations with a global footprint rather than a

national one, a key aspect of "globalization" covered in the *Health and "The Economy"* chapter. For all these reasons, trade, and associated policies governing and responding to that trade, have become increasingly recognized as a critical driver of health issues.

It is important to note that the current era of globalization has been associated largely with a rapid increase and liberalization of trade and the regimes governing it (Dollar, 2001a, 2001b; Dollar & Kraay, 2004; Hanefeld, 2015). Trade may not be growing in some countries, however, and the populist politics evident in the 2016 election of Donald Trump to the U.S. presidency have been associated with a backlash against globalization and specifically trade (McKee, Greer, & Stuckler, 2017). Nevertheless, the vast majority of evidence focused on the link between trade and health that has emerged over the past two decades and is discussed in this chapter has concentrated on the increase and liberalization of trade.

This chapter provides an overview of these issues. We first look at the relationships between international trade, determinants of health, and health care itself, along with evidence of the impact on health and health equity. This discussion provides an indication of the changes and current levels of trade in key areas. The legal foundations, in the form of trade agreements, are outlined, focusing especially on the wider political and policy landscape and the ways in which this shapes global health governance. We next examine the different pathways through which trade affects health generally (not focusing on the health system beyond the effects of government revenue generated from trade). Attention is paid to how trade in goods, for example, affects communicable and noncommunicable diseases. We then focus on the impact of trade on health systems, including trade in health services and its likely current and future effects. The concluding section explores the response from the global health community to date, and raises likely future issues relating to trade and health. There is synergy between this chapter and the *Health and "The Economy"* chapter in terms of their focus on governance, negotiation, and diplomacy, especially in the concluding sections, and because of this they are best read in conjunction.

Relationship Between Trade and Health

Increased trade in goods, services, people, and capital—whether directly health related or not—affects health through a number of channels, including the cross-border flow of infectious disease, advertising of unhealthy lifestyles, health professional migration, and so forth. Health and health care are affected by changes in general trade liberalization, international legislation, and international institutions, as well as those changes specific to health; in turn, they affect national economies.

Building on the analysis presented in the *Health and "The Economy"* chapter, we can summarize these main proximal and distal determinants and linkages between trade and health as shown in **FIGURE 20-1**. The lower half of the figure represents the country under consideration, and the upper half indicates the aspects of international trade expansion and liberalization that affect the country, with the three arrows between them indicating the major linkages. (A more comprehensive exposition is provided by Woodward, Drager, Beaglehole, and Lipson [2001].)

Taking the lower half of the figure first, the standard influences on health are as follows:

- Risk factors, representing genetic predisposition to disease, environmental influences on health, infectious disease, and other factors.
- The household economy, representing factors associated with human capital and the investment in health by individuals and households.
- The health system, representing the impact of goods and services consumed principally to improve health status.
- The national economy, representing the meta-influences of government structures and general economic well-being.

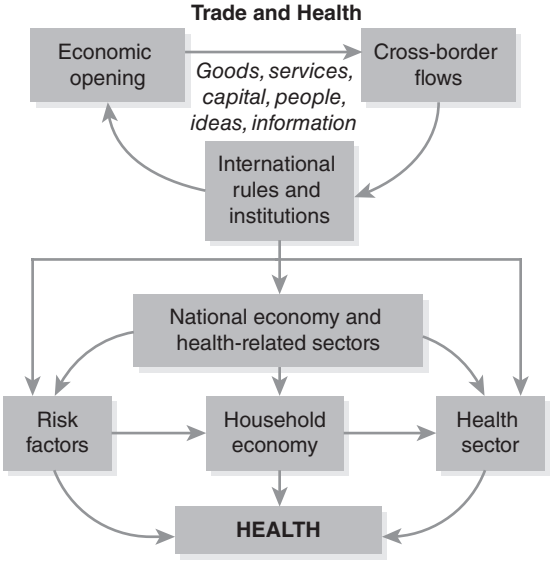


FIGURE 20-1 A simplified view of how trade affects health.

Modified from Smith, R. D., & Hanson, K. (Eds.). (2012). *Health systems in low-and middle-income countries: An economic and policy perspective*. Oxford, UK: Oxford University Press.

The range of interlinkages between these factors is also illustrated in Figure 20-1, many of which are covered in the chapters *Health and "The Economy"* and *The Design of Health Systems*.

The upper half of Figure 20-1 depicts the influences of factors outside the national economy, or national "control." For example, a wide variety of international influences directly affect risk factors for health, including increased exposure to infectious disease through the rapid cross-border transmission of communicable diseases via the movement of people, animals, and products (especially food products); increased marketing of unhealthy products and behaviors; and increased environmental degradation as a result of increased industrialization. Trade also affects health through its influences on the national economy. An extensive literature addresses the relationship between health and wealth: To the extent that trade influences economic growth, and the distribution of positive and negative aspects of this growth, it also influences health (e.g., changes in consumption of various goods that are either beneficial or detrimental to health) (Blouin, Chopra, & van der Hoeven, 2009; Sachs, 2002). Finally, trade affects health through the direct provision and distribution of health-related goods, services, and people, such as through access to pharmaceutical products, health-related knowledge and technology (e.g., new genomic developments), and the movement of patients and professionals.

Also in the upper half of Figure 20-1, we see the importance of international trade agreements in the governance of this system. This aspect of trade has gained particular prominence over the past decade with a newer generation of trade and investment agreements (TIAs) whose provisions have often limited the ability of national governments to introduce public health regulations (Walls, Smith, & Drahos, 2015).

Each of the elements is covered in the remaining sections of this chapter, but we begin by looking at the legal foundation for trade: trade agreements.

Trade Agreements Creating a Governance for Health and Trade

Trade is far from a modern phenomenon. The exchange of goods and services, along with the movement of people, predates the establishment of nation states, and indeed contributed to the development of what are now considered to be major cities such as Venice, New York, and Tokyo. Developments in transport and communications assisted trade in expanding in scope (e.g., the advent of refrigeration opened up the trade of perishable goods), scale (e.g., super-tankers and cargo ships increased the volumes that can be traded), and speed (e.g., air travel for goods, the Internet for financial transactions) over the centuries. Indeed, trade contributed as large a share of global gross domestic product (GDP) at the end of the 19th century as it did at the end of the 20th century. So, what happened?

The events of the early 20th century have important parallels today. That era saw a closing in of borders, especially by the United States, under a protectionist nationalist agenda. This, at least in part, generated the Great Depression of the 1930s, which has been identified as a major contributor to World War II. Following the war, a key component of the new United Nations system was the General Agreement on Tariffs and Trade (GATT), which was to lead to successive reductions in global tariffs and other barriers to trade. It established a global trade regime whose current incarnation is the WTO, which superseded GATT in 1995 and has a much expanded remit compared to its predecessor (Lee, Sridhar, & Patel, 2009). Although in recent years the focus has moved away from such global initiatives and toward more regional, multilateral, and bilateral trade and investment agreements, trade agreements remain the glue that holds trade together.

Despite the prevailing rhetoric, no trade is truly "free." Instead, it is organized through trade agreements of myriad kinds that provide the regulatory framework—or rules—under which trade is conducted. Trade agreements are important because they are at the core of the trade–health nexus. Critically, trade is structured and has a highly formalized governance system, whereas global health governance is almost completely unstructured, features a greater diversity of actors and approaches, and imposes weaker legal obligations on states. These differences help explain why the trade–health nexus is dominated by trade agreements rather than "health regulation" (Fidler, Drager, & Lee, 2009). This topic is covered in more detail in the *Health and "The Economy"* chapter.

The WTO is the center of authority for the governance of trade, as indicated by its large number of member states and the substantive reach of its agreements. The WTO provides the overarching legal and normative framework for trade between its 164 member states, and sets the tone for other agreements beyond its membership (**EXHIBIT 20-1**).

Multilateral and Bilateral Trade Agreements

Although much of the interest regarding trade and health stems from, and has related to, the WTO,

EXHIBIT 20-1 The World Trade Organization and Health

In an effort to regularize an increasingly complicated system of international trade, the WTO was created in 1995 as the successor to GATT, which had been drafted in 1947. The WTO is a global agency that sets the rules for international trade (including trade in services, goods, and intellectual property). In June 2017, 164 countries were members of the WTO. The aim of the WTO is to stimulate economic activity and promote economic development, through progressive trade liberalization. Its philosophical basis is that liberalization will encourage a global increase in efficiency, through the traditional economic arguments relating to comparative advantage, thereby ensuring continued product availability for consumers and reducing the economic power of individual economic operators (Chanda, 2002). The WTO seeks to achieve this goal by creating a credible, reliable, and binding system of international trade rules to ensure transparency, consistency, and predictability in international economic policies (Adlung & Carzaniga, 2001).

The WTO negotiates and implements new trade agreements, serves as a platform for trade negotiations, settles trade disputes, reviews national trade policies, assists low- and middle-income countries (LMICs), and cooperates with other international organizations. Its role has five aspects (Lee et al., 2009):

1. Provision of a forum for negotiations between WTO members about their multilateral trade relations in matters dealt with under WTO agreements
2. Administration of multilateral trade agreements
3. Promotion of the transparency of WTO members' trade policies and actions regarding the implementation of WTO obligations, through regular monitoring and surveillance
4. Provision of a process for WTO members to mediate and settle trade disputes
5. Working in cooperation with relevant international organizations to achieve greater coherence in global economic policy making

To become a WTO member, a country has to agree to accept 17 main multilateral agreements and 60 agreements, annexes, decisions, and understandings that contain binding obligations on, among other things, tariffs and nontariff barriers on industrial and agricultural goods; trade in all kinds of services; application of measures to protect human, animal, and plant health (sanitary and phytosanitary measures); implementation of technical barriers to trade; use of trade-related investment measures; imposition of additional tariffs on dumped or subsidized imports; and protection of intellectual property rights. The large number of WTO member states means that most of the international community has committed itself to implementing this vast array of obligations. Why? Members benefit because the WTO stimulates increased world trade by securing stability and predictability through its legally enforceable and indefinitely binding nature, although this binding nature has raised concerns (Pollock & Price, 2000).

A key element of WTO membership is that all countries commit to the "most favored nation" principle. This means that privileges, such as reductions in tariffs or market access, granted to one country must also apply to all other WTO members.

The Dispute Settlement Body is an important element of the WTO that highlights the recognition of the importance of health as a political and social objective. In various cases, the WTO Dispute Settlement Body has ruled that (1) the protection of health is a national objective of vital importance; (2) WTO members are free to select the level of health protection they believe appropriate for their populations; (3) health effects should be included in the case-by-case analysis of whether products or services are alike; and (4) in cases analyzing measures to protect human health under the necessity test, the potential effectiveness of less trade-restrictive alternatives should receive strict scrutiny. In this respect, public health has a much stronger profile in international trade law than the protection of human rights, which is not an objective that trade treaties recognize as a legitimate reason for restricting trade. In short, in implementing trade treaties, national policy makers and legislators have a foundation in the treaties on which to build coherency between trade and health.

The WTO has drafted several agreements that directly involve and include health and health systems. These include:

- The General Agreement on Trade in Services (GATS)
- The agreement on Technical Barriers to Trade (TBT), which requires that technical product regulations and standards be applied in a nondiscriminatory manner, be harmonized where possible on the basis of recognized international standards, and be the least trade-restrictive measures possible to achieve the level of health protection sought
- The agreement on Sanitary and Phytosanitary measures (SPS), which comprise national measures that aim to reduce hazards to animal, plant, and human health, including food safety regulations
- The agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) (Blouin, Drager, & Smith, 2006)

We will return to these agreements in the relevant sections later in the chapter.

Over the past decade, the WTO has seen a decline in its influence owing to an increase in regional and bilateral trade agreements and the lack of conclusion to the Doha trade round (Bown, 2017).

significant developments have also occurred in the form of regional and bilateral agreements. For instance, there has been a recent flurry of activity around the proposed Trans-Pacific Partnership (TPP) and the Transatlantic Trade and Investment Partnership (TTIP) between the United States and the European Union (Thow et al., 2014). These agreements have a much broader scope and wider provisions that might potentially limit national sovereignty, including constraints on the ability of national governments to regulate areas including public health (Labonté, Schram, & Ruckert, 2016; Ruckert et al., 2017). These agreements have engendered much opposition, including protests from the public health community and public advocates. Indeed, the controversy led to the TTIP talks being suspended in the second half of 2016. Further uncertainty over both the TPP and the TTIP emerged with the election of Donald Trump to the U.S. presidency. Trump campaigned on an openly anti–free trade platform, and one of his first actions in office was to cancel the TPP (Labonté, Schram, & Ruckert, 2017). Despite the apparent slowdown in implementation of some of the newer trade and investment agreements in the last few years, regional and bilateral trade agreements have generally regained importance over the past decade.

Often liberalization is more easily negotiated at the regional level, which tends to involve countries with geographic proximity, cultural ties, and similar levels of economic development. Regional and bilateral trade agreements also often involve a more limited number of participants and, therefore, allow for greater reciprocity, as well as often being undertaken for reasons other than purely economic motivations, such as strategic, cultural, and political reasons. At least in part, regional and bilateral agreements of recent years have also been pursued in response to the slow pace of WTO negotiations (Crawford & Fiorentino, 2004).

Indeed, since the creation of the WTO in 1995, member states have notified the organization of more than 400 additional regional trade agreements (RTAs). This is in addition to 124 RTAs that were in place when the organization was founded (WTO, 2017a). An estimated 55% of all international trade occurs through RTAs (Jones, 2007; Organisation for Economic Co-operation and Development [OECD], 2003). Within these kinds of regional or bilateral agreements, countries are typically obliged to ease or eliminate restrictions on entry of people and foreign investment, and to protect investors against adverse government regulation or confiscation. Multinational organizations dealing in tobacco, alcohol, and arms can challenge governmental regulations that hinder trade and potentially recover monetary compensation under RTAs. Indeed, these agreements often provide more predictable protection for multinational companies than either WTO law or regional legislation (Walls, Smith, & Drahos, 2015). In addition to the elimination of tariff and nontariff barriers, bilateral agreements facilitate the establishment of industry presence, which can further aid the marketing of foreign products such as alcohol and tobacco, and significantly reduce barriers to trade by breaking up protected markets, thereby contributing to increased consumption (Onzivu, 2002).

Regional, and increasingly bilateral, trade agreements do pose a danger of benefiting developed countries and being detrimental to LMICs. This discrepancy in impact arises because they create the potential for trade diversion, as the "most favored nation" principle (see Exhibit 20-1) does not apply in this sort of agreement. Often LMICs engage in regional and bilateral agreements in hopes of securing access to the market for their goods, without taking into account the impact that engaging in such agreements could have on their health systems (Mattoo & Fink, 2004). In addition, a significant body of evidence has emerged over the past decade suggesting that trade and investment agreements, including investor protection mechanisms, may allow corporations to seek to avoid public health regulations. Cases here include the court case brought by Philip Morris (which the company lost) against the Australian government's introduction of plain packaging for cigarettes (Hawkins & Holden, 2016).

The Impact of Trade on Health: Different Pathways

Impact of Trade on "The Economy"

Trade affects health through its general macroeconomic impact on countries, including government revenue, living standards, and types of employment and other determinants of health.

But why is trade encouraged in the first place? Simply speaking, because it promotes economic growth. Macroeconomic policy is concerned with GDP, which measures the value of all goods and services produced in a country during a given time period (see the *Health and "The Economy"* chapter). An increase in GDP represents economic growth. International trade is a key factor leading to economic growth through specialization, given that specialization allows each country to

exploit its absolute or comparative advantage in particular areas. Differences between countries in terms of their technology, labor skills, climate, institutions, and other factors mean that their production functions across different goods and services also differ (Maneschi, 1998). If each country specializes in the production of those goods and services that best suit its relative endowment of skills and resources, and if it trades these for other goods and services from countries that are relatively more efficient at producing those goods and services, all parties stand to benefit. In this way, specialization and international trade change relative prices and allow countries to produce and consume more than they would otherwise. Trade liberalization increases competition, which leads not only to lower prices but also increased specialization. Overall, those countries that engage in trade will see increasing GDP, lower prices, a wider selection of available goods and services, higher employment, and higher government revenues (due to higher income and the ability to tax that income) (Dollar, 2001b; Mattoo, Rathindran, & Subramanian, 2001; United Nations Conference on Trade and Development [UNCTAD], 1995).

Trade liberalization is about reducing barriers—financial or otherwise—to the movement of goods, services, capital, and people, with countries being aligned along a spectrum of what may be termed "openness" to trade (Smith, Lee, & Drager, 2009). As the raison d'être of trade is to increase both wealth and the availability of goods, changing trade patterns can influence both income (buying power) and its distribution. In turn, it may influence health through determinants including employment, nutrition, environmental factors, social capital, and education (Blouin et al., 2009).

General analyses suggest that "wealthier countries are healthier countries" (Bloom & Canning, 2000; Pritchett & Summers, 1996). The relevant factors in this relationship are generally improved nutrition, sanitation, water, and education (Becker & Tomes, 1986; Deaton, 2003; Fogel, 1994; Frankel & Rose, 2005; Hunt, 2007; Subbarao & Raney, 1995). The direction of causation is subject to some debate, although it is widely accepted that a "virtuous circle" exists between increasing wealth and health such that both directions of causation are valid (Suhrcke, McKee, Arce, Tsolova, & Mortensen, 2006). For instance, evidence shows a solid positive relation between a country's wealth and its population's health, defined in terms of increased life expectancy (Preston, 1975) or reduced child mortality (Pritchett & Summers, 1996). Importantly, open economies appear to grow more rapidly than restricted ones, such that the health gains from higher growth will accrue more quickly in those countries that engage in more international trade (Dollar & Kraay, 2004; Edwards, 1993; Frankel & Romer, 1999; Levine & Rothman, 2006; Winters, 2004; Winters, McCulloch, & McKay, 2004).

Although there is generally a positive relationship between trade liberalization and national income (Ben-David, 1993; Dollar, 1992; Feachem, 2001; Sachs, Warner, Åslund, & Fischer, 1995), there is increasing recognition, including among international financial institutions such as the World Bank, that trade liberalization in the absence of other policies will not necessarily lead to higher growth (Winters, 2004). In recent years especially, in the aftermath of the global financial crisis and subsequent recession from 2008, analysis has become more nuanced and less certain. Considerable evidence also indicates that poverty and income inequality are associated with poorer health (Deaton, 2003). Thus, although trade liberalization generally is advantageous, the crucial factor in how advantageous and to whom is "how countries manage the process of integrating into the global economies" (Lall, 2004). For example, trade liberalization often triggers both creation and destruction of jobs within and across sectors, as firms adjust to the new competitive environment (Ghose, 2003). To ensure that the movement of labor is as smooth as possible and to avoid excessive unemployment, effective labor market policies are required (Hoekman, 2005; Milanovic & Squire, 2005; Mundial, 2005; Winters, 2004; Wood & Ridao-Cano, 1999).

Distributional Impacts Are Critical for Health

It is important to remember that analyses of trade and the economy relate to *aggregate* populations and incomes, rather than to individuals (Ravallion, 2006). Although evidence supports the proposition that trade liberalization will be poverty alleviating in the long run and on average, it is important to remember that this outcome implies adjustment and distributional impacts and that trade reform is not always the most important determinant of poverty reduction (Winters et al., 2004). Most macroeconomic indicators do not account for these distributional impacts; that is, they focus on the aggregate of total income, trade volume, employment, and so on, rather than the composition of that total.

Expanding the availability of goods, reducing prices, and increasing access to high-speed communication and travel may benefit those who produce these goods and services and the wealthy. For others, these trends may mean lower wages, greater pollution, and insecurity. Much of the benefits of trade "trickle-down" may be dammed at specific points by those with power and influence. Consider the example of food security:

Between 2006 and 2008, the cost of food overall increased by 57%, generating major crises in many LMICs as people struggled to buy even the basic foodstuffs. Although several factors were associated with this increase, including devaluation of the U.S. dollar, rising oil prices, and increased demand from emerging economies, a significant factor was increased support for biofuels. Biofuel development was promoted primarily by the United States and European countries as a cost-effective means to meet greenhouse gas emissions targets as well as to support their own embryonic biofuel industries, which were subject to sizable government subsidies (Hallam, 2009). But it also meant that food crops were being diverted to produce ethanol. For example, Brazil took this course and became a major producer of ethanol-producing crops over this period (Tenenbaum, 2008).

Lowering tariffs (taxes on imports) may generally benefit consumers, who can purchase imported goods at lower prices, but it may also create losses in competing domestic firms and reduce import tax revenues. Trade reforms inevitably create both winners and losers: Some sectors of the economy might not be able to compete with new imported goods, whereas others get access to new markets and opportunities. Moreover, individual incomes can change because jobs can be created or lost and prices of and external demand for goods produced can rise or fall. Therefore, some losers from trade liberalization might be poor households, whose incomes will decline even further (Winters, 2006). Losses to domestic firms can, in turn, negatively impact the employees and other stakeholders of the firms, such as via lost jobs, reduced wages (Arbache, Dickerson, & Green, 2004), or reduced revenues from both corporate and personal income taxes. For these reasons, countries are often careful in opening sectors to international competition. Depending on the economic geography of a country, some regions might benefit more than others from liberalized trade. For instance, it might be the case that export goods are produced in urban areas. Increased income from trade accruing to these urban areas may benefit those areas and surrounding rural areas (through higher demand due to rising incomes), but may not reach areas farther away. However, if managed wisely by the government, the new income opportunities could, in principle, benefit the entire population.

Thus, although in aggregate trade liberalization may generate economic benefits for a country, it is also likely to require adjustment at various levels, including at the individual level. This adjustment process can cause a feeling of increased economic insecurity with implications for psychological well-being (Cornia, 2001; Marmot, 2000; Scheve & Slaughter, 2004).

Effect of Trade on Government Revenue for Health Care

Trade liberalization will also affect government income and hence the ability to finance or provide public services, including those related to health. Trade directly affects government revenues via taxes (tariffs) imposed on imports to the country. However, trade can also indirectly affect government revenues via income taxes paid by local businesses and their employees, although these streams are harder to trace. Thus, establishing the net contribution of trade to government income—where gains in tax revenues traceable to increased export income exceed losses in tax revenues due to increased imports—is difficult. Nonetheless, understanding the potential impacts of trade on the different types of taxes is important because certain types of tax income are relatively easy to collect (such as tariffs), whereas other types may be harder to collect (such as consumption taxes, income taxes, and value-added taxes) (Aizenman & Jinjarak, 2009).

In many LMICs, tariffs on imports are an important source of public revenues and contribute to the capacity of these countries to adopt policies that affect the various determinants of health. The contribution of tariff revenues to total government revenues may be less than 1% for members of the OECD, but is as much as 80% in Guinea, with more typical examples including Cameroon at 28% and India at 18% (Laird, Vanzetti, & Fernandez de Cordoba, 2005). The concern in such cases is that trade liberalization may change the tax revenue structure, reducing the proportion of government income from easy-to-collect sources such as tariffs. Although governments should theoretically be able to shift tax bases from tariffs to domestic taxes, such as sales or income taxes, in practice many (especially low-income) countries find this difficult due to the informal nature of their economies, which often have large subsistence sectors (Aizenman & Jinjarak, 2009). For instance, high-income countries (HICs) are usually able to recover 100% of the tariff revenues lost through trade liberalization, while middle-income countries accrue approximately 40% to 60%, and low-income countries can gather only 30% or so (Baunsgaard & Keen, 2010; Blouin et al., 2009; Glenday, 2006; Khattry & Rao, 2002). Ultimately, the reduced import tariffs in the latter cases result in a decline of government income available to pursue public policies, whether related to health care, education, water, sanitation, or social safety nets.

Conversely, lower import prices might benefit the public provision of health care, as inputs into health services become less expensive and more widely

available. Furthermore, consumers are likely to benefit from lower prices, which enable them to spend more on health care, education, and other beneficial services.

Finally, it is important to note that lower tariffs on imports are not necessarily correlated with lower total revenues (Glenday, 2002). In fact, trade liberalization can increase the volume of international trade such that the base expansion may exceed the rate reduction and hence yield higher revenues, at least where tariffs are reduced rather than eliminated. For instance, halving import tariffs may be entirely offset if trade doubles. This appears to have been the case in a few countries at the early stage of implementation of World Bank trade reform programs (Bamou, 1999). Import tariffs can also have complex structures. Lower average duty rates are not necessarily the result of across-the-board rate reductions. For example, high duty rates may be lowered, while low duty rates may be raised. Interestingly, the impact of liberalization on tariffs may be an argument for increased multilateral, rather than bilateral, trade negotiations, as the latter often involve more severe tariff concessions than the former.

Trade and the Risk Factors for Disease

Trade also affects health through its direct impacts on risk factors for disease. These effects occur as a result of the pathogens that travel with goods, including livestock and plants traded, and through the commodities traded and their health effects—for example, medicines, alcohol, food, or tobacco. Trade significantly affects the spread of disease either by acting as a "vector" for pathogens, or by increasing or limiting the availability of goods that affect health. Some of the concerns relating specifically to the spread of infectious disease are further elaborated on in the *Public Health Infrastructure* chapter.

Trade and Communicable Disease

As indicated elsewhere (see the *Environmental and Occupational Health* chapter), concerns have arisen that changes in climate will facilitate the spread of vectors for infectious disease, especially mosquitoes. Trade may play a role in this relationship through increased production, consumption, and transportation of goods globally contributing to climate change. This is a much wider issue than trade itself, and the interested reader is referred to the *Environmental and Occupational Health* chapter for more details. Here, the focus is on the more direct relationship between trade and infectious disease.

The most worrisome scenario remains the impact of trade—in terms of the rapid and widespread movement of people, animals, and goods—on possible pandemic disease. The prospect of an infection emerging with the transmissibility of the common cold but with a much higher pathogenicity is the public health community's worst nightmare. This fear was stoked especially by the SARS outbreak in 2002, with emergence since then of possible pandemic influenza strains raising visions of the Spanish flu, which killed an estimated 40 million people in 1918–1919, and the recent outbreaks of SARS and Ebola (**EXHIBIT 20-2**). Minor changes to the influenza virus (antigenic drift) happen continually over time, which is why it can reinfect populations over and over again, but this change is usually minimal such that vaccine created from the previous year can offer protection. Perhaps every decade or two, a major change in the virus (antigenic shift) occurs, against which most people have little or no immunity. The Spanish flu was one such event, as was H1N1 ("swine flu") in 2009. The danger arising from antigenic shift may be exacerbated by increased trade: Once strains emerge, they can travel worldwide at great speed using animals, people, and even goods as vectors. Whereas a century ago spread of the new strains may have taken months, and by 1918 took weeks, it now takes only a matter of hours.

It is, of course, of major importance to public health to have a system that can respond to infectious disease outbreaks. Notably, the rapid spread of cholera associated with increased global trade in the 19th century led to a considerable death toll; while governments imposed containment measures to control the disease, the economic interests of the mercantile elite suffered as these measures slowed the circulation of people and goods that was so crucial to commercial profit. These economic interests prompted the assembly of 12 governments in Paris in 1851 to establish a basis for common agreements on quarantine and the containment of cholera, leading eventually to the first International Sanitary Convention (ISC) in 1892. Other international treaties in the following decades ultimately resulted in the International Health Regulations (IHR), which were adopted at the fourth World Health Assembly (WHA) in 1951.

The IHR have historically coordinated the public health response to infectious disease threats, although countries were originally obliged to report only outbreaks of cholera, plague, and yellow fever (Liverani & Coker, 2012). However, compliance has been notoriously weak because of the economic consequences of notification, in the form of trade barriers. In short, the IHR have lacked the means and incentives to work effectively. Recognition of this problem prompted

EXHIBIT 20-2 SARS and Ebola

Severe Acute Respiratory Syndrome

The outbreak of SARS in 2002–2003 demonstrated how the modern world can be more vulnerable to disease spread. SARS is a respiratory illness caused by previously unknown type of coronavirus. Coronaviruses usually cause mild to moderate upper respiratory tract symptoms, but SARS led to high fever, cough, shortness of breath, and eventually severe pneumonia or respiratory failure that could be fatal. Between November 2002 and July 2003, there were 8,096 cases of SARS, causing 774 deaths in 27 countries. Although the disease burden ultimately proved to be low (in comparison, seasonal influenza causes 250,000 to 500,000 deaths each year worldwide), the outbreak caused alarm due to its unprecedented speed and extent of global spread. Its airborne transmission, the lack of diagnostic technology, and the lack of a vaccine, combined with this rapid dissemination, led to mass concern and widespread measures to restrict movement, including WHO's unprecedented step in issuing an advisory against travel to Toronto. It is believed that air travelers led to the spread of SARS from China, where it originated, to 16 countries (Bonn, 2003). Whereas the plague may have taken 3 years to cross from Asia to Europe in the Middle Ages, SARS took just 15 hours to get from Hong Kong to Toronto.

As outlined in the *Health and "The Economy"* chapter, the overall global macroeconomic burden of the outbreak was between $30 billion and $100 billion—a sum vastly out of proportion to the health impact (Keogh-Brown & Smith, 2008). The Canadian Public Health Agency suggested SARS costs included losses of $350 million from tourism, $220 million from reduced airport activity, and $380 million from non-tourism retail sales (Naylor, 2003). Overall estimates of the direct and indirect costs to Canada of SARS were placed at as much as $2 billion (Keogh-Brown & Smith, 2008).

Costs were distributed across a wide range of countries, some of which had no confirmed cases of SARS, illustrating the vulnerability of the global economy to a public health threat. Had the outbreak been more serious, more widespread, or of a longer duration, it is possible that quarantines and greater restrictions on movement of people and goods could have led to the closure of factories or worse in many areas. Moreover, this outbreak served as an illustration of the shared interests globally in ensuing effective surveillance and management of outbreaks.

Ebola

The recent outbreak of Ebola virus disease in West Africa during 2013–2016 confirmed the devastating link between trade and the spread of disease, as well as the effects of infectious disease outbreaks on trade. The lack of adequate surveillance and countries' slow initial responses, paired with the movement of people, including traders, allowed the initial outbreak to spread from Guinea to Liberia and Sierra Leone (Moon et al., 2015). In reaction to the outbreak, some neighboring countries closed their borders and sought to suspend trade against the explicit advice of WHO. The human costs aside, the World Bank estimated that the outbreak cost the three countries most affected—Liberia, Guinea, and Sierra Leone—approximately $2 billion.

substantial review and revision starting in 1995 to widen the IHR's remit beyond the three diseases and to address issues of means and incentives for compliance (WHO, 2002). The SARS outbreak provided the impetus for the final push, and the final revision was endorsed at the 2005 WHA and came into effect in June 2007. Under these revised IHR, countries are obligated to report any "public health emergency of international concern"—basically any adverse health event that has the potential to spread beyond borders—and allows notification from nongovernmental as well as governmental sources.

Yet, while the scope of IHR was expanded, incentive and repercussion issues remain: The suggestion to integrate some form of reparation, financial support, or other measures in relation to the timing and extent of possible barriers to cross-border movement from affected countries was not implemented. The economic imperative to protect trade stemming from the ISC in 1892 also remains, as the IHR are to "prevent, protect against, control and provide a public health response to the international spread of disease in ways that are commensurate with and restricted to public health risks, and which *avoid unnecessary interference with international traffic and trade*" (emphasis added). Some of these issues have since been highlighted by the H1N1 and Ebola outbreaks. Most critical of these is the ability of member states to impose trade and travel restrictions against countries experiencing outbreaks, to an extent that might be excessive and harmful to the latter's economies. In 2009, for example, the H1N1 outbreak led several countries to ban live pig imports or even pork products that originated in Mexico, despite the absence of scientific evidence that humans might contract the virus from animals or animal products. Under competing international law from WTO, the burden of proof is on health science to show that there is no chance that humans could catch the disease—that is, the burden to provide evidence is always with the health authorities and countries may adopt a

"precautionary principle" of "if in doubt, ban." Thus, there is no system of governance to regulate the economic impact of reporting or experiencing outbreaks, and hence no incentive to report them, fundamentally undermining surveillance and control strategies. The recent outbreak of Ebola has led to much greater emphasis on the importance of an effective global governance to respond to disease outbreaks. Further discussion of these approaches and implications is found in the *Public Health Infrastructure* and *Global Health Governance and Diplomacy* chapters.

Movement of goods, including livestock and food, has also important implications for health in two other ways. First, the noncommunicable disease (NCD) epidemic has highlighted the link between obesity and trade. In this case, recent research has focused on the role of trade in shifting patterns of consumptions as less healthy, high-fat and -calorie, processed foods, for example, undercut the price of local fresh produce (Friel et al., 2013; Hawkes & Thow, 2008; Thow, 2009; Thow & Hawkes, 2009). Second, the transmission of disease through livestock and food has drawn attention. This area is regulated by the WTO Agreement on the Application of Sanitary and Phytosanitary Measures (SPS Agreement), which states that countries have the right to protect human, animal, and plant health on the basis of scientific evidence; in other words, countries are allowed to ban products that are thought to endanger health. Similarly, the Codex Alimentarius of the Food and Agriculture Organization (FAO) is a set of international standards to ensure that food products are safe and can be traded. Recent attention has also focused on the role of trade in livestock and food products in creating AMR (Hanefeld, Khan, Tomson, & Smith, 2017).

Trade Liberalization and Noncommunicable Diseases

The *Chronic Diseases and Risks* chapter highlights the economic impact of NCDs. Trade—including not just movement of goods, animals, and people, but also foreign investment and marketing—has important implications for the spread and control of NCDs. Trade liberalization, through reduced tariff and nontariff barriers, can be beneficial when it leads to an expansion in the consumption of the goods that improve health, such as clean water, safe food, and education. However, it also facilitates the increased consumption of goods that may be harmful or hazardous to health, such as tobacco, alcohol, and unhealthy food items. Furthermore, increased trade in high-technology products such as mobile phones and computers increases the volume of toxic products that cross borders, and ultimately leads to an increase in toxic wastes as obsolete products are discarded. This trend may directly impact the health of those who handle such products and waste, and indirectly affect health via environmental degradation. In this section, we consider each of the four core areas where trade and risk of NCDs intersect.

Trade and Tobacco. Strengthening of tobacco controls in Western Europe and North America, coupled with LMICs' reductions of their trade barriers and expanded advertising by transnational tobacco companies, has shifted the market share for tobacco products toward the emerging economies across Asia, Latin America, the Middle East, and Eastern Europe (Chaloupka & Corbett, 1998; Grise, 1990; Hagihara & Takeshita, 1995; Hsieh, Hu, & Lin, 1999). Given the inverse relationship between price and consumption, as well as the positive relationship between advertising and demand, rates of cigarette smoking and tobacco use can be expected to increase as tobacco markets become more open, all else equal. In turn, the mortality and morbidity attributable to tobacco use can also be expected to increase.

As early as the 1980s, the United States used bilateral trade agreements to exert pressure on countries such as Thailand, Taiwan, and South Korea to open their domestic economies to U.S. cigarette imports. In fact, U.S. market share in these countries was approximately 600% higher, on average, in 1991 than it would have been had these markets remained closed (Chaloupka & Laixuthai, 1996). Furthermore, opening of the cigarette markets in these countries led to a significant increase in rates of cigarette smoking, with per capita cigarette consumption being 10% higher by 1991, on average, than it would have been in the absence of the bilateral agreements (Chaloupka & Laixuthai, 1996). The Uruguay Round of GATT, which concluded in 1994, focused on liberalizing trade in "raw" tobacco; subsequently, 1994–1997 saw a 12.5% increase in nonmanufactured tobacco exports globally, after a decade of virtually no growth (Chaloupka & Warner, 2000).

Since the WTO replaced GATT in 1995, it has significantly reduced tariff and nontariff barriers to tobacco trade. In response to "plain packaging" requirements and minimum sizes for health warnings, which some condemn as unjustifiably encumbering trademark rights (Bettcher et al., 2003; Callard, Chitanondh, & Weissman, 2001), pressure has been brought to bear in manufacturers' home countries to include

EXHIBIT 20-3 Regulating Trade in Tobacco: The Framework Convention on Tobacco Control

The FCTC, which came into force in 2005, requires contracting parties to do the following:

- Implement public health protection measures by addressing issues such as tobacco advertising, promotion, sponsorship, packaging, and labeling.
- Improve regulation of the domestic tobacco industry, specifically by requiring disclosure of the contents of tobacco products and tobacco smoke.
- Address such issues as illicit trade in tobacco products, price and tax measures, sales of tobacco to and by young persons, government support for tobacco manufacturing, treatment of tobacco dependence, passive smoking and smoke-free environments, surveillance, research and exchange of information, and scientific, technical, and legal cooperation.

Its final text contains provisions that have implications for trade rules, with potential opportunities as well as challenges for domestic tobacco control. The first clause of the preamble to the FCTC provides that parties to the convention are determined to give priority to the right to protect public health. Consistent with the free trade principle allowing the imposition of nondiscriminatory taxes, the FCTC requires parties to prohibit or restrict, as appropriate, the tax-free and duty-free sale of tobacco. The FCTC's regulation of advertising, packaging, and labeling of cigarettes is relevant to trade law. For instance, cancellation (or prohibition of registration) of the terms "light" and "mild" should not contravene the relevant provisions of international intellectual property law, especially the TRIPS Agreement (McGrady, 2004). FCTC Article 9 provides for the development of international standards for the testing and measuring of the contents and emissions of tobacco products, which can be recognized under the WTO TBT agreement, thereby helping to advance tobacco control within the trade agenda (Shaffer, Brenner, & Houston, 2005). However, public health advocates must continue to monitor and advocate transparency in the implementation of the FCTC, as measures to implement the FCTC could be affected by other trade obligations, especially through regional, multilateral, and bilateral trade agreements.

tobacco in negotiations to liberalize the agricultural sector and to reduce tobacco control measures under the terms of trade agreements.

Tobacco companies have also sought to use provisions within trade agreements to fight national tobacco control policies. An example is the use of intellectual property rights relating to trademarks to contest plain packaging for cigarettes—for example, in Australia (Gleeson & Friel, 2013; Thomas & Gostin, 2013). In such cases, trade agreements provided the opportunity for companies to contest the issues in a national court. Another avenue is to contest public health regulation through trade agreements, and to report countries to the WTO dispute settlement mechanism or to bilaterally threaten or impose sanctions on them. Again, these mechanisms have been employed by transnational tobacco companies (Alemanno & Bonadio, 2011).

Such moves have meant that the regulation of trade in tobacco has been a high priority for the public health community for many years. Two competing views have emerged regarding the best way to regulate tobacco under international trade rules. First, tobacco may be excluded from schedules of liberalization and treated as a special product for stricter regulation or even prohibition (Weissman, 2003). Second, tobacco may be regulated like any other good, in accordance with international treaties such as those of the WTO (Bettcher et al., 2003; McGrady, 2007). In the end, the latter approach was adopted, through the WHO Framework Convention on Tobacco Control (FCTC),[1] described in **EXHIBIT 20-3**.

Trade and Alcohol. Excess consumption of alcoholic beverages has many detrimental effects on health, and various proposals have been made to prevent excess consumption, including imposition of a minimum legal purchasing age, restrictions on hours or days of alcohol sale, outlet density restrictions, public service educational campaigns, and alcohol taxes. Of these measures, a common call is to tax alcoholic beverages so that they become more expensive and hence less accessible. However, the "national treatment" provisions embodied in many trade treaties make it clear that governments must be careful not to tax products, including alcohol, in ways that (even unintentionally or incidentally) protect domestic production from foreign competition. Thus, although governments may appear to have the option to discourage alcohol consumption by imposing taxes, they face challenges in doing so in practice.

[1] http://www.who.int/fctc/en/index.html

An alternative is to regulate the time, place, and manner in which alcohol is sold. For example, as a general matter it is permissible for a government to restrict the hours or days during which alcohol can be sold, require retail businesses to obtain licenses (authorization) before selling alcohol, or require retail businesses to demand proof of age from those seeking to purchase alcohol. Again, policy makers must be careful to design such measures so as to avoid discrimination, indirectly or inadvertently, based on the national origin of the product.

The regulation of advertising is another way for governments to prevent harm caused by alcohol. In 2001, the European Court of Justice (preliminary ruling C-405/98) ruled that Sweden's law limiting alcohol advertising affected foreign alcohol products more adversely than more familiar domestic products. However, the Court also held that the ban could be justified if it were proportionate to the public health objective and not a disguised restriction on trade; thus, it allowed the case to be further processed within the Swedish court system. Sweden's domestic court ultimately upheld the ban, finding that it was indeed "proportionate" to its purpose (Center for International Environmental Law, 2003). This example illustrates that trade treaties can accommodate appropriately designed policies intended to further public health objectives.

The issues with strongest relevance for alcohol policies within the WTO system are now covered by the discussions under GATS. These negotiations do not cover the products themselves, but rather the services related to the products (such as sales and marketing). For example, GATS provisions provide a framework for how, when, and why governments can impose regulations on alcohol licensing procedures, state monopolies, the number of suppliers or outlets, and advertising and promotion activities. Alcohol is treated like any other good under WTO law, and any regulatory action must meet the nondiscriminatory principles. However, recognition of the harmful effects of trade in alcohol has led to the inclusion in some RTAs of provisions that allow some restrictions on alcohol trade. For example, Association of Southeast Asian Nations (ASEAN) trade agreements have enabled the group's members to exclude alcohol and guns from the list of goods subject to trade liberalization (Onzivu, 2006).

Trade, Food, and Agriculture. Trade liberalization has the potential to influence food-related nutrition and health issues through a series of pathways, and measures designed to liberalize trade influence the entire food supply chain. These measures include those to reduce financial and regulatory barriers to food imports and exports across national borders; harmonize or remove national food-related regulations; remove limits on the percentage of domestic companies that can be owned by foreign businesses; decrease government support to domestic food production; and support the development of infrastructure and capacity for trade and investment, such as transportation routes and storage facilities. Such efforts create changes along the food supply chain that influence the environment in which consumers make food choices, such as the availability of foodstuffs (amount, type, and nutritional quality), the safety of that food, what it costs, and how it is marketed. These factors, which represent important components of national and household food security, influence the choices people make about the food they eat. Thus, they affect the diets of consumers and, therefore, the prevalence of malnutrition, diet-related NCDs, and foodborne illness (De Vogli, Kouvonen, & Gimeno, 2014; Popkin, 2001; Stuckler, McKee, Ebrahim, & Basu, 2012). In addition, trade liberalization may affect human nutrition and health through a range of indirect effects—for example, the inadvertent entry of emerging human, animal, and plant diseases via the food supply, as well as effects on income, employment, and industry restructuring in response to trade liberalization.

Similar to tobacco and alcohol, food has seen major shifts in patterns of trade across the world. According to the OECD, the share of global agri-food trade between countries with RTAs rose from 20% to 40% between 1998 and 2009 (Bureau & Jean, 2013). Between 1980–1981 and 2006, world agricultural trade rose from $243 billion to $945 billion in real terms (Peterson, 2009). Between 1970 and 2001, gross world food imports, measured in terms of calorie equivalents, grew by almost 60%. The share of agricultural production that is exported increased from 19% in 1971 to 40% in 2003 (Bloom, Canning, & Sevilla, 2001). Exports from high-income OECD countries—which are responsible for the vast majority of trade in processed products—more than doubled, increasing from $169 billion in 1995 to $363 billion in real terms in 2008. Exports from LMICs increased even more, tripling and even quadrupling during this time.

While increased trade can make energy- and nutrient-rich food more widely available, with consequent benefits for the undernourished, this outcome also carries risks associated with over-consumption. Reductions in the prices of unhealthy foods (i.e.,

foods that are calorie rich, nutrient poor, and/or high in saturated fats and salt) compared with healthy foods, increased desirability and availability of unhealthy foods, worsening asymmetry between consumers and suppliers of foodstuffs, and growing urbanization and changes in lifestyle are all possible means by which trade liberalization could affect popular diets, especially those of poor populations. For instance, changes in trade policy in Central America have been directly linked to changes in the availability of meat, dairy products, processed foods, and temperate (imported) fruits (Thow & Hawkes, 2009). Similar trends have been observed with the lowering of trade barriers between Mexico and the United States following the signing of the North American Free Trade Agreement (NAFTA), after which imports of corn, soybeans, sugar, snack foods, and meat products into Mexico increased significantly (Clark, Hawkes, Murphy, Hansen-Kuhn, & Wallinga, 2012). Such an impact has been especially documented in the Pacific Islands, where several studies have noted displacement of traditional diets with high-fat imported foodstuffs and a concomitant increase in obesity rates and chronic diseases (Cassels, 2006; Evans, Sinclair, Fusimalohi, & Liava'a, 2001; Hughes & Lawrence, 2005; Thow & Snowdon, 2010).

Although the effects of the liberalization of food trade on food safety, availability (volume and variety), prices, quality, and marketing are not straightforward, and depend on the nature of the implementing legislation and other contextual factors (Thow, 2009), trade clearly plays a major role in influencing food prices. Prices are important because they affect incentives for food production and consumption. In practice, the effects of trade liberalization have varied, but trade has generally led to a lowering of the relative cost of energy-dense foods and diets (Drewnowski, Hanks, & Smith, 2010).

Finally, increased trade in advertising and other telecommunications services facilitates the commercial promotion of highly processed foods (Nestle, 2006). The ability to advertise and promote is a major "pull" factor for inward investment by large transnational food corporations into LMICs. Processed foods are commonly advertised and marketed all over the world, using a wide range of communication channels and marketing techniques. Estimates from Asia suggest that food accounts for a significant proportion of child-targeted advertising, ranging from 25% in the Republic of Korea to 70% in Malaysia (Escalante de Cruz, Phillips, Visch, & Bulan Saunders, 2004). Studies in Latin America suggest a high proportion of advertisements during children's programming are for processed foods, such as sweetened beverages, candy, sugar-sweetened cereals, and chips (Chopra, Galbraith, & Darnton-Hill, 2002).

Trade and Toxic Substances. Trade in toxic substances (e.g., heavy metals, persistent organic pollutants, and other substances that may have adverse health effects) has a long and well-documented history (Agege, 1985; Broad, 1981; Smith, 1980). In an unintended side effect, it is the stricter environmental protection in the industrialized countries that is contributing to the buildup of these hazardous wastes in the developing world. In the 1990s, governments in the European Union, Japan, and some states of the United States set up recycling systems for hazardous substances, but many did not have the capacity to deal with the sheer quantity of material collected or with its hazardous nature—in particular, the increased waste consisting of electronic devices (e-waste). Those devices often comprise a complex mixture of several hundred materials, many of which contain heavy metals such as lead, mercury, cadmium, and beryllium, as well as hazardous chemicals (Singhal, 2005).

The prospect of earning foreign currency in return for importing international waste is attractive to LMICs. Unfortunately, these countries often have neither the technical expertise nor adequate facilities for safely recycling or disposing of such waste. Substandard facilities contribute to the improper handling of hazardous waste in these countries. In addition, many employees at recycling facilities lack adequate protective equipment. As a result, workers in the recycling industry may develop a variety of health problems. For example, laborers who melt down car batteries may develop lead poisoning. When e-waste or polyvinyl chloride (PVC) is burned, employees are exposed to cancer-causing dioxins and other toxic chemicals. Adverse health effects attributable to unsafe handling of hazardous wastes include mercury poisoning, birth defects and miscarriages, kidney disease, cancer, and even death. Exports of e-waste to LMICs are likely to increase unless greater support is given to initiatives that ensure that companies take greater responsibility for the recycling, reuse, or disposal of potentially hazardous electronic products.

Trade and its effects on health—especially NCDs—specifically through trade in products has been facilitated to an extent by financial liberalization and foreign investment. While liberalization of financial markets lies beyond the scope of this chapter, some of

EXHIBIT 20-4 Facilitating Trade Liberalization: Foreign Investment

The expansion of trade has been facilitated by concomitant liberalization of financial markets and regulations related to advertising and promotion. Foreign direct investment (FDI) is an investment by an enterprise from one country into an entity in another, in which the parent firm owns a substantial, but not necessarily majority or controlling, interest. The foreign enterprise becomes an affiliate of the parent company, thereby creating or joining a transnational corporation (TNC). FDI is one of the mechanisms through which TNCs enter new markets, and it reflects a company's intention to remain invested in the market over the longer term. It has played an unprecedented role in recent decades as a source of capital for economic growth and development in many LMICs, facilitated by increased opening of global capital markets since the 1990s.

The global financial crisis of 2008 saw a sharp decline in FDI, from approximately $2 trillion in 2007 to $1 trillion in 2009 overall, although the drop-off was less marked in emerging economies. These trends are important for incidence and prevalence of NCDs, as FDI often takes the form of strong investments in tobacco, unhealthy food, and alcohol in emerging markets. For example, 10 of the top 100 TNCs (ranked by value of foreign assets) manufacture tobacco, food, or alcohol, as do a large proportion of the largest affiliates of foreign TNCs in emerging economies. There are also wider associated investments in other areas that may be hazardous for health, such as the disposal and recycling of hazardous waste.

Clearly, substantial FDI can bring much-needed capital, skills, technology, and other benefits to local economies, encouraging further investment and providing support for infrastructure developments, increased foreign exchange earnings, and a general lift to the host economy. However, they may also lead to reduced health and safety standards, or compromise policies to protect public health. For example, FDI by TNCs involved in food processing may exert pressure against enhanced food labeling, development of nutritional standards, or costly improvements in production methods. With respect to tobacco, a study showed that British American Tobacco and Philip Morris—the world's largest transnational tobacco corporations—are among the largest foreign investors in Russia and Moldova, where the companies have been able to negotiate tax breaks, exemptions from monopoly legislation, and other conditions beneficial to their business but not beneficial to public health (Gilmore & McKee, 2004).

the main links between trade and foreign investment are summarized in **EXHIBIT 20-4**.

Trade and Its Effects on Health Systems

In addition to having varied impacts on health status and risk factors for disease, trade has led to significant changes in the way in which health systems function, interact, and are financed. These effects have been created through trade in both goods—principally pharmaceuticals—and services; both of which are explored in this section.

Trade in Pharmaceuticals

As outlined in the *Pharmaceuticals* chapter, pharmaceuticals are a significant element of healthcare expenditures, accounting for some 7% to 21% of such spending in HICs and 24% to 66% in LMICs. They are the single most important health-related product traded, representing some 55% of all health-related trade by value; by comparison, the share of the next most significant health-related goods traded, small devices and equipment, is 19% (Smith & Hanson, 2012).

Over the last few decades the pharmaceutical industry, as a mainstay of modern medical care, has grown in both size and concentration. At the turn of this century, the industry earned some $500 billion in global sales, but by 2013 this amount had increased to $950 billion in current prices (Hanefeld, 2015). It is an industry dominated by TNCs, with the top 10 earning more than half of these total revenues (increasing from earning just one-third in 1995). This concentration of the industry in fewer, larger firms (**TABLE 20-1**) has been accelerated by a need to exploit economies of scale, given the increasing costs of research and development (R&D); market positioning (especially with tighter regulation of pricing in HICs and emergence of markets in LMICs); the firms' desire to broaden their portfolios (with less reliance on a single blockbuster drug); and acquisition of generic manufacturers to both reduce the cost of production and remove competition, especially in India and in Asia more widely (Busfield, 2003; Lee, Yach, & Kamradt-Scott, 2011).

In general, however, despite increases in production and consumption of pharmaceuticals in LMICs, the pharmaceutical industry retains a strong HIC focus. Both production and consumption remain concentrated in HICs. Moreover, many companies retain a national base and even identity in HICs (e.g.,

TABLE 20-1 Top Ten Pharmaceutical Companies in Rank Order by Sales Value, 2012

Corporation	Headquarters	Sales Value (U.S. $ million)
Novartis	Switzerland	50,761
Pfizer	United States	46,930
Merck & Company	United States	40,115
Sanofi	France	37,780
Roche	Switzerland	35,069
GlaxoSmithKline	United Kingdom	32,714
AstraZeneca	United Kingdom	31,893
Johnson & Johnson	United States	27,933
Abbott	United States	26,715
Teva	Israel	24,846

Data from IMS Health (2013)/IQVIA.

Pfizer as a U.S. company), with 75% of the industry concentrated in Europe, North America, and Japan. Importantly, the size and concentration of the industry creates specific issues for trade policy and governance, compared to other areas covered in this chapter that tend to be more diffuse and national in character.

A critical issue is the business model of pharmaceutical R&D: It front-loads the costs of drug discovery, development, and licensing, with returns coming only after this substantive investment has been made, through the temporary monopoly provided by intellectual property (IP) rights. IP rights—and patents more specifically—grant legal exclusivity as a compromise to remove the disincentive to R&D, but also provide only a limited amount of time in which to recoup costs before the product faces competition (Smith, 2003). Patents have been the mainstay of policy to ensure investment in pharmaceutical R&D, acting as guarantor of monopoly rents. However, the high price charged at a product's launch and for several years thereafter creates problems for poorer countries and populations as they seek to access these medicines. Additionally, patents generate investments only when profitable markets exist; they do not work for drugs needed to address many of the diseases that prevail in LMICs (such as malaria).

This clear tension between the need for incentivizing private-sector funding of R&D and the need to use the drugs developed as widely as possible to meet public health goals was expressed as early as the 1970s. On one side, multinational pharmaceutical producers based in HICs expressed growing concern over alleged "misappropriation" of their patented technology by enterprises based in LMICs. On the other side, LMICs expressed deep concerns about the imbalance between technological capacity and ownership of technology as occurred between HICs and LMICs. This dialogue played out in a series of negotiations at UNCTAD, the World Intellectual Property Organization (WIPO), and ultimately the GATT during the Uruguay Round of trade negotiations. The last culminated in 1993, with the WTO's Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) entering into force in 1995 (Smith, Correa, & Oh, 2009).

The TRIPS agreement served as the catalyst for the change from the GATT to the WTO, as it moved beyond goods and into services trade. TRIPS established minimum standards for protecting and enforcing nearly all forms of IP (patents, trademarks, and copyrights) for WTO member states across all products, with standards derived from legislation in HICs. All member states are required to comply with these

standards, and where necessary to alter their national legislation to do so. In an important departure from previous conventions, pharmaceutical products were accorded full IP rights under TRIPS, and companies were granted the legal means to prevent others from making, using, or selling the new invention for a period of time—specified as 20 years.

There is a clear conflict between TRIPS and public health. By their very nature, the monopoly rents afforded by patents are reflected in the final product's pricing, so they act as a barrier to affordability. A flashpoint on this issue occurred toward the end of the 1990s, concerning HIV/AIDS medications. The annual cost for treatment of persons with HIV/AIDS with life-saving medications at the beginning of the 2000s was approximately $12,000 per year, putting the treatment beyond the reach of most public health systems and patients in the LMICs where the greatest burden of HIV is found (Roemer-Mahler, 2010). At the same time, the patent provisions within TRIPS initially prevented the export and use of generic, less expensive versions of such intellectual property being made in India, where the largest capacity for generic pharmaceutical manufacturing exists. Even in countries where domestic manufacturing capacity existed, notably in South Africa, drug companies initially sought the protection of their intellectual property.

For example, in South Africa, manufacturers of HIV/AIDS medications took the government to court over this issue in 1998. Subsequently, in 2001 the South African government sought to amend the South African Medicines and Related Substances Control Amendment Act, which would allow the import and use of less expensive generic versions of prescription drugs required to treat HIV/AIDS. The key clause stated that the government could find and "parallel import" the cheapest drug available globally and also grant "compulsory licensing" to other companies to allow them to make copies of the patented drugs. The government argued that this policy was justified and acceptable under Article 8 of TRIPS, which states that in framing national laws, members "may adopt measures necessary to protect public health and nutrition, and to promote the public interest." The prevalence of HIV/AIDS was such that the South African government felt that this article applied in this circumstance. Thirty-nine pharmaceutical companies, including GlaxoSmithKline, Merck, and Roche, launched legal action against this amendment, arguing that it was a direct violation of their patent rights. Supported by nongovernmental organizations (NGOs) such as Health Action International and Médecins sans Frontières, the case generated huge public pressure on the pharmaceutical companies, and the industry eventually withdrew the case. However, it left a lasting legacy in terms of its effects on TRIPS (Berger & Kapczynski, 2009).

Doha and Beyond

Attempts to address continued concerns related to public health protection under TRIPS led to the Declaration on the TRIPS Agreement and Public Health—known as the "Doha Declaration"—in November 2001. This affirmed the right of WTO member states to interpret and implement TRIPS in a manner supporting the protection of public health and, in particular, access to medicines (**EXHIBIT 20-5**). Critically, the Doha Declaration gave rights to countries to grant compulsory licenses and engage in parallel importation where a "public health crisis" made this necessary. Interestingly, countries have remained reluctant to use these available "TRIPS flexibilities" due to political pressures (Songkhla, 2009).

The comparatively limited application by countries of the TRIPS flexibilities has been repeatedly noted, including by the UN Secretary General's High Level Panel on Access to Medicines, and countries' right to use these flexibilities was explicitly affirmed as a target (3B) in the Sustainable Development Goals:

> Support the research and development of vaccines and medicines for the communicable and noncommunicable diseases that primarily affect developing countries, provide access to affordable essential medicines and vaccines, in accordance with the Doha Declaration on the TRIPS Agreement and Public Health, which affirms the right of developing countries to use to the full the provisions in the Agreement on Trade Related Aspects of Intellectual Property Rights regarding flexibilities to protect public health, and, in particular, provide access to medicines for all.

Specific concerns have arisen regarding those countries unable to access medicines due to their lack of manufacturing capacity. Paragraph 6 of the Doha Declaration recognized that "WTO members with insufficient or no manufacturing capacities in the pharmaceutical sector could face difficulties in making effective use of compulsory licensing under the TRIPs Agreement," and instructed the Council for TRIPS to find an expeditious solution to this problem. In 2003, a waiver was temporarily introduced that allowed countries with limited or no manufacturing capacity to import medicines produced under a compulsory license. For example, countries neighboring

EXHIBIT 20-5 TRIPS Flexibilities

The patent represents a balance that governments establish between inventors and the wider public. Sometimes the government will determine that the balance between the patent holder and the public needs to be adjusted. That is, the exclusive rights being exercised by the patent holder are perceived as causing a difficulty that the public should not be expected to bear. The TRIPS agreement and the Doha Declaration set out a series of flexibilities designed for countries to redress these imbalances.

Compulsory licensing is a core component of a country's patent law that is necessary to ameliorate potentially harsh consequences of patents on public health. The TRIPS agreement ultimately incorporated rules on compulsory licensing in Article 31. Article 31 addresses compulsory licensing by establishing procedural requirements for the granting of such licenses, without imposing restrictions upon the circumstances or grounds that justify the granting of such licenses. In terms of procedural requirements, Article 31 provides an accelerated or "fast track" mechanism when compulsory licenses are granted to address national emergency, in other circumstances of extreme urgency, or for public noncommercial use. The Doha Declaration on the TRIPS Agreement and Public Health, adopted on November 14, 2001, re-emphasized that governments may grant compulsory licenses on grounds of their own choosing, and that HIV/AIDS, malaria, TB, and other epidemics may constitute national emergencies.

Governments may also opt for *parallel importation,* which refers to authorization of importation of drugs under patent within the importing country that have been lawfully placed on the market outside the country. The TRIPS agreement authorizes WTO members to allow parallel importation of patented medicines, and this was confirmed by Paragraph 5(d) of the Doha Declaration. Parallel importation may help a country lower the cost of patented medicines because it allows for the purchase and importation of the lowest-priced version of the same medicine from anywhere in the world.

Although the basic concept of parallel importation is widely accepted, there are different schools of thought regarding the precise practices that are permitted. It is generally accepted that parallel importation covers patented products put on the market in any country by or with the consent of the patent holder. Some respected commentators take the view that pharmaceutical products lawfully placed on the market under compulsory license may also be parallel imported. Others have rejected this view on grounds that such products have not been placed on the market with the patent holder's consent.

In some countries, such as South Africa, authorization of parallel importation of patented medicines is expressly incorporated in national legislation. In most countries, rules on parallel importation are made by judicial decision. If a country authorizes parallel importation of patented medicines, it typically must also authorize parallel importation using the patent holder's trademark, since most pharmaceutical product packaging will also include a trademark or "brand name" of the patent holder. This is also permitted by the TRIPS agreement. Because originators may claim copyright protection for brochures or leaflets accompanying pharmaceutical products, it may also be useful to recognize parallel importation for copyrighted medical information as part of an authorization.

It is sometimes argued that parallel importation of pharmaceutical products does not actually reduce the price of medicines to consumers, because parallel traders and middlemen retain the price differential between imported and locally supplied products that should otherwise benefit consumers. However, the mere possibility of parallel imports of lower-priced medicines from foreign sources should exert downward pressure on local prices even if middlemen may absorb part of the price differential.

South Africa were permitted to import medicines that had been produced in South Africa under a compulsory license. This "temporary waiver" became a formal amendment to the TRIPS agreement in 2017, when two-thirds of WTO member states ratified the amended protocol (WTO, 2017b). It is hoped that this significant step will lead to much greater uptake of TRIPS flexibilities by the poorest countries facing the greatest barriers to access of medicines.

Three key issues continue to hinder improved access to medicines. The first issue is the inability of many countries, especially the very poor, to implement the "flexibilities" given their stark resource constraints. Even when medications are available at reduced cost, increased funding is still required for their purchase.

The second issue is the undermining of the flexibilities by provisions adopted under bilateral and regional trade agreements, known as "TRIPS-plus" measures, which are more restrictive than TRIPS and not bound by the flexibilities attached to TRIPS (Correa, 2006). TRIPS-plus refers to unilateral or bilateral trade negotiations in which standards beyond TRIPS (e.g., increasing patent life from 10 to 20 or 30 years, restricting compulsory licensing, removing the option to parallel import, and/or increasing exclusivity rights

over test data) are incorporated in exchange for trade concessions, particularly the promise of free access to markets for agricultural goods. Free trade agreements (FTAs), signed by the United States and European Union especially with a growing number of LMICs, have constituted one of the main routes for TRIPS-plus standards to be implemented. For instance, both China and Argentina are on the "Priority Watch List" of Special Section 301 of the U.S. Trade Act, which applies pressure to those countries to introduce their own TRIPS-plus standards (Office of the United States Trade Representative, 2007).

The third issue preventing greater access to pharmaceuticals is the need to achieve an appropriate balance between ensuring access to medicines and creating incentives for drug development and innovation. To date, this balance has remained elusive. Most recently, the concerns over new antibiotic production—how to incentivize R&D for new antibiotics when policy is also geared toward restricting their use—has promoted new energy and processes to develop systems that move away from current IP restrictions, although none has yet been implemented (Hanefeld, Khan, Tomson, & Smith, 2017; O'Neill, 2016).

Trade in Health Services

Trade does not only affect health care through goods. Advances in information, communication, and travel technologies have meant that increasingly services, such as telecommunications and electricity, as well as health and finance, are traded across borders (Smith, Chanda, & Tangcharoensathien, 2009). Trade in health services covers a wide range of areas, typically categorized according to the classification used in the WTO GATS (**EXHIBIT 20-6**). This model identifies four

EXHIBIT 20-6 General Agreement on Trade in Services

GATS, like all other WTO agreements, is binding on all member countries that sign it. There are some obligations and exemptions associated with GATS. Member countries have the right to restrict trade in services if they find that it poses a threat to human health. Although GATS is legally binding, countries have the opportunity to pull out after three years of officially adopting the agreement. However, they may have to compensate trading partners.

GATS permits WTO members to design the scope of their market access and national treatment commitments ("commitments" here refers to the obligations that countries enter into under the trade agreements), should they decide to make such commitments. From the national implementation perspective, the specific commitments are more important than the general obligations because GATS makes modification of specific commitments difficult to achieve. In general, the number of sectors for which commitments are made is positively related to the level of economic development, although some anomalies exist with respect to the health sector. In that area, overall commitments are lower than in any other sector except education, yet far more LMICs than HICs have made commitments in health. This suggests that trade in health services is more common between LMICs. The low overall level of commitments may exist for a number of reasons: Many countries simply ratified existing trade arrangements, which in health have historically been rare; health is seen by pace setters for trade—the United States and the European Union—as having less comparative economic value (compared with telecommunications or financial services, for example); and for most countries both health and education are recognized as public goods and have historically been subject to significant government involvement.

It is difficult to measure the effects of GATS on health systems. Not only are data lacking, but it is also difficult to isolate the influence of GATS from the many other factors that affect health systems simultaneously. Trade through all of the GATS modes has the potential to benefit health systems, but there are concerns that these benefits may be concentrated among the wealthy, increasing inequities within the health sector. In this respect, three specific aspects of GATS have particular importance:

1. GATS applies to services provided in competition, so it excludes services provided in the "exercise of government authority." While health services not supplied competitively fall outside GATS, if there is a combination of public and private providers, then the public providers are covered by GATS. This point is significant, as many health-sector reform programs have introduced some element of health care that is provided through private entities but publicly funded.
2. Sectors that are opened may contain qualifications, such as limitations on foreign equity investment, size limitations on facilities, or exclusion from subsidy. However, GATS requires as an absolute general obligation that members extend to all other members the best treatment that they give to their most favored trading partner (the "most favored nation" principle).
3. GATS does not affect member rights concerning market regulation for social policy purposes, such as universal access to care, but it could affect which types of regulations are allowed. For example, the "necessity test"

allows governments to deal with economic and social problems as they wish, provided that the measures and regulations are not more trade restrictive or burdensome than necessary. Obviously, much depends here on how "necessary" is defined.

Levels of activity and policy engagement in health-related services trade are growing under the combined influence of rising incomes in LMICs; demographic change, particularly population aging in LMICs and the ensuing pressures to contain health budgets; technological applications that facilitate the remote supply of an increasing range of health-related services, including to isolated populations in LMICs; continued liberalization of investment in services; and growing demand for skilled medical personnel and their cross-border mobility. A growing number of LMICs, particularly middle-income countries, regard health services, especially those that can be combined with tourism-related activities, as a potentially significant source of foreign exchange earnings, foreign investment, and skills upgrading. In turn, they are devoting significant policy attention to building health-related export clusters, with some having developed targeted investment promotion strategies. Current, and especially future, rounds of negotiations may yield much greater engagement and commitments by countries under GATS.

Information on GATS negotiations and up-to-date information on levels of commitments can be found at https://www.wto.org/english/tratop_e/serv_e/gatsqa_e.htm.

modes of supply of services, each of which is overviewed in this section (Blouin et al., 2006):

- Mode 1: Cross-border supply of services. This mode regulates the remote provision of health services, such as diagnostics and radiology, from one country to another.
- Mode 2: Consumption of services abroad. This mode covers "health tourism," the movement of people across borders to receive medical services.
- Mode 3: Commercial presence. Mode 3 involves FDI from one country to another. In the health sector, it usually involves setting up private hospitals and the presence of foreign insurance companies.
- Mode 4: Presence of natural persons. This mode covers the temporary movement of people from one country to another. In the case of health services, it would involve healthcare workers.

TABLE 20-2 summarizes these modes and the associated types of trade.

Mode 1: E-Health

Mode 1 refers to the cross-border provision of medical services. This is primarily telemedicine, in which diagnosis or expertise is provided by, for example, a physician in a different country, or in which medical records are transcribed by a service provider in another country (Smith, Chanda, & Tangcharoensathien, 2009). For instance, the shift from hard copy to digital imaging has enabled the remote storage, interpretation, and access of radiology images. Companies in China, for example, now provide hospital diagnostic services via this kind of electronic means for patients in Macao, Taiwan, and other countries (Blouin et al., 2006). A recent review of telemedicine found that examples documented in the literature mainly relate to telepathology, telesurgery, emergency and trauma telemedicine, and teleradiology, and that exchanges across borders were mainly between physicians consulting on patient cases, rather than between patients and physicians (Saliba et al., 2012). It identified cross-border supply of health services as driven by the need to address shortages in expertise predominately in LMICs, highlighting the positive potential of this form of trade.

Trade within this mode presents several opportunities: increased healthcare delivery to remote and under-serviced areas; alleviating (some) human resources constraints in service delivery; enabling more extensive and cost-effective disease surveillance; improving quality of diagnosis and treatment; and upgrading skills, by disseminating knowledge through interactive electronic means. It also raises some risks—namely, that such capital-intensive activity may divert resources from basic preventive and curative services, and that by largely catering to urban affluent populations, it may adversely impact on equity of access to services. Moreover, the implementation and use of e-health, even within national health systems, raises issues about the recognition of credentials and licensure; legal liability and malpractice considerations, including provider insurance coverage; provider remuneration; patient privacy and confidentiality; the existence of enabling infrastructure and compatibility of standards in areas such as data, images, and medical records; infrastructure and future operating costs; and the quality and appropriateness of care (Van Doosselaere, Wilson, Herveg, & Silber, 2007). The primary concern, though, is that electronic trade requires appropriate telecommunications and power

TABLE 20-2 Characterization of Trade in Health Services by GATS Modes of Supply

	Trade in Health Services	Trade in Ancillary Services	Trade in Goods Associated with Health Services
Mode 1: Cross-border supply	Telemedicine, including diagnostics, radiology	Distance medical education and training	Healthcare equipment
		Medical transcription, back office	Pharmaceuticals
		Medical research tools and databases	Medical waste
		Medical insurance	Prosthesis
Mode 2: Consumption abroad	"Medical tourism" (i.e., voluntary trip to receive medical treatment abroad)	All activities associated with health tourism (e.g., transport, hotel, restaurant, paramedical, local purchases)	
	Medically assisted residence for retirees	Local medical education and training of foreign nationals	
	Expatriates seeking care in country of residence		
	Emergency cases (e.g., accident when abroad)		
Mode 3: Commercial presence	Foreign participation or ownership of hospital/clinic or medical facilities (e.g., capital investments, technology tie-ups, collaborative ventures)	Foreign-sponsored education or training centers	
		Foreign-sponsored medical research facilities	
Mode 4: Presence of natural persons	Temporary movement of doctors and health personnel for the purpose of commercial medical practice	Temporary movement of doctors and health personnel for other purposes (e.g., education or training)	

infrastructure, which has led to developments in this mode becoming concentrated in countries that have this type of infrastructure, especially middle- and high-income countries.

Such remote supply of health services has the potential for impact on the health systems of both the countries contracting the services and the countries supplying them. The countries contracting the services can potentially improve the efficiency and the flexibility of their health systems if services can be carried out overnight—that is, if they are contracted to countries at a considerable time difference. More importantly, this has the potential to alleviate some of the staff shortages that some diagnostic services face. As some of these services are contracted to countries where salaries are lower, such as India, there is also potential for cost savings. In addition, this type of trade can increase the capacity of the country delivering the

services. The health systems of these countries, therefore, may benefit from increased profits and higher employment. At the same time, concerns have arisen that these services, and the higher salaries associated with them, may drive medical staff away from domestic services. Questions also exist about the quality of the services provided; the recognition of foreign degrees; and legal, malpractice, and liability issues related to telemedicine. **EXHIBIT 20-7** provides an illustrative case study of e-health trade.

EXHIBIT 20-7 Offshore Medical Transcription and Other E-Services in the Philippines and India

Medical transcription (MT) is the process of interpreting/encoding electronically the oral dictation of health professionals regarding patient assessment, therapeutic procedures, diagnosis, and so forth. U.S. outsourcing is the main driver of the global MT business. Since the entry into force of the Health Insurance Portability and Accountability Act (HIPAA) of 1996, which provides guidelines for safeguarding patient medical data, the need for medical transcription in the United States has expanded rapidly, growing at 20% per year. At the same time, the number of transcriptionists in the United States has declined at a rate of 10% per year.

The Philippines

The first large MT company, Outsource Transcription Philippines, started in the late 1990s. As the third-largest English-speaking nation in the world, with a large workforce, 94% literacy rate, and a strategic location with an ideal 12-hour time difference from the United States, the Philippines possessed key inherent advantages as a "first choice MT outsourcing destination" for the United States (Smith & Hanson, 2012, p. 183). Medical transcription is one of the five subsectors identified by the Philippines Department of Trade and Industry in its campaign to promote the country as a global hub for outsourced information technology (IT)–enabled services, carried out through the Center for International Trade Expositions and Missions (CITEM). The Philippines government has lent strong support to the MT industry by implementing an important e-commerce law, the Data Protection Bill (2006), and setting up the Information Technology and Electronic Commerce Council. The government has also sought to develop and expand the IT infrastructure of the country. Furthermore, the vast majority of companies exporting these services from the Philippines are ultimately owned by U.S. investors.

In the case of Philippines–United States outsourcing, privacy concerns have recently begun to increase. While patient information is protected though service contracts between importing hospitals in the United States and exporting transcription companies in the Philippines, it is expected that an increasing number of U.S. insurers will move to a model that legally requires information to be stored in the United States.

India

India has considerable opportunities in many aspects of e-health, including teleradiology, telediagnostics, telepathology, intensive care, ophthalmology, dermatology, psychiatry, and, to some extent, continuous online remote monitoring. Telemedicine is provided to other countries by independent telemedicine providers as well as reputed Indian IT companies such as Wipro. The Apollo Group in India exports e-health services. It provides telemedicine services (e.g., consultation, diagnostic, telepathology, teleradiology) from its Apollo Gleneagles Hospital in Kolkata to patients in Bangladesh, Nepal, Bhutan, and Myanmar. It also provides telediagnostic and teleconsultation services from its center in Karaganda Oblastu in Kazakhstan to the Central Asian region.

The United States and Singapore are also important markets for Indian telemedicine. The client base in the United States for leading Indian telemedicine providers has expanded from one hospital to 60 hospitals in just a few years. The National Healthcare Group of Singapore has contracted with telemedicine institutions in India to provide teleradiology services to designated hospitals in Singapore. Contracts have also been signed with Malaysia and with Dubai. India's potential in cross border e-health provision is driven by its cost advantages and the fact that specialized medical staff and doctors are involved in reporting and interpretation work (Chanda, 2008).

In addition to telemedicine, India exports health services in other forms through mode 1. It is an attractive market for outsourcing of healthcare business processes such as medical transcriptions, billing, coding, and data conversion, with potential to provide cost savings of 20% to 30% for client companies. For instance, Apollo has partnered with Health Services America and Med Staff International in the United States for billing, documentation of clinical and administrative records, coding of medical processes, and insurance claims processing for Medicare/Medicaid-claim policy holders and third-party administrators.

Outsourcing of pathology services is also emerging as a huge opportunity due to the high cost differential in India. Some Indian research labs and contract research organizations provide sophisticated tests such as molecular diagnostics for autoimmune disorders, cytogenetics, and diseases related to abnormalities in chromosomes and hormones. Some laboratories are able to offer a wide test menu of more than 1,500 tests under one roof.

Mode 2: Medical Tourism

Mode 2 is defined as the consumption of health services abroad. It is the most prevalent and long-standing form of trade in health services. Commonly referred to as medical travel or medical tourism, mode 2 involves the movement of consumers from one country to another for purposes of diagnosis, treatment, and rehabilitation and follow-up services.

Estimates of the number of medical tourists globally per year vary tremendously depending on the source, from a lower bound of 5 million to an upper bound of more than 40 million, with intermediate estimates putting the number at approximately 14 million per year. The financial value of mode 2 in health services trade is difficult to pin down, but conservative estimates place it in the range of $60 billion to $100 billion annually. According to McKinsey, 25% to 30% of these patients are expatriates, another 30% to 35% are seeking emergency care, and the remainder are patients who go abroad to seek treatment (Lunt, Horsfall, & Hanefeld, 2015). These numbers and values are much debated, as highlighted by Helble (2011), and currently no reliable data are available for global volume and flow of medical tourism (Lunt et al., 2015). This is particularly the case because medical treatment of travelers often takes place in the private sector, where companies may be reluctant to share commercial information.

Trade in mode 2 in health services is driven by differences in cost, quality, and availability of treatment across countries. Other factors include natural endowments; existence of alternative medicines and treatment procedures; long waiting lists for treatment in the source country; and cultural, linguistic, and geographic proximity between the sending and receiving countries. An in-depth study of outbound U.K. patients revealed that while factors such as cost, perceived quality and availability (including regulation and laws around treatment), and cultural affinity are important, so are interpersonal networks in determining where patients travel for treatment (Hanefeld, Lunt, Smith, & Horsfall, 2015).

In principle, medical travel occurs in all directions among high-, middle-, and low-income countries. In the absence of "hard" data, an initial common assumption has been that patients from HICs travel to LMICs to access cheaper health services there (Lunt, Horsfall, Exworthy, Smith, & Hanefield, 2013). Nevertheless, a small but increasing number of regional and national studies have revealed a much more diverse pattern (Bustamante, 2014; Crush & Chikanda, 2015; Hanefeld, Horsfall, Lunt, & Smith, 2013; Lautier, 2008; Noree, Hanefeld, & Smith, 2016). It is common for affluent patients in LMICs to seek specialized high-quality treatment overseas in HIC hospitals or in neighboring LMICs with superior healthcare standards (Hanefeld et al., 2013). For example, over the past decade there has been an increase in Nigerian patients seeking treatment in the United Kingdom, and in patients from the Middle East seeking treatment in Thailand (Noree et al., 2016). It is also common for persons in HICs to seek quality treatment at a fraction of the cost in LMICs, or to seek alternative medicines and treatments and take advantage of natural endowments in LMICs. For instance, patients from developed countries such as the United States and the United Kingdom can get bypass surgeries or transplants done at one-fourth or one-fifth of the cost in high-quality corporate and super-specialty hospitals in middle-income countries such as India, indicating the tremendous scope for gains from trade due to cost differences (**EXHIBIT 20-8** describes medical tourism in Thailand as an example). In addition, legislative differences and the availability of types of treatment have led to certain countries becoming known for areas of treatment—for example, fertility treatment and surrogacy in India (Whittaker, 2010).

Along with the travel of more affluent patients, increasing evidence suggests that a significant amount of cross-border trade in services involves the general population moving between LMICs, such that people cross land borders in search of treatment either at small-scale private providers or in the public sector (Crush & Chikanda, 2015). Research suggests that these patients are far from the empowered affluent consumers commonly associated with medical tourism, but rather are poorer members of the population seeking to access services and medicines unavailable in their countries of origin. Studies have, for example, focused on patients from Indonesia accessing services in Malaysia (Ormond, 2013), or patients from neighboring countries seeking treatment in South Africa (Vearey et al., 2016; Walls et al., 2015).

In addition to this trade in health services driven by individual motivation, governments have developed their own mode 2 schemes, whereby a country seeks to import health services for its population to address unmet needs. This has been the case in small island states such as the Maldives, for example, as they seek to achieve universal health coverage (Suzana & Chongsuvivatwong, 2015).

In sum, the different types of trade in health services under mode 2 and the pathways by which people travel or services are imported vary greatly, as do the health and health systems impacts associated with these flows. Common concerns have related to the equity impact on health systems of countries receiving medical

EXHIBIT 20-8 Medical Tourism in Thailand

Thailand has long been recognized as prime medical tourism destination. Since 2003, the Thai government has explicitly sought to make Thailand a destination for foreign patients, in an effort to increase government revenues.

An initial evaluation of the Thai policy reported that income from medical tourists contributed as much as 0.4% of Thai's GDP, with 1.2 million international patients visiting the country annually by 2006 (NaRanong & NaRanong, 2011). However, a large cross-sectional study of foreign patients in Thailand's five largest private hospitals published in 2016 provided greater detail and nuance on the shape and size of medical tourism in Thailand (Noree et al., 2016). Importantly, it suggested that the actual number of patients visiting Thailand was smaller than previously thought.

The 2016 data revealed that, on a regional basis, the largest number of medical tourists came from the Middle East, followed by Southeast Asia and Europe. The United States ranked third in terms of country of origin, however, with 7.5% of all medical tourists coming from that country. These findings highlighted the regional aspect of medical tourism.

Analysis of the conditions for which patients traveled also revealed a clear pattern. Middle-aged male patients from the Middle East predominantly traveled for heart procedures, while women from Southeast Asia traveled for cosmetic and fertility treatments, and Europeans tended to purchase digestive and orthopedic procedures. Other patterns included the finding that almost all medical tourists from Australia were female, visiting Thailand for cosmetic surgery only.

In terms of medical tourism's contribution to the Thai economy, the research found that many patients engaged in significant tourism activities, with short medical treatments as "add-ons." Indeed, a survey among medical tourists identified that their tourist expenditures (on accommodation, food, and activities) were higher than the tourism expenditures of nonmedical tourists, indicating the former are particularly lucrative. Using the estimated numbers and expenditures, Noree et al. (2016) concluded that the gross tourism revenues from medical tourists and their companions in 2012 was approximately $900 million.

Taken together, these study findings confirm that Thailand is a key destination for medical travel, albeit on a modest scale. They also suggest net benefits to the economy may be substantially smaller than previously thought, highlighting the need for good-quality data. As for health systems impact, there has been little evidence of a substantial internal brain drain from this practice: Doctors working within the private sector continue to also work in the public sector, and private clinics cater to both domestic and foreign patients. At the same time, private hospitals do draw on publicly trained human resources.

tourists (i.e., exporting health services). Other criticisms have focused on potential cost increases in the private sector for domestic patients (Noree, Smith, & Hanefeld, 2014), internal brain drain within countries from the public to the private sector owing to the focus on serving wealthier medical tourists (Lautier, 2008), and the ethical issues when gametes or organ donations are sought from poorer people in LMICs (Whittaker, 2011). A persistent challenge to addressing some of these issues has been the lack of an international governance structure tackling the issues arising in mode 2 (Hanefeld, Mandeville, & Smith, 2017).

Mode 3: Commercial Presence/FDI

Health services can be traded through commercial presence (mode 3), wherein hospitals, clinics, diagnostic and treatment centers, and nursing homes may be established across countries. In some cases, joint ventures, alliances, and management tie-ups may be forged between healthcare organizations across countries and regional networks of healthcare providers that are engaged in delivering health care through modes 1 and 2. Such arrangements may involve acquisition of facilities, management contracts, and licensing arrangements with some degree of local participation. Although conventional data from balance of payments statistics do not provide information on this mode of trade in health services, available information such the Foreign Affiliates Trade in Services Statistics (FATS), which cover a variety of indicators (exports, imports, sales, turnover, and employment) regarding the activities of foreign companies in overseas markets, indicate that health services FDI has been growing in significance over time.

Traditionally, HICs have been the leading sources and destinations for FDI in health service. However, over the past decade, a growing number of LMICs have also emerged as host nations. The value of mergers and acquisitions in health and related social services reached $14 billion in 2006 in terms of sales transactions, with diversification of source and recipient markets occurring over the years. Transnational hospitals and healthcare providers with both regional and global presence have emerged in recent years, with several middle-income country health services firms based in Asia and Africa entering foreign markets through investment, mergers, and acquisitions of service providers and hospitals. It is also interesting to note the emergence of South–North and South–South

TABLE 20-3 Recent Acquisitions of Healthcare Providers Involving Non-OECD Countries

Year	Investor	Subsidiary	Exporting Country	Importing Country	Value of Investment	Nature of Investment
2006	Netcare	General Healthcare	South Africa	United Kingdom	£2.2 billion (~ $2.5 billion)	52.6% stake
2007	Mediclinic	Emirate Healthcare Holdings	South Africa	United Arab Emirates	$46.4 million	49% stake
2007	Mediclinic	Hirslanden	South Africa	Switzerland	$2.4 billion	100% stake
2005	Bumrungrad International	Asian Hospitals	Thailand	Philippines		45.5% stake
2006	Bumrungrad International	Bumrungrad Hospital Dubai	Thailand	United Arab Emirates		49% stake (joint venture with Istithmar)
2007	Bumrungrad International	Asia Renal Care	Thailand	Singapore (operates clinics in 6 Asian countries)	$75 million	100% stake
2005	Apollo Hospitals	Apollo Hospitals Dhaka	India	Bangladesh	$35 million	100% stake
2005	Parkway Healthcare	Pantai Hospitals	Singapore	Malaysia	$139 million	31% stake

Reproduced from Mortensen, J. (2008). International trade in health services-assessing the trade and the trade-offs: DIIS Working Paper no 2008/11. https://www.diis.dk/node/1587

flows of capital in health services, as highlighted in **TABLE 20-3**. The bilateral pattern of investment indicates the significance of factors such as geographic and cultural proximity, regional markets, growth dynamics, and market size.

EXHIBIT 20-9 highlights the growing regional and global presence of healthcare providers from selected high-, low-, and middle-income countries, and provides more detail on some of the firms shown in Table 20-3. Entry into overseas markets has taken place through joint ventures, franchises, greenfield investments, acquisitions, tie-ups, contractual arrangements, and public–private partnerships. Linkages are also evident with other forms of health services trade.

The growth of the foreign commercial presence in health services has been mainly driven by the relatively recent opening of the health sector to participation by foreign hospitals, diagnostic centers, and clinics, as well as privatization and deregulation of the healthcare sector in many countries. For instance, Indonesia allows Singaporean, Australian, and Canadian firms to operate in its market. India has permitted automatic approval for 100% FDI in hospitals since 2000. Between 2000 and 2006, nearly 100 FDI projects in hospitals and diagnostic centers were approved in India, representing a total of $53 million in investment from high-, middle-, and low-income country sources. Thailand's open FDI regime for hospital services has resulted in several part-foreign-owned hospitals, mainly in the Bangkok area, with investments from Japan, Singapore, China, Europe, and the United States being funneled to the country's healthcare sector. Cambodia permits cross-border investment in hospital services and foreign ownership; it also permits management of private hospitals and clinics as long as at least one director is a national. Foreign firms are allowed to provide dental services through joint ventures with Cambodian legal entities.

EXHIBIT 20-9 International Health Services Provider Firms

Singapore

The Parkway Healthcare Group is the biggest investment group for health care in Singapore and one of the largest healthcare organizations in Asia. It created Gleneagles International as an international brand. The company has been interested in acquisition of hospitals in Singapore, building up a base there, and entering countries such as India, Indonesia, Malaysia, Sri Lanka, and the United Kingdom, mostly through joint ventures with local partners. It entered the Indian healthcare market in 2003 through a joint venture with the Apollo Group and built the Apollo Gleneagles Hospital, a multispecialty hospital at a cost of $29 million. It has formed a joint venture with the Mumbai-based Asian Heart Institute and established a research center to provide medical excellence. It is in the process of setting up a specialized heart hospital in London.

The Singapore-based Raffles Medical Group is building strategic alliances through triangular business associations with healthcare organizations from high- and middle-income countries. It is venturing into low-income countries in partnership with host country investors.

Thailand

Bumrungrad Hospital in Thailand has entered into management contracts with hospitals in Bangladesh and Myanmar. It has formed a joint venture with a hospital in the Philippines. Bangkok Hospital has 12 branches in Southeast and South Asia, located mostly in tourist towns.

India

The Apollo Group of Hospitals has centers of excellence in several countries, including Nepal, Sri Lanka, Ghana, and Bangladesh. It has also entered into contract-based management of hospitals or clinics in the United Arab Emirates, Oman, Kuwait, Mauritius, Malaysia, Sri Lanka, and Nigeria. It has established a telemedicine center in Kazakhstan. Apollo Hospitals has a joint venture with Amcare Labs, an affiliate of Johns Hopkins International, to set up a diagnostic laboratory in Hyderabad.

South Africa

South African health services firms are present in the United Kingdom, Switzerland, and the United Arab Emirates, and are also the main source of regional FDI in southern Africa. Some major firms include Netcare, MediClinic, Life Healthcare, and the Afrox Health Care Group. MediClinic owns private hospitals in Namibia; Life Healthcare operates private hospitals and clinics in Botswana; and the Afrox health care group has operations in Botswana, Namibia, Zambia, and Mozambique. Netcare has a public–private partnership with the Lesotho government to build a hospital, refurbish two feeder clinics, and run clinical services for the government.

China

The public and private sectors in China have jointly developed a strategy to attract foreign health providers to set up a commercial presence. Chinese institutions have entered into joint ventures with partners in the medical profession and with local authorities overseas. Traditional Chinese medicine (TCM) facilities have been established in more than 20 countries. Such joint ventures help spread TCM overseas, enable the deployment of Chinese health workers and their exposure to other systems under contractual arrangements, and help attract patients to China.

Other Countries

The large U.S. corporate healthcare sector, which owns for-profit hospitals, has entered foreign markets mainly through small specialized clinics, such as eye clinics, rehabilitation centers, and outpatient clinics. U.S. firms typically enter these markets through joint ventures with local specialist doctors or surgeons.

Columbia Asia Group, a Seattle-based hospital services company, is a worldwide developer and operator of community hospitals. It has started the first American-style medical center in Bangalore.

The Fresenius Medical Care group (FMS) is headquartered in Germany and is one of the leading foreign healthcare providers in the United States. FMS has operations in Belgium, France, Italy, the Netherlands, Portugal, Switzerland, and the United Kingdom. It has affiliates in Australia, Singapore, Malaysia, Thailand, Korea, Taiwan, Philippines, Hong Kong, and Japan, and representative or branch offices in New Zealand, India, Indonesia, and China.

Nevertheless, numerous regulatory barriers restrict mode 3 trade in health services. These constraints include limits on foreign equity participation; limits on type of foreign commercial presence; economic needs tests; authorization, certification, and licensing requirements; discriminatory taxes; and technology collaboration and transfer conditions.

The impact of foreign commercial presence in health services can be seen in its effects on capacity, quality, and equity. The literature indicates that these effects are shaped by the existing structure of the health system, its public–private mix, and the national regulatory environment. In terms of its implications for efficiency, mode 3 trade in health services can augment a country's health resources by bringing in additional financial resources, enabling capacity expansion, and alleviating the pressure on government budgets. On the downside, it could create inefficiencies by encouraging overinvestment of resources in high-end and highly capital-intensive and specialized treatments and procedures with lower cost-effectiveness, while diverting funding from basic healthcare services. Likewise, it could lead to long-term outflows of payments to foreign investors, along with direct and indirect subsidization costs for incentives given to foreign investors.

On the quality front, foreign commercial presence in hospitals and health management may improve the quality of national health systems through the introduction of better management techniques and information systems; better technology, equipment, and infrastructure; and improved standards and accreditation. There could also be positive spillover effects on the standards and practices of domestic establishments. Such improvements in the quality of the domestic healthcare system and presence of foreign healthcare providers of global standards could, in turn, benefit the country by reducing its residents' spending on expensive treatments overseas. These gains are contingent on the regulatory environment's ability to ensure quality and standards, the ease with which technology and equipment can be accessed, and the spillovers to the rest of the healthcare system.

Likewise, mode 3 trade in health services has mixed implications for equity. If such establishments cater to the urban and affluent segments of the population who can afford to pay, then they may aggravate existing inequities in access to health care. Such establishments are more likely to focus on tertiary care, specialized treatments, and curative and intervention-oriented procedures rather than primary and preventive healthcare needs. There may be cost implications as well, if foreign-owned and managed health providers charge higher fees. Foreign investment in health services, particularly in hospitals, could also distort the healthcare market by encouraging a brain drain—that is, the movement of the most qualified and specialized health personnel toward such establishments and away from domestic establishments owing to offers of better pay and facilities. This could adversely affect the quality of medical human resources in competing institutions, particularly public-sector hospitals.

On the bright side, there could be positive effects on equity through the creation of employment opportunities in the healthcare sector at all levels, with better remuneration, especially for specialized and senior medical professionals. Such establishments may attract overseas medical professionals and returnees who are internationally accredited, thereby augmenting the human resources capacity and quality in the host country. Once again, factors such as the extent of health insurance penetration, the public–private mix in the healthcare system, and regulatory requirements on pricing and access for the poor will shape the equity consequences of mode 3 trade in health services.

Mode 4: Migration of Healthcare Workers

Health services can also be traded through the temporary movement of health personnel (mode 4), including doctors, specialists, nurses, paramedics, midwives, technicians, consultants, trainers, health management personnel, and other skilled and trained professionals. Today, much of cross-border mobility of health providers does not constitute mode 4, but rather involves permanent migration. Mode 4 trade in health services is a subset of such movement, which is temporary in nature, usually occurring under bilateral contracts between institutions and/or governments and aimed at addressing shortages such as of nurses or specialists in the receiving market (Kingma, 2007). As with other modes of health services trade, it is difficult to estimate the value of mode 4 trade in health services, as statistics on these flows are rare. This is especially the case in terms of delineating temporary from permanent cross-border movement in the health sector aligned with the guidelines laid down in the Manual on Services of International Trade in Services (MSITS).

Notwithstanding the limited data, existing mobility patterns clearly indicate that both HICs and LMICs are engaged in health services trade via mode 4. There are mode 4 exports on short-term contracts from LMICs to HICs, such as from India and the Philippines to countries in the Persian Gulf region, or from Cuba to countries in Africa and the Caribbean (Eastwood et al., 2005). The Middle East is an important host market for a wide range of health professionals from high-, middle-,

and low-income countries, including doctors, nurses, x-ray technicians, lab technicians, dental hygienists, physiotherapists, and medical rehabilitation workers. **TABLE 20-4** highlights some of the individual "push" and "pull" factors that help explain the pattern of movement of health workers under mode 4 (Eastwood et al., 2005).

An important aspect of mode 4 trade in health services is managed mobility through bilateral and regional trade and economic cooperation arrangements (Buchan, 2006; Plotnikova, 2014). Some of these arrangements are narrow in their focus, with their objective being to facilitate movement of health providers. Others go beyond the issue of mobility to address a wide range of associated issues including recruitment, return and reintegration, recognition of qualifications, harmonization of standards and training, and capacity building. Some arrangements also address aspects such as cooperation in health systems development and best practices for induction, training, and treatment of foreign health professionals. These types of agreements reduce the need to use commercial recruitment agencies by relying on cooperation between recognized agencies/institutions in the sending and receiving countries, thereby narrowing the risk of exploitation of migrating professionals, enabling greater transparency in the mobility process, and shifting much of the mobility costs from the migrating individual to the client. A broader objective of some of these arrangements is to create win–win possibilities for both the receiving and sending countries through health worker mobility and to avoid the negative consequences that arise from long-term migration of health services providers.

Examples of arrangements that address health worker mobility in a broad-ranging manner are CARICOM and the Caribbean Free Trade Agreement (CARIFTA) (Stilwell et al., 2004). The focus of these agreements in the context of mode 4 trade in health services is regional collaboration and regionally managed migration. This includes a nursing campaign within the Caribbean to increase the supply of nurses within the region and measures to improve human resources management so as to retain and attract back nurses (Salmon, Yan, Hewitt, & Guisinger, 2007). To augment the regional supply of nurses, the member countries are collaborating on training capacity, including establishing distance education programs at the BSc and MS levels in nursing, a mentorship program directed at recruitment and retention of nurses, and a regional training program for nurses. To improve human resources management of nurses, a regional convention on nursing has sought to address issues related to the work environment. In addition, regionally coordinated incentives have been created to value nursing and to improve the utilization and deployment of nurses through a workload measurement system.

There is a regionally managed migration program between countries such as St. Vincent, which trains healthcare providers for export, and countries such as Jamaica and Trinidad and Tobago, which have shortages of such workers (Salmon et al., 2007). Skilled professionals across many sectors, including health care, are encouraged to work in other countries of the region on a rotational basis for three years, and then return to the home country. To facilitate the movement of health professionals within the region without

TABLE 20-4 Push and Pull Factors in International Markets for Skilled Health Workers

Push Factors	Pull Factors
Low pay; late payment and non-payment	Higher pay
	Opportunities for remittances
Low job availability	Vacancies
Poor working conditions (inadequate resources; high workload; poor human resources planning)	Better working conditions
Lack of resources to work effectively	Better resourced health systems
Limited career opportunities	Career opportunities
Limited educational opportunities for workers and children	Provision of post-basic education; good general education system
High prevalence of HIV/AIDS	Low prevalence of HIV/AIDS
Unstable/dangerous work environment	
Unstable general political environment: risks of political and criminal violence	Political stability
	Low crime rates
Economic instability	Economic stability

Reproduced from Smith, R. D., & Hanson, K. (Eds.). (2012). *Health systems in low- and middle-income countries: An economic and policy*. Oxford, UK: Oxford University Press.

compromising the quality of health services, initiatives have been launched, with the involvement of nursing associations and regulatory bodies in member countries, to harmonize educational standards and entry requirements in the profession. For instance, nurses in all of the English-speaking Caribbean countries sit for the same registration examination.

Overall, the Caribbean strategy is multifaceted. It covers aspects such as recruitment, retention, training, utilization, deployment, work terms and conditions, management practices, and long-term human resources development in the health sector. It also attempts to pool regional resources in the health sector and to address human resources imbalances within the region through managed migration policies, supplemented by efforts to develop regional capacity through training programs (with the aid of technology) and improved human resources management.

Other preferential agreements that address the issue of temporary mobility of healthcare providers, but are not as broad ranging as the aforementioned Caribbean arrangements (Kanchanachitra et al., 2011), include the Japan–Philippines Economic Partnership Agreement (JPEPA) and the India–Singapore Comprehensive Economic Cooperation Agreement (CECA). Under the JPEPA, one of the major issues negotiated was the easing of restrictions in Japan's labor market to accommodate more Filipino healthcare professionals. Following difficult negotiations on this issue, Japan agreed to allow a limited number of nurses—100 in the first year—to stay beyond the current four-year time limit if they acquired a Japanese license and underwent vocational training. Japan also insisted on Japanese language proficiency requirements as an eligibility condition for entry. However, the program was not deemed a success, in part due to low uptake (Yagi, Mackey, Liang, & Gerlt, 2014).

An important regulatory issue that has been addressed in some economic integration arrangements is mutual recognition and harmonization of qualifications. In the European Union, several directives relate to the facilitation of free movement of doctors, nurses, pharmacists, and other health professionals and mutual recognition of their diplomas, certificates, and evidence of other formal qualifications, among the member countries. These directives provide mechanisms for convergence and coordination. They entitle any EU physician who has completed basic training in a member state and who holds a recognized qualification to be automatically registered in any other member state. The sectoral directives are based on the principle of mutual confidence and comparability of training levels. This mindset is reflected in the "Recognition of Foreign Professional Qualifications Act," which requires EU member states to consider the practical experience of an individual in the process of recognition of qualifications. Where there are major structural differences in training requirements, the host country is entitled to require compensation through an adaptation period or an aptitude test.

In other regions, mutual recognition of qualifications has also been an important issue in the context of mode 4 trade in health services. Under the ASEAN Framework Agreement on Services, which aims at achieving an FTA in services by 2020, one of the seven priority sectors is health services (Kanchanachitra et al., 2011). Nevertheless, challenges have emerged in operationalizing this mutual recognition among the member countries due to differences in educational standards and programs, regulatory systems and licensing practices, and language and culture.

Likewise, under the India–Singapore CECA (Chandra, 2010), visa restrictions have been eased for Indian professionals in 127 categories, including for medicine and nursing. The two governments have decided to mutually recognize degrees for the purposes of issuing multiple entry visas. This decision reflects a synergy of interests regarding the mobility of health workers: India is interested in overseas deployment of its healthcare professionals, while Singapore needs foreign health personnel to meet its local requirements for care (Chandra, 2010).

In addition to trade and cooperation agreements, sector-specific bilateral labor agreements may be created to address health worker mobility. For instance, Cuba has signed bilateral agreements with various countries in the Caribbean and in Africa. It provides doctors to these countries to address domestic shortages in underserved areas (Feinsilver, 2009). It also sends health personnel to these countries to help enhance host country capacity through the provision of training to local personnel. Cuba has such agreements with Botswana, South Africa, Namibia, and Zimbabwe. The agreement between Cuba and Guyana involves broad cooperation in terms of human resources deployment and educational exchange in the health sector. For example, for more than 20 years, a Cuban medical brigade has provided its services in Georgetown, Guyana, and other locations of that country. Cuban specialists founded the medical school in Georgetown, Guyana (Blouin et al., 2006).

Interlinkages Between Modes

Although each delivery mode for health services trade has been discussed separately in the preceding

subsections, the considerable links between these modes should be recognized. For example, Cuba has used joint ventures with Canadian, German, and Spanish companies (mode 3) to attract patients from these countries for specialized treatments (mode 2). This trade has helped establish Cuba as a hub for teleconsultation and telediagnostic services (mode 1) to the Central American and Caribbean market and has led to the establishment of specialized Cuban clinics in Central and Latin America (mode 3), where Cuban physicians and nurses are employed (mode 4). Thus, health services trade should be considered as an integrated package to appreciate fully the potential risks and opportunities presented (Chanda, 2006).

Other examples highlight the scope for such linkages. A mix of modes is evident where South African hospital companies have succeeded in winning healthcare contracts abroad, including with the United Kingdom's National Health Service. Netcare established its presence in the United Kingdom in 2001. The company has sent teams of medical personnel from South Africa to its U.K.-based establishments for fixed periods, thereby enabling its employees to work 4 to 6 weeks at a time abroad, get exposure to opportunities overseas, and supplement their income with fixed-term contracts abroad.

India's Apollo Group of Hospitals engages in mode 2–based health services exports, which have been facilitated by its overseas marketing offices and management contracts with hospitals in the United Arab Emirates, Saudi Arabia, Oman, Kuwait, Mauritius, Tanzania, United Kingdom, Sri Lanka, Bhutan, Nigeria, Pakistan, and Bangladesh. Apollo Gleneagles, which is a joint venture with the Singapore-based Parkway Group, exports health services to patients from neighboring countries such as Bangladesh, Nepal, Bhutan, and Myanmar. It also provides telemedicine services such as medical consultation, diagnostic, telepathology, teleradiology, and scanning services. In addition, Apollo provides contract research and medical education and training services through its overseas subsidiaries, using a combination of cross-border supply (online training and research services) and temporary on-site deployment of professionals at its subsidiaries, thereby benefiting its own professionals as well as host country professionals.

Overall, this overview of trade in health services and the different modes highlights the message that such trade is an area of significant growth. At the same time, it is one where data are limited, and our understanding of how these practices will affect health and health systems is still underdeveloped and in much need of further research.

Conclusion: From Conflict Between Trade and Health to Coherence

In the final paper in the 2009 *The Lancet* series on trade and health, three cross-cutting priority areas were outlined to cut across the many issues related to health and trade: "(i) highlighting the need for stronger evidence concerning trade and health links; providing (ii) the foundation for building trade and health engagement and capacity; in order to (iii) facilitate the assertion of health goals in trade policy" (Smith, Lee, & Drager, 2009). These priority areas still stand.

There remains a lack of sound empirical evidence demonstrating how trade and growth in trade are linked directly and indirectly to health. Even though the positive links among increased trade, poverty reduction, and economic growth are widely accepted, evidence regarding the impact of increases in trade on the determinants of health varies from one national context to another. Hence, adapting trade to national conditions is important in ensuring desired outcomes. This requires consideration of "flanking" policies that may be undertaken to mitigate adverse consequences, and the design of trade policies that reduce the potential health risks associated with freer trade while maximizing the positive impact of more trade on the social determinants of health.

Trade policies adopted by national governments affect health and health systems through a diverse set of channels and intermediate variables. Unfortunately, these causal linkages can be difficult to track and monitor. Moreover, in many political systems, health authorities are not in a position to directly affect trade policy decisions at the national level. Even so, their existing knowledge of the determinants of population health and their jurisdiction over social and health policies place health policy makers in a privileged position to ensure that, in an increasingly global economic environment, domestic policies and regulations are designed to maximize the potential of, and to minimize the risks presented by, increasing levels of trade.

Although more evidence on this topic is necessary, it is not sufficient to ensure that health is more integrated in trade negotiations and decision making. A common theme in this chapter is the dynamics of the *linkages* between the various constituent elements. For example, the delivery of health care is subject to pharmaceutical availability, which depends

on local price negotiations, international legislation, and domestic exchange rates. At the regional and global levels, changes in food production processes through trade liberalization will affect health and consequently demands for health care. Thus, those with a health focus will need to engage with agricultural reforms and policies to target agricultural commodity and food retail prices and other influences over food security, such as food safety, security within the food supply chain, and the balance in the use of agricultural land for biofuels and other "cash crops" versus nutritional foodstuffs. Another example is the enormous potential for increases in zoonotic outbreaks: It will require health-sector involvement in the development of international rules and legislation concerning movement of people, animals, and food products. Linkages such as these imply a substantive requirement for those with a health remit to engage in negotiation with those from other sectors and from other geographic locations.

Ultimately, perhaps the biggest challenge is breaking the "silo mentality" that suggests trade and health operate as separate policy spheres. Developments such as those mentioned in this chapter mean that these areas will face an ever more expanded agenda of issues. Although some issues have produced closer cooperation between the two sectors, others have exposed tensions between the goals of protecting health and promoting trade. Indeed, a former Health Minister of Thailand has said that "The lessons from compulsory licensing in Thailand reflect poor governance and double standards in global trade and health" and that "inadequate support was provided by WHO and the World Trade Organization" in the country's dispute over TRIPS (Songkhla, 2009). In this respect, the potential risks associated with trade and health are further increased by the added complication of conflicts, or misunderstandings, between the trade and health sectors, and therefore by further confusion in estimating the potential benefits and risks of trade liberalization for health. When making trade commitments, national ministries of trade (and perhaps finance and foreign affairs) often do so in isolation from health ministries—yet those decisions have an impact on health, of which the trade officials have limited knowledge. Conversely, ministries of health typically have limited knowledge in trade issues. A critical issue in trade and health is the need to address this asymmetry of information, by enabling ministries of health to make informed and comprehensive presentations to ministries of trade concerning decisions to be taken with respect to trade and trade agreements.

The pace of change, especially at the global level, has affected the ability of national domestic health sectors to set policy and to engage fully in such activities without wider understanding and capacity in negotiations at the international level. Indeed, it has been suggested that WHO should develop a "Committee C"[2] to provide a forum for debate about major health initiatives by WHO member states and non-WHO organizations, including international agencies, philanthropic organizations, multinational health initiatives, and representatives from major civil-society groups (Kickbusch, Hein, & Silberschmidt, 2010). This idea, colloquially termed "global health diplomacy," is covered in more detail in the *Global Health Governance and Diplomacy* chapter.

Of course, the question is where to begin. Many national and international information sources discuss trade and the general macroeconomy, such as public expenditure reviews and trade policy reviews. The key challenge facing health policy makers is how to interpret this information with respect to the health sector. Which information needs to be extracted? How might this information be interpreted? Which health and trade policies should be adopted in light of the information? To help guide policy makers through this quagmire, Smith and Chanda (2005) developed a framework to investigate these issues, which is provided in **EXHIBIT 20-10**, and which could offer a first step in the desired direction.

In sum, health and health care are affected by trade and trade liberalization in a number of direct and indirect ways: Healthcare activities, products, and services can themselves be traded; health systems are impacted by changes in exchange rates; the health profile of a country is affected by global impacts on communicable and noncommunicable diseases; and numerous aspects of trade policy influence health system flexibility. Trade liberalization has both beneficial and detrimental effects for the health systems of the countries engaging in it. The challenge facing national governments is to capitalize on the positive effects and regulate their healthcare markets appropriately to avoid the negative ones. To achieve this goal requires that all those engaged in health understand the importance of trade and engage with their counterparts involved in trade and trade policy.

[2] Currently the World Health Assembly prepares resolutions and decisions to be taken in two main committees: Committee A deals with program matters, and Committee B focuses on budget and managerial concerns.

EXHIBIT 20-10 Framework for Collecting Data on the Impact of Trade in Health Services on Health Systems

The following steps represent a framework that national governments may adopt to evaluate the impact of trade on their health system.

1. Establish the country's general background on macroeconomic and trade environment. This requires determining the following:
   a. Country's macroeconomic status and stability
   b. Trade and balance of payments
   c. Degree of openness to trade and investment regime
   d. Overall policy objectives
2. State of domestic health system. This is done by collecting data on:
   a. Amount of investment in the healthcare sector
   b. Demand and supply conditions
   c. Balance between public and private provision of care
   d. Policy environment
   e. Infrastructure conditions
   f. Regulatory framework
   g. Human resources capabilities
   h. Labor market conditions in the health sector
3. Agreement (or particular mode within agreement) specific elements. This requires determining the following:
   a. Current state of trade and investment in the health sector
   b. Direction of policy
   c. Current status of international commitments and proposed liberalization by the agreement or mode in question
   d. Institutional capacity with regard to trade in health services
   e. Data sources and availability of information on the health sector

Reproduced from Smith, R. D., & Hanson, K. (Eds.). (2012). *Health systems in low- and middle-income countries: An economic and policy*. Oxford, UK: Oxford University Press.

Discussion Questions

1. Identify at least two positive examples and two negative examples of traded goods or services that affect health.
2. How do you think your own field of work or interest within public health is being affected by trade? Which regulatory mechanisms or incentive systems do you think may be needed to tackle the issues this trade raises?
3. Looking across the different modes of trade in services, which ones do you think will be key for achieving universal health coverage, and why?

References

Adlung, R., & Carzaniga, A. (2001). Health services under the General Agreement on Trade in Services. *Bulletin of the World Health Organization, 79*(4), 352–364.

Agege, C. O. (1985). Dumping of dangerous American products overseas: Should Congress sit and watch. *Journal of World Trade, 19*, 403.

Aizenman, J., & Jinjarak, Y. (2009). Globalisation and developing countries: A shrinking tax base? *Journal of Development Studies, 45*(5), 653–671.

Alemanno, A., & Bonadio, E. (2011). Do you mind my smoking? Plain packaging of cigarettes under the TRIPS Agreement. *John Marshall Review of Intellectual Property Law, 10*(3), 10:i.

Arbache, J. S., Dickerson, A., & Green, F. (2004). Trade liberalisation and wages in developing countries. *Economic Journal, 114*(493), F73–F96.

Bamou, E. (1999). *Trade liberalization and economic performance of Cameroon and Gabon*. Nairobi, Kenya: African Economic Research Consortium.

Baunsgaard, T., & Keen, M. (2010). Tax revenue and (or?) trade liberalization. *Journal of Public Economics, 94*, 563-577.

Becker, G. S., & Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of Labor Economics, 4*(3 pt 2), S1–S39.

Ben-David, D. (1993). Equalizing exchange: Trade liberalization and income convergence. *Quarterly Journal of Economics, 108*(3), 653–679.

Berger, J., & Kapczynski, A. (2009). The story of the TAC case: The potential and limits of socio-economic rights litigation in South Africa. In D. Hurwitz, M. Satterthwaite, & D. Ford (Eds.), *Human rights advocacy stories*. New York, NY: Thomson Reuters/Foundation Press.

Bettcher, D. W., Subramaniam, C., Guindon, E. G., Soll, L., Grabman, G., Joossens, L., . . . Initiative, W. T. F. (2003). Confronting the tobacco epidemic in an era of trade liberalization. Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/iris/handle/10665/42441

Bloom, D. E., & Canning, D. (2000). The health and wealth of nations. *Science, 287*(5456), 1207–1209.

Bloom, D., Canning, D., & Sevilla, J. (2001). *The effect of health on economic growth: Theory and evidence. National Bureau of Economic Research Working Paper,* No. 8587. Cambridge, MA: National Bureau of Economic Research.

Blouin, C., Chopra, M., & van der Hoeven, R. (2009). Trade and social determinants of health. *The Lancet, 373*, 502–507.

Blouin, C., Drager, N., & Smith, R. D. (2006). *International trade in health services and the GATS: Current issues and debates*. Washington DC: World Bank.

Bonn, D. (2003). Closing in on the cause of SARS. *The Lancet Infectious Diseases, 3*(5), 268.

Bown, C. P. (2017). Mega-regional trade agreements and the future of the WTO. *Global Policy, 8*(1), 107–112. doi:10.1111/1758-5899.12391

Broad, W. J. (1981). Export of hazardous goods raises congressional ire. *Science, 211*(4489), 1404–1405.

Buchan, J. (2006). Migration of health workers in Europe: Policy problem or policy solution. In C. A. Dubois, M. McKee, & E. Nolte (Eds.), *Human resources for health in Europe*: *European observatory on health systems and policies* (pp. 41–62). New York, NY: Open University Press.

Bureau, J. C., & Jean, S. (2013). The impact of regional trade agreements on trade in agricultural products. *OECD Food and Fisheries Papers, No. 65*. OECD Publishing. Retrieved from https://www.oecd-ilibrary.org/agriculture-and-food/the-impact-of-regional-trade-agreements-on-trade-in-agricultural-products_5k3xznkz60vk-en

Busfield, J. (2003). Globalization and the pharmaceutical industry revisited. *International Journal of Health Services, 33*(3), 581–605.

Bustamante, A. V. (2014). Globalization and medical tourism: The North American experience: Comment on "Patient Mobility in the Global Marketplace: A Multidisciplinary Perspective." *International Journal of Health Policy and Management, 3*(1), 47–49. doi:10.15171/ijhpm.2014.57

Callard, C., Chitanondh, H., & Weissman, R. (2001). Why trade and investment liberalisation may threaten effective tobacco control efforts. *Tobacco Control, 10*(1), 68–70.

Cassels, S. (2006). Overweight in the Pacific: Links between foreign dependence, global food trade, and obesity in the Federated States of Micronesia. *Global Health, 2*(1), 10.

Center for International Environmental Law. (2003). Second guessing national policy choices: Necessity, proportionality, and balanace in the WTO services negotiations. Retrieved from http://www.ciel.org/Publications/Necessity_3Sep03.pdf

Centers for Disease Control and Prevention (CDC). (2009). *Reported tuberculosis in the United States, 2008.* Atlanta, GA: Author.

Chaloupka, F., & Corbett, M. (1998). Trade policy and tobacco: Towards an optimal policy mix. In *The economics of tobacco control* (pp. 129–145). Cape Town, South Africa: Applied Fiscal Research Centre.

Chaloupka, F. J., & Laixuthai, A. (1996). *US trade policy and cigarette smoking in Asia*. Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w5543

Chaloupka, F. J., & Warner, K. E. (2000). The economics of smoking. *Handbook of Health Economics, 1*, 1539–1627.

Chanda, R. (2002). Trade in health services. *Bulletin of the World Health Organization, 80*(2), 158–163.

Chanda, R., & Smith, R. D. (2005). Trade in health services and GATS: A framework for policy makers. In C. Blouin, N. Drager, & R. Smith (Eds.). International trade in health services and the GATS: Current issues and debates. The World Bank, 2005.

Chanda, R. (2006). *Inter-modal linkages in services trade, OECD Trade Policy Papers,* No. 30. Paris, France: OECD Publishing. doi:10.1787/747731586224

Chanda, R. (2008). *Services trade and investment liberalisation, and domestic regulation: A summary of six country studies.* Jaipur, India: Consumer Unity & Trust Society.

Chandra, R. (2010). Indian professional and skilled migration to Australia and Singapore: A comparative study. *Journal of Sociology and Anthropology*, *1*(1–2), 71–80.

Chopra, M., Galbraith, S., & Darnton-Hill, I. (2002). A global response to a global problem: The epidemic of overnutrition. *Bulletin of the World Health Organization, 80*(12), 952–958.

Clark, S. E., Hawkes, C., Murphy, S. M. E., Hansen-Kuhn, K. A., & Wallinga, D. (2012). Exporting obesity: US farm and trade policy and the transformation of the Mexican consumer food environment. *International Journal of Occupational and Environmental Health, 18*(1), 53–64. doi:10.1179/1077352512Z.0000000007

Cornia, G. A. (2001). Globalization and health: Results and options. *Bulletin of the World Health Organization, 79*(9), 834–841.

Correa, C. M. (2006). Implications of bilateral free trade agreements on access to medicines. *Bulletin of the World Health Organization, 84*(5), 399–404.

Crawford, J.-A., & Fiorentino, R. V. (2004). *The changing landscape of regional trade agreements.* Geneva, Switzerland: World Trade Organization.

Crush, J., & Chikanda, A. (2015). South–South medical tourism and the quest for health in Southern Africa. *Social Science & Medicine, 124*, 313–320. doi:10.1016/j.socscimed.2014.06.025

Deaton, A. (2003). Health, inequality, and economic development. *Journal of Economic Literature, 41*(1), 113–158.

De Vogli, R., Kouvonen, A., & Gimeno, D. (2014). The influence of market deregulation on fast food consumption and body mass index: A cross-national time series analysis. *Bulletin of the World Health Organization, 92*(2), 99–107A.

Dollar, D. (1992). Outward-oriented developing economies really do grow more rapidly: Evidence from 95 LDCs, 1976–1985. *Economic Development and Cultural Change, 40*(3), 523–544.

Dollar, D. (2001a). *Globalization, inequality, and poverty since 1980.* Washington, DC: World Bank.

Dollar, D. (2001b). Is globalization good for your health? *Bulletin of the World Health Organization, 79*(9), 827–833.

Dollar, D., & Kraay, A. (2004). Trade, growth, and poverty. *Economic Journal, 114*(493), F22–F49.

Drewnowski, A., Hanks, A. S., & Smith, T. G. (2010). International trade, food and diet costs, and the global obesity epidemic. In C. Hawkes, C. Blouin, S. Henson, N. Drager, & L. Dube (Eds.), *Trade, food, diet and health: Perspectives and policy options* (pp. 77–90). Chichester, England: Wiley-Blackwell.

Eastwood, J. B., Conroy, R. E., Naicker, S., West, P. A., Tutt, R. C., & Plange-Rhule, J. (2005). Loss of health professionals from sub-Saharan Africa: The pivotal role of the UK. *The Lancet, 365*(9474), 1893–1900. doi:10.1016/S0140-6736(05)66623-8

Edwards, S. (1993). Openness, trade liberalization, and growth in developing countries. *Journal of Economic Literature, 31*(3), 1358–1393.

Escalante de Cruz, A., Phillips, S., Visch, M., & Bulan Saunders, D. (2004). *The junk food generation: A multi-country survey of the influence of television advertisements on children*. Kuala Lumpur, Malaysia: Consumers International Asia Pacific Office.

Evans, M., Sinclair, R. C., Fusimalohi, C., & Liava'a, V. (2001). Globalization, diet, and health: An example from Tonga. *Bulletin of the World Health Organization, 79*(9), 856–862.

Feachem, R. G. (2001). Globalisation is good for your health, mostly. *British Medical Journal, 323*(7311), 504–506.

Feinsilver, J. M. (2010). Fifty years of Cuba's medical diplomacy: From idealism to pragmatism. *Cuban Studies*, *41*(1), 85–104.

Fidler, D. P., Drager, N., & Lee, K. (2009). Managing the pursuit of health and wealth: The key challenges. *The Lancet, 373*(9660), 325–331. doi:10.1016/s0140-6736(08)61775-4

Fogel, R. W. (1994). The relevance of Malthus for the study of mortality today: Long-run influences on health, mortality, labor force participation, and population growth. Cambridge, MA: National Bureau of Economic Research.

Frankel, J. A., & Romer, D. (1999). Does trade cause growth? *American Economic Review, 89*(3), 379–399.

Frankel, J. A., & Rose, A. K. (2005). Is trade good or bad for the environment? Sorting out the causality. *Review of Economics and Statistics, 87*(1), 85–91.

Friel, S., Gleeson, D., Thow, A. M., Labonte, R., Stuckler, D., Kay, A., & Snowdon, W. (2013). A new generation of trade policy: Potential risks to diet-related health from the Trans Pacific Partnership agreement. *Global Health, 9*, 46. doi:10.1186/1744-8603-9-46

Ghose, A. K. (2003). *Jobs and incomes in a globalizing world.* Geneva, Switzerland: International Labour Organization.

Gilmore, A. B., & McKee, M. (2004). Moving east: How the transnational tobacco industry gained entry to the emerging markets of the former Soviet Union. Part I: Establishing cigarette imports. *Tobacco Control, 13*(2), 143–150.

Gleeson, D., & Friel, S. (2013). Emerging threats to public health from regional trade agreements. *The Lancet, 381*(9876), 1507–1509.

Glenday, G. (2002). Trade liberalization and customs revenue: Does trade liberalization lead to lower customs revenue? The case of Kenya. *Journal of African Finance and Economic Development, 5*(2), 89–125.

Glenday, G. (2006). *Towards fiscally feasible and efficient trade liberalization*. Durham, NC: Duke University. Retrieved from https://fds.duke.edu/db/attachment/828

Grise, V. N. (1990). *The world tobacco market: Government intervention and multilateral policy reform* (AGES 9014). Washington, DC: U.S. Department for Agriculture, Economic Research, Service Statistics Report.

Hagihara, A., & Takeshita, Y. J. (1995). Impact of American cigarette advertising on imported cigarette consumption in Osaka, Japan. *Tobacco Control, 4*(3), 239.

Hallam, D. (2009). *The state of agricultural commodity markets: High food prices and the food crisis: Experiences and lessons learned.* Rome, Italy: Food and Agriculture Organization.

Hanefeld, J. (2015). *Globalisation and health.* New York, NY: McGraw-Hill.

Hanefeld, J., Horsfall, D., Lunt, N., & Smith, R. (2013). Medical tourism: A cost or benefit to the NHS? *PLoS One, 8*(10), e70406. doi:10.1371/journal.pone.0070406

Hanefeld, J., Khan, M., Tomson, G., & Smith, R. (2017). Trade as central to achieving the Sustainable Development Goals: A case-study of antimicrobial resistance. *British Medical Journal, 358,* 3505. doi: 10.1136/bmj.j3505

Hanefeld, J., Lunt, N., Smith, R., & Horsfall, D. (2015). Why do medical tourists travel to where they do? The role of networks in determining medical travel. *Social Science & Medicine, 124*, 356–363. doi:10.1016/j.socscimed.2014.05.016

Hanefeld, J., Mandeville, K., & Smith, R. (2017). Making "health tourists" pay for care. *British Medical Journal, 356*. doi:10.1136/bmj.j771

Hawkes, C., & Thow, A. (2008). Implications of the Central America–Dominican Republic free trade agreement for the nutrition transition in Central America. *Revista Panamericana de Salud Pública, 24*, 345–360.

Hawkins, B., & Holden, C. (2016). A corporate veto on health policy? Global constitutionalism and investor–state dispute settlement. *Journal of Health Politics, Policy and Law, 41*(5), 969–995.

Helble, M. (2011). The movement of patients across borders: Challenges and opportunities for public health. *Bulletin of the World Health Organization, 89*(1), 68–72. doi:10.2471/Blt.10.076612

Hoekman, B. M. (2005). *Trade and employment: Stylized facts and research findings* (Vol. 3676). Washington, DC: World Bank Publications.

Hsieh, C. R., Hu, T. W., & Lin, C. F. J. (1999). The demand for cigarettes in Taiwan: Domestic versus imported cigarettes. *Contemporary Economic Policy, 17*(2), 223–234.

Hughes, R., & Lawrence, M. (2005). Globalization, food and health in Pacific Island countries. *Asia Pacific Journal of Clinical Nutrition, 14*, 298–306.

Hunt, P. (2007). Right to the highest attainable standard of health. *The Lancet, 370*(9585), 369–371. doi:10.1016/s0140-6736(07)61173-8

Jones, E. (2007). *Signing away the future: How trade and investment agreements between rich and poor countries undermine development.* Oxford, UK: Oxfam International.

Kanchanachitra, C., Lindelow, M., Johnston, T., Hanvoravongchai, P., Lorenzo, F. M., Huong, N. L., . . . dela Rosa, J. F. (2011). Human resources for health in southeast Asia: Shortages, distributional challenges, and international trade in health services. *The Lancet, 377*(9767), 769–781. doi:10.1016/S0140-6736(10)62035-1

Keogh-Brown, M. R., & Smith, R. D. (2008). The economic impact of SARS: How does the reality match the predictions? *Health Policy, 88*(1), 110–120.

Khattry, B., & Rao, J. M. (2002). Fiscal faux pas? An analysis of the revenue implications of trade liberalization. *World Development, 30*(8), 1431–1444.

Kickbusch, I., Hein, W., & Silberschmidt, G. (2010). Addressing global health governance challenges through a new mechanism: The proposal for a Committee C of the World Health Assembly. *Journal of Law, Medicine, and Ethics, 38*(3), 550–563.

Kingma, M. (2007). Nurses on the move: A global overview. *Health Services Research, 42*(3 pt 2), 1281–1298.

Labonté, R., Schram, A., & Ruckert, A. (2016). The Trans-Pacific Partnership: Is it everything we feared for health? *International Journal of Health Policy and Management, 5*(8), 487.

Labonté, R., Schram, A., & Ruckert, A. (2017). The TPP is dead, long live the TPP? A response to recent commentaries. *International Journal of Health Policy and Management, 6*(4), 245–247. doi:10.15171/ijhpm.2017.12

Laird, S., Vanzetti, D., & Fernandez de Cordoba, S. (2005). *Smoke and mirrors: Making sense of the WTO industrial tariff negotiations* (UNCTAD Blue Series Papers 30). Herndon, VA: United Nations Publications. Retrieved from itcdtab31_en.pdf

Lall, S. (2004). The employment impact of globalization in developing countries. In E. Lee & M. Vivarelli (Eds.), *Understanding globalization, employment and poverty reduction* (pp. 73–101). New York, NY: Palgrave Macmillan.

Lautier, M. (2008). Export of health services from developing countries: The case of Tunisia. *Social Science & Medicine, 67*(1), 101–110. doi:10.1016/j.socscimed.2008.01.057

Lee, K., Sridhar, D., & Patel, M. (2009). Bridging the divide: Global governance of trade and health. *The Lancet, 373*(9661), 416–422.

Lee, K., Yach, D., & Kamradt-Scott, A. (2011). Globalization and health. In M. Merson, R. Black, & A. Mills (Eds.), *Global public health: Diseases, programs, systems and policies,* (3rd ed., pp. 885–913). Burlington, MA: Jones & Bartlett.

Levine, D. I., & Rothman, D. (2006). Does trade affect child health? *Journal of Health Economics, 25*(3), 538–554.

Liverani, M., & Coker, R. (2012). Protecting Europe from diseases: From the International Sanitary Conferences to the ECDC. *Journal of Health Politics, Policy and Law, 37*(6), 915–934. doi:10.1215/03616878-1813772

Lunt, N., Horsfall, D., Exworthy, M., Smith, R., & Hanefeld, J. (2014). Market size, market share and market strategy: Three myths of medical tourism. *Policy and Politics, 42*(4), 597-614.

Lunt, N., Horsfall, D., & Hanefeld, J. (2015). *Handbook on medical tourism and patient mobility.* Northampton, MA: Edward Elgar Publishing.

MacDonald, R., & Horton, R. (2009). Trade and health: time for the health sector to get involved. *The Lancet, 373*(9660), 273–274.

Maneschi, A. (1998). *Comparative advantage in international trade: A historical perspective.* Northampton, MA: Edward Elgar Publishing.

Marmot, M. (2000). *Psychoscoial and biological mechanisms behind the recent mortality crisis in Central and Eastern Europe.* Oxford, UK: Oxford University Press.

Mattoo, A., & Fink, C. (2004). Regional agreements and trade in services: Policy issues. *Journal of Economic Integration*, *19*(4),742–779.

Mattoo, A., Rathindran, R., & Subramanian, A. (2001). *Measuring services trade liberalization and its impact on economic growth: An illustration* (Vol. 2655). Washington, DC: World Bank Publications.

McGrady, B. (2004). TRIPs and trademarks: The case of tobacco. *World Trade Review, 3*(1), 53–82.

McGrady, B. (2007). Trade liberalisation and tobacco control: Moving from a policy of exclusion towards a more comprehensive policy. *Tobacco Control, 16*(4), 280–283. doi:10.1136/tc.2006.019141

McKee, M., Greer, S. L., & Stuckler, D. (2017). What will Donald Trump's presidency mean for health? A scorecard. *The Lancet, 389*(10070), 748–754.

Milanovic, B., & Squire, L. (2005). *Does tariff liberalization increase wage inequality? Some empirical evidence.* Cambridge, MA: National Bureau of Economic Research.

Moon, S., Sridhar, D., Pate, M. A., Jha, A. K., Clinton, C., Delaunay, S., . . . Piot, P. (2015). Will Ebola change the game? Ten essential reforms before the next pandemic. The report of the Harvard–LSHTM Independent Panel on the Global Response to Ebola. *The Lancet, 386*(10009), 2204–2221. doi:10.1016/S0140-6736(15)00946-0

Mundial, B. (2005). DR-CAFTA: Challenges and opportunities for Central America. Washington, DC: World Bank Publications.

NaRanong, A., & NaRanong, V. (2011). The effects of medical tourism: Thailand's experience. *Bulletin of the World Health Organization, 89*(5), 336–344.

Naylor, C. D. (2003). *Learning from SARS: Renewal of public health in Canada: A report of the National Advisory Committee on SARS and Public Health.* Ottawa, Canada: National Advisory Committee.

Nestle, M. (2006). Food marketing and childhood obesity: A matter of policy. *New England Journal of Medicine, 354*(24), 2527–2529.

Noree, T., Hanefeld, J., & Smith, R. (2016). Medical tourism in Thailand: A cross-sectional study. *Bulletin of the World Health Organization, 94*(1), 30–36. doi:10.2471/blt.14.152165

Noree, T., Smith, R., & Hanefeld, J. (2014). UK medical tourists in Thailand: They are not who you think they are. *Globalisation and Health, 10*(29),10–29.

Office of the United States Trade Representative. (2007). *Special 301 report.* Washington DC: Author.

O'Neill, A. (2016). Review on antimicrobial resistance, tackling drug resistant infections globally. In *Infection prevention, control and surveillence: Limiting the development and spread of drug resistance* (pp. 1–32). Retrieved from https://amr-review.org/sites/default/files/160518_Final%20paper_with%20cover.pdf

Onzivu, W. (2002). The public health implications of the Association of South East Asian Nations (ASEAN) legal regime on tobacco control. *Australian Journal of Asian Law, 4*(2), 160.

Onzivu, W. (2006). Globalism, regionalism, or both: Health policy and regional economic integration in developing countries, an evolution of a legal regime. *Minnesota Journal of Internation Law, 15*, 111.

Organisation for Economic Co-operation and Development (OECD). (2003). *Regionalism and the multilateral trading system.* Paris, France: Author.

Ormond, M. (2013). *Neoliberal governance and international medical travel in Malaysia*. London, UK: Routledge.

Peterson, E. W. F. (2009). *A billion dollars a day: The economics and politics of agricultural subsidies.* Hoboken, NJ: John Wiley & Sons.

Plotnikova, E. (2014). The role of bilateral agreements in the regulation of health worker migration. In J. Buchan, M. Wismar, I. A. Glinos, & J. Bremner (Eds.), *Health professional mobility in a changing Europe: New dynamics, mobile individuals and diverse responses, Vol II* (pp. 325–344). Copenhagen, Denmark: World Health Organization.

Pollock, A. M., & Price, D. (2000). Rewriting the regulations: How the World Trade Organisation could accelerate privatisation in health-care systems. *The Lancet, 356*(9246), 1995–2000.

Popkin, B. M. (2001). The nutrition transition and obesity in the developing world. *Journal of Nutrition, 131*(3), 871S–873S.

Preston, S. H. (1975). The changing relation between mortality and level of economic development. *Population Studies, 29*(2), 231–248.

Pritchett, L., & Summers, L. H. (1996). Wealthier is healthier. *Journal of Human Resources*, *31*(4), 841–868.

Ravallion, M. (2006). Looking beyond averages in the trade and poverty debate. *World Development, 34*(8), 1374–1392.

Roemer-Mahler, A. (2010). Business strategy and access to medicines in developing countries. *Global Health Governance, IV*(1), 1939–2389.

Ruckert, A., Schram, A., Labonté, R., Friel, S., Gleeson, D., & Thow, A.-M. (2017). Policy coherence, health and the Sustainable Development Goals: A health impact assessment of the Trans-Pacific Partnership. *Critical Public Health, 27*(1), 86–96.

Sachs, J. D. (2002). Macroeconomics and health: Investing in health for economic development. *Revista Panamericana de Salud Pública, 12*(2), 143–144.

Sachs, J. D., Warner, A., Åslund, A., & Fischer, S. (1995). Economic reform and the process of global integration. *Brookings Papers on Economic Activity, 1*, 1–118.

Saliba, V., Legido-Quigley, H., Hallik, R., Aaviksoo, A., Car, J., & McKee, M. (2012). Telemedicine across borders: A systematic review of factors that hinder or support implementation. *International Journal of Medical Informatics, 81*(12), 793–809. doi:10.1016/j.ijmedinf.2012.08.003

Salmon, M. E., Yan, J., Hewitt, H., & Guisinger, V. (2007). Managed migration: The Caribbean approach to addressing nursing services capacity. *Health Services Research, 42*(3 pt 2), 1354–1372.

Scheve, K., & Slaughter, M. J. (2004). Economic insecurity and the globalization of production. *American Journal of Political Science, 48*(4), 662–674.

Shaffer, E., Brenner, J., & Houston, T. (2005). International trade agreements: A threat to tobacco control policy. *Tobacco Control, 14*(suppl 2), ii19–ii25.

Singhal, P. (2005). Life cycle environmental issues of mobile phones. In *Stage I final report: Integrated product policy pilot project* (pp. 16-39). Espoo, Finland: Nokia.

Smith, R. D. (2003). Global public goods and health. *Bulletin of the World Health Organization, 81*(7), 475–475.

Smith, R. D., Chanda, R., & Tangcharoensathien, V. (2009). Trade in health-related services. *The Lancet, 373*(9663), 593–601. doi:S0140-6736(08)61778-X

Smith, R. D., Correa, C., & Oh, C. (2009). Trade, TRIPS, and pharmaceuticals. *The Lancet, 373*. doi:10.1016/S0140-6736(08)61779-1

Smith, R. D., & Hanson, K. (2012). *Health systems in low-and middle-income countries: An economic and policy perspective.* Oxford, UK: Oxford University Press.

Smith, R. D., Lee, K., & Drager, N. (2009). Trade and health: An agenda for action. *The Lancet, 373*(9665), 768–773. doi:10.1016/s0140-6736(08)61780-8

Smith, R. J. (1980). Hazardous wastes cause international stink. *Science, 207*(4434), 962–963.

Songkhla, M. N. (2009). Health before profits? Learning from Thailand's experience. *The Lancet, 373*(9662), 441.

Stilwell, B., Diallo, K., Zurn, P., Vujicic, M., Adams, O., & Dal Poz, M. (2004). Migration of health-care workers from developing countries: Strategic approaches to its management. *Bulletin of the World Health Organization, 82*(8), 595–600.

Stuckler, D., McKee, M., Ebrahim, S., & Basu, S. (2012). Manufacturing epidemics: The role of global producers in increased consumption of unhealthy commodities including processed foods, alcohol, and tobacco. *PLoS Medicine, 9*(6), e1001235.

Subbarao, K., & Raney, L. (1995). Social gains from female education: A cross-national study. *Economic Development and Cultural Change, 44*(1), 105–128.

Suhrcke, M., McKee, M., Arce, R. S., Tsolova, S., & Mortensen, J. (2006). Health economics: Investment in health could be good for Europe's economies. *British Medical Journal, 333*(7576), 1017.

Suzana, M., & Chongsuvivatwong, V. (2015). Changes in access to and costs of overseas treatment after the introduction of universal health coverage in the Maldives: Findings from surveys and the analysis of claims data. *Health Services Management Research, 28*(3–4), 47–57. doi:10.1177/0951484815611147

Tenenbaum, D. J. (2008). Food vs. fuel: Diversion of crops could cause more hunger. *Environmental Health Perspectives, 116*(6), A254.

Thomas, B. P., & Gostin, L. O. (2013). Tobacco endgame strategies: Challenges in ethics and law. *Tobacco Control, 22*(suppl 1), i55–i57.

Thow, A. (2009). Trade liberalisation and the nutrition transition: Mapping the pathways for public health nutritionists. *Public Health Nutrition, 12*, 2150–2158.

Thow, A., & Hawkes, C. (2009). The implications of trade liberalization for diet and health: A case study from Central America. *Global Health, 5*, 5.

Thow, A., Labonte, R., Gleeson, D., Stuckler, D., Hattersley, L., Kay, A., & Friel, S. (2014). *Will the next generation of preferential trade and investment agreements undermine implementation of the Global Action Plan for Prevention and Control of Noncommunicable Diseases?* Presentation at 142nd Americal Public Health Association Annual Meeting and Exposition, New Orleans, LA.

Thow, A., & Snowdon, W. (2010). The effect of trade and trade policy on diet and health in the Pacific Islands. In C. Hawkes, C. Blouin, S. Henson, N. Drager, & L. Dube (Eds.), *Trade, food, diet and health: Perspectives and policy options* (pp. 147–168). Chichester, England: Blackwell Publications.

United Nations Conference on Trade and Development (UNCTAD). (1995). *World investment report: Transnational corporations and competitiveness.* New York, NY/Geneva, Switzerland: Author.

Van Doosselaere, C., Wilson, P., Herveg, J., & Silber, D. (2007). eHealth . . . But is it legal? *Eurohealth–London, 13*(2), 1.

Vearey, J., de Gruchy, T., Kamndaya, M., Walls, H. L., Chetty-Makkan, C. M., & Hanefeld, J. (2018). Exploring the migration profiles of primary healthcare users in South Africa. *Journal of Immigrant and Minority Health*, *20*(1), 91–100.

Walls, H. L., Smith, R. D., & Drahos, P. (2015). Improving regulatory capacity to manage risks associated with trade agreements. *Global Health, 11*(1), 14. doi:10.1186/s12992-015-0099-7

Walls, H. L., Vearey, J., Modisenyane, M., Chetty-Makkan, C. M., Charalambous, S., Smith, R. D., & Hanefeld, J. (2016). Understanding healthcare and population mobility in southern Africa: The case of South Africa. *South African Medical Journal*, *106*(1):14–15. doi:10.7196/SAMJ.2016.v106i1.10210

Weissman, R. (2003). International trade agreements and tobacco control: Threats to public health and the case for excluding tobacco from trade agreements. *Essential Action*, 1–28. Retrieved from http://www.takingontobacco.org/trade/tobacco.trade.v02.backgrd.pdf

Whittaker, A. (2010). Challenges of medical travel to global regulation: A case study of reproductive travel in Asia. *Global Social Policy, 10*(3), 396–415.

Whittaker, A. (2011). Cross-border assisted reproduction care in Asia: Implications for access, equity and regulations. *Reproductive Health Matters, 19*(37), 107–116.

Winters, L. A. (2004). Trade liberalisation and economic performance: An overview. *Economic Journal, 114*(493), F4-F21.

Winters, L. A. (2006). International trade and poverty: Cause or cure? *Australian Economic Review, 39*(4), 347–358.

Winters, L. A., McCulloch, N., & McKay, A. (2004). Trade liberalization and poverty: The evidence so far. *Journal of Economic Literature, 42*(1), 72–115.

Wood, A., & Ridao-Cano, C. (1999). Skill, trade, and international inequality. *Oxford Economic Papers, 51*(1), 89–119.

Woodward, D., Drager, N., Beaglehole, R., & Lipson, D. (2001). Globalization and health: A framework for analysis and action. *Bulletin of the World Health Organization, 79*(9), 875–881.

World Health Organization (WHO). (2002). Global crises, global solutions: Managing public health emergencies of international concern through the revised International Health Regulations. Retrieved from www.who.int/csr/resources/publications/ihr/whocdsgar20024.pdf?ua=1

World Health Organization (WHO). (2016). *Global tuberculosis report 2016*. Geneva, Switzerland: Author.

World Trade Organization (WTO). (2017a). Regional trade agreements: Facts and figures. Retrieved from https://www.wto.org/english/tratop_e/region_e/regfac_e.htm

World Trade Organization (WTO). (2017b). WTO IP rules amended to ease poor countries' access to affordable medicines. Retrieved from https://www.wto.org/english/news_e/news17_e/trip_23jan17_e.htm

Yagi, N., Mackey, T. K., Liang, B. A., & Gerlt, L. (2014). Policy review: Japan–Philippines Economic Partnership Agreement (JPEPA): Analysis of a failed nurse migration policy. *International Journal of Nursing Studies, 51*(2), 243–250.

CHAPTER 21

Global Health Governance and Diplomacy

Kelley Lee

Introduction

There is wide-ranging evidence that human societies worldwide, along with the natural environment, are being fundamentally altered by processes collectively referred to as globalization. Although globalization has been unfolding for many centuries, if not longer, it is broadly recognized that the changes occurring since the late 20th century have been faster paced, farther reaching, and more profound in their impacts (Beck, 2015; Scholte, 2005; Stiglitz, 2007).

Within the public health community, understanding of globalization and its population health consequences has led to a shift in attention from *international* health to *global* health. Many of the chapters in this text concern the direct impacts of globalization on public health. For example, there is clear evidence of new patterns of communicable disease outbreaks since 1980, including a four-fold increase in their number and a 20% rise in the specific types of diseases causing them (Smith et al., 2014). Similarly, a shifting and growing burden from noncommunicable diseases (NCDs) is rapidly emerging in populations across the world (see the *Chronic Diseases and Risks* chapter). These new epidemiologic patterns have created a "double burden" in many countries, as they struggle to cope with both communicable and noncommunicable diseases, as well as new distributions of health and disease for populations that cut across national borders (Lewis & MacPherson, 2013; World Health Organization [WHO], 2009).

The need to better understand how globalization is affecting population health has, in turn, opened up new avenues of inquiry and action. Foremost are renewed efforts to identify and explain the social determinants of health related to globalization, such as changes in the world economy, patterns of migration, changes to natural and human-built environments, and trends in human conflict. There is now a thriving body of scholarship conceptualizing and empirically demonstrating these complex linkages between globalization and population health as a precursor to strengthening public health policies (Huynen, Martens, & Hilderink, 2005; Labonte & Schrecker, 2007; Martens, Akin, Maud, & Mohsin, 2010). Overall, global health research, policy, and practice have made significant advances since the late 1990s, with key areas of focus being diagnoses of the changing public health problems faced and the likely causes posed by globalization (Birn, Pillay, & Holtz, 2017).

Alongside this growing evidence of the direct impacts of globalization, and awareness of the linkages between the social determinants of health and

globalization, there have been efforts to understand the collective actions needed to better protect and promote population health in a globalizing world. What should be the agreed and shared responses by societies to the health impacts of globalization? How can these responses be best achieved within and across societies?

This chapter is about the forms of collective action needed to address public health issues arising from globalization, broadly known as global health governance (GHG), and the best ways to achieve them. GHG concerns the institutional arrangements—in the form of rules, responsibilities, and processes—for achieving agreed goals related to global health needs. This chapter begins by briefly setting out clear definitions for two concepts underlying GHG: globalization and global health. While widely used, these terms are often poorly defined, which in turn contributes to confusion regarding the needed scope and nature of collective action. Not all changes taking place since the late 20th century have been due to globalization. Not all public health issues are global in nature. Moreover, while globalization is having profound impacts on a vast array of health determinants and outcomes, not all public health issues require GHG. A clear understanding of these concepts, therefore, helps to define the parameters of this chapter.

Focusing on GHG, this chapter then describes the existing institutional arrangements that structure collective responses to global health needs. These arrangements are located within a historical background of international health cooperation, established after World War II, as well as contemporary efforts to transition to new arrangements that enable new forms of collective action. This discussion includes the limitations of existing institutions, in terms of governance gaps. In this context, global health diplomacy (GHD) is understood as a key feature of GHG. GHD concerns emerging processes of political negotiation and consensus building that enable the creation of the institutional arrangements needed for GHG.

Finally, this chapter reviews the enduring challenges faced by the global health community in strengthening GHG. These challenges stem from the diverse interests, ideas, and institutions shaping global health and the invariably political nature of governance. Importantly, these challenges are not intractable. The changing context within which GHD takes place, including the changing roles of particular state and non-state actors, brings new opportunities. The chapter concludes by discussing ways that institutional innovation is beginning to facilitate the transition from international to global health governance. This paradigm change requires the global health community to navigate between longstanding efforts to reform the World Health Organization, which have yielded limited success, and the creation of an entirely new institutional architecture, which has been hindered by a lack of political consensus on what this structure should look like.

Globalization and Global Health Needs

The widely used term "globalization" articulates a shared sense that the contemporary world, beginning in the last decades of the 20th century, has become more interconnected, with events in one part of the world having potentially far-reaching consequences elsewhere. This phenomenon has spawned a vast and multidisciplinary scholarly literature, spanning the natural and applied sciences, social sciences, arts, and humanities, characterized by diverse theoretical perspectives, normative frameworks, and empirical debates about whether globalization is really happening, what the key drivers are, and, perhaps most controversially, whether it is having positive or negative impacts on human societies and the natural world.

It is beyond the scope of this chapter to review this substantial and diverse literature with its many points of contestation. At the same time, if we are to understand the collective action needed to manage the health impacts of globalization, a clear and shared understanding of the nature of the changes occurring, as a distinct phenomenon, is an essential starting point. To achieve this goal, we can begin by locating globalization within a historical context.

Although the "globalization" term has been coined relatively recently, contemporary shifts can be linked to social changes over the course of centuries, if not millennia. The earliest beginnings of globalization can be traced to the migration of *Homo erectus* out of the African continent 1 million years ago, or perhaps more specific to modern humans, the migration of *Homo sapiens* around 80,000 years ago. The latter hominids eventually colonized much of the planet, forming societies that interacted with other societies across ever-increasing distances and with greater frequency. Globalization, in this historical sense, can be understood as a continuum from lesser to greater social formation and integration across societies over time. Earliest human societies were circumscribed in membership and territory, infrequently interacting with each other, and often competing for available resources. Over time, the invention and use of new modes of transportation (from horse-drawn carts to ships and railroads and then cars and airplanes) and

communication (from papyrus scrolls to telegraphy to the Internet) enabled people to connect across greater distances in ever-larger numbers. Social, political, and economic history is correspondingly marked by events denoting this gradual interconnectedness of human societies—the opening of the Silk Road between Asia and Europe, the arrival of Christopher Columbus in the Americas, the formation of the modern state system, the colonial empires of European powers, the establishment of the slave trade, and the Industrial Revolution. Characterizing all of these developments was the increased mobility of people (voluntary or otherwise) and other life forms (plants and animals); flows of capital, goods, and services; and dissemination and exchange of knowledge and ideas. Thus, what many refer to as contemporary globalization can be understood as the acceleration and intensification of a historical process resulting in even more frequent and deeper interactions across human societies.

Yet, while globalization can be understood as a historical process, it is also necessary to identify what is distinct about the term. The work of Jan Aarte Scholte (2008) is helpful in this respect. Scholte begins by arguing that the term "globalization" is often used when people really mean something else. He identifies this "something else" to be several existing terms—internationalization, liberalization, universalization, and Westernization—that are not synonymous with globalization, but rather "redundant concepts of globalization." Each of these terms already has an existing and particular meaning:

- *Internationalization.* This is the most commonly used meaning attached to the term "globalization," referring to the process of closer interaction and exchange between people located in different countries. Various cross-border measures, such as trade and investment, communications, and migration, are cited as evidence of the increase in exchanges across national borders. Because the focus of these interactions and exchanges is how they occur across countries (or nations), the term "internationalization" is the appropriate term. Historically, since the establishment of the international system of sovereign states some 500 years ago, there have been a growing scope and frequency of economic, political, and social connections across groups of two or more countries.
- *Liberalization.* This is another commonly used meaning attached to "globalization." Liberalization concerns the process, in a world composed of sovereign states, of lessening government "regulatory barriers to transfers of resources between countries." Historically, the various rounds of negotiations under the General Agreement on Tariffs and Trade (GATT), spanning from the end of World War II to 1995 and including the creation of the World Trade Organization (WTO), led to successive reductions in trade tariffs. The continuation of this process since the mid-1990s, under multilateral, regional, and bilateral trade and investment agreements, has frequently been referred to as globalization. However, Scholte argues that the term "liberalization" is more accurate, with "little need now to invent a new vocabulary for this old phenomenon."
- *Universalization.* This term, which refers to the process of a social phenomenon spreading to all parts of the world, is often the intended meaning of "globalization" for some users of this term. When defined in this way, however, it might be argued that many examples of universalization predate globalization. Several world religions, for example, have won followers worldwide over the centuries. In this sense, Scholte argues that the term "universalization" is deemed adequate.
- *Westernization.* This term describes the process by which culture, values, ideas, and behaviors characteristic of Western societies, and particularly American society, are increasingly adopted in other countries worldwide. In this sense, globalization refers to a largely negative process of cultural colonization, and even homogenization, through a steady diet of American media, fashion, language, and consumer products. Although it is undoubtedly the case that Westernization is taking place, Scholte (2008) argues that "intercontinental westernization . . . has unfolded since long before the recent emergence of globe-talk." He suggests that the concepts of modernization and imperialism readily capture the ideas of Westernization and "[w]e do not need a new vocabulary of globalization to remake an old analysis."

Scholte (2008) casts aside these "redundant" terms and reserves the term "globalization" to describe social phenomena that do not simply cross national borders, but rather transcend them. Only when territorial boundaries, based on physical geography, are circumvented or become irrelevant can we speak of globalization. Satellite communications, climate change, social media, illicit drug trafficking, and undocumented migration are examples of globalization in this strict sense. It is this strict definition of globalization, as the accelerated transcendence of territorial boundaries, that distinguishes the changes being experienced across human societies since the late 20th century.

Thus, globalization is a historical process characterized by changes in the nature of human interaction across a range of social spheres, including the economic, political, technological, cultural, and environmental arenas. These changes are globalizing in the sense that boundaries hitherto separating us from each other are being transformed. These boundaries—spatial, temporal, and cognitive—can be described as the dimensions of globalization. Briefly, the spatial dimension concerns changes in how we perceive and experience physical space or geographical territory, the temporal dimension concerns changes in how we perceive and experience time, and the cognitive dimension concerns changes in how we think about ourselves and the world around us (Lee, 2003).

The loose and varied uses of the term "globalization" in public health have led to similar imprecision in the application of the term "global health." The diverse definitions of global health, in tandem with its prevalent use, risk rendering this term meaningless. Although such definitions abound, many lack conceptual rigor. As a result, the term "global health" is also now associated with a broad range of other meanings. Some are redundant, given preexisting terms with the same meaning, such as health cooperation between governments (international health) or health in low- and middle-income countries (LMICs) (health development). Others are based on seeing the world from a self-centric (largely Western) viewpoint, such that domestic health issues in one's home country are considered separate from those of all other countries. The problem with adopting this perspective, in which "global" is viewed as everywhere but home, is that global health becomes completely relative. Moreover, this view supports the "othering" of nondomestic populations, and overlooks the key defining feature of globalization—namely, the closer linkages and interconnections among populations across territorial space. Other definitions focus on normative goals such as reducing differences in health status or distribution of health determinants caused by an increasingly globalized world (health inequalities) or the systematic, socially produced and unfair causes of health inequalities (health inequities). Thus, the understanding of the term "global health" can depend as much on normative frameworks and aspirations as it does material reality.

In sum, "global health" has become an umbrella term, subsuming previously used terms describing existing health needs, as well as embracing emerging kinds of health issues arising from contemporary globalization. It is important to recognize and interrogate these varying definitions as a key part of the task of strengthening global health governance and diplomacy.

An alternative approach is to focus on the nature or characteristics of a health issue, and then to categorize it as a certain type of "global health" problem requiring collective action. Building on the previously described distinction between globalization, as strictly defined, and other common uses of the term, we can arrive at a clearer understanding of the health needs arising within an increasingly globalized world that require collective action. Within the public health community, there has been a broad shift in nomenclature between the terms "international health" and "global health." The reasons for this shift, and the conceptual difference between the two terms, have elicited much discussion. Brown et al. (2006) locate the historical roots of international health in 19th century European imperialism, when the priority for collective health action was controlling epidemic diseases spreading from colonized territories. In contrast, global health encompasses a shift to putting "the health needs of the people of the whole planet above the concerns of particular nations." Koplan et al. (2009) distinguish among public (or population), international (interstate) health, and global health, with the last variably "thought of as a notion (the current state of global health), an objective (a world of health people, a condition of global health), or a mix of scholarship, research, and practice (with many questions, issues, skills, and competencies)."

Following Scholte's work, a strict definition of global health would limit the term to issues characterized by the transcendence of territorial boundaries. The issue must be caused by factors that transcend territorial geography, or possess the capability to occur or spread in ways that do so—that is, it must be "transboundary"—to be classified as a global health problem. Needless to say, not every health problem qualifies. Guinea worm infection, for example, remains a residual health threat in many parts of Africa. By comparison, significant political and financial capital has been expended on human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS). The distinction drawn might be explained, at least in part, by the fact that the former is transmitted via an animal vector that is prevalent only in certain parts of the world. The other is a disease spread by close contact with infected bodily fluids, an event that can occur in any part of the world.

Another good example is communicable diseases having the potential to become pandemic. The migration of wild birds across territorial boundaries on a regional and even planetary scale, for instance, makes avian influenza a global health issue. The increased distribution of the *Aedes aegypti* mosquito over the past 25 years, due to global climate change, local environmental

conditions, and increased transport links (e.g., ship, airplanes) worldwide (European Centre for Disease Prevention and Control, n.d.), means that associated diseases such as yellow fever, dengue, and chikungunya should be considered global health issues.

The large-scale trade in illicit goods such as counterfeit medicines, tobacco products, and narcotics, enabled by a worldwide network of criminal organizations, represents another example of a global health issue. Likewise, the use of social media—whose content is accessible worldwide and pose challenges for national authorities to regulate—to promote unhealthy lifestyles and health-harming products can be considered a global health issue.

Focusing on health determinants, lung cancer and liver cirrhosis could be described as global health issues because their causal factors are transboundary (i.e., transnational tobacco and alcohol companies). In contrast, the nutritional taboos practiced in some cultures, which may lead to nutritional deficiencies in pregnant women and children, pose a serious problem but are not a global health issue by virtue of their localized practice.

Collectively, these examples, when a stricter definition of globalization is applied, suggest that there are relatively few truly global health issues. In this chapter, by balancing the stricter and looser definitions of globalization and global health, we can describe health issues requiring collective action in a globalizing world as falling into four main categories (Lee, Hawkins, & Wiist, 2016):

- *Cross-border or transborder problems* originate in one locale but have wider ramifications for other locales (e.g., large-scale communicable disease outbreaks; advertising, marketing, and promotion of health-harming products via social media).
- *Commons problems* concern disputed access to or use of shared resources across societies (e.g., intellectual property rights over genetic sequencing, access to vaccines and other essential medicines).
- *Shared problems* are widely experienced across geographies due to changes to social and natural environments from intensifying globalization (e.g., increased popularity of computer gaming and childhood obesity, lack of safe water and sanitation).
- *Planetary problems* arise from the cumulative actions of individuals and populations in many countries (e.g., spread of antibiotic resistance, impact of climate change on distribution of disease vectors).

The forms of collective action needed to address these categories of global health issues are discussed in the next section.

What Is Global Health Governance?

Governance concerns the agreed rules for interaction and procedures for decision making adopted by an organized society to address its collective needs and deliver shared solutions. The agreed rules and procedures relate to such considerations as membership in the society, obligations and responsibilities of members, forms of engagement and decision making, mobilization and distribution of resources, dispute settlement, and implementation and evaluation of actions taken. Defined in this way, governance should be recognized as occurring across many types, and at many different levels, of social organization (Dodgson, Lee, & Drager, 2002). In global health, governance is essential to groups ranging from local community groups (e.g., patient advocacy groups, community health centers) to globally operating organizations (e.g., global public–private partnerships, transnational corporations).

Importantly, as Rosenau (1995, p. 4) writes, governance is not the same as government:

> Both refer to purposive behavior, to goal oriented activities, to systems of rule; but government suggests activities that are backed by formal authority . . . whereas governance refers to activities backed by shared goals that may or may not derive from legal and formally prescribed responsibilities and that do not necessarily rely on police powers to overcome defiance and attain compliance.

Government, in other words, is a particular form of governance, where a recognized body applies agreed rules and procedures in ways that are binding on the society's members. The formation of the international states system during the mid-17th century, for example, is based on the principle that national governments are sovereign, having ultimate authority over their own territories and populations. Above the state, where formally binding authority exists to a more limited degree, collective action is heavily reliant on governance. International law exists but is circumscribed in scope and authority. Where formal authority to govern does not exist, compliance is based on custom, common law, cultural norms and values, and public opinion.

Health governance, in turn, can be defined as the agreed rules for interaction and procedures for decision making adopted by an organized society to address collective health needs and deliver shared health solutions. Such rules and procedures can be

binding, such as various types of health-related legislation adopted and enforced by governments, or they can be reliant on nonbinding mechanisms (e.g., Hippocratic Oath, resolutions, codes of practice). Historically, the locus of health governance has been at the national level (and delegated to the subnational level), as the governments of sovereign states have held primary responsibility for protecting and promoting the health and well-being of their own populations. Where health determinants and outcomes spill over national borders and create cross-border or transborder impacts, international (intergovernmental or interstate) governance mechanisms have facilitated cooperation between two or more governments.

Indeed, there is a long and established history of *international health governance*. For example, the recurrence of bubonic plague, which caused as many as 50 million deaths during the 6th to 8th centuries, resulted from burgeoning trade links between Asia and Europe. The disease was spread by rat-infested ships along the Mediterranean coast and then inland via transported goods. The disease returned during the 14th century, transported by trade from Asia; it reached first Italy and then the rest of Europe, eventually killing one-third of the continent's population. The outbreaks led to closer cooperation among ruling authorities, including the introduction of quarantines (Cohn, 2002; Watts, 2003).

Similarly, the transition of cholera from an endemic disease limited to South Asia, to a pandemic disease reaching Europe in the 1830s, and then recurring on various continents over the next two centuries, was initially caused by the social, political, and economic upheavals of European imperialism (Lee & Dodgson, 2000). The cholera epidemics of the 19th century prompted European countries and the United States to hold a series of International Sanitary Conferences and adopt the International Sanitary Conventions (forerunner of the International Health Regulations). The formation of the League of Nations Health Organization in 1919 after World War I, amid the devastating 1918–1920 influenza pandemic, and then the founding of WHO after World War II, are other key examples of institutional arrangements for international health governance led by sovereign states (for more on this history, see the Introduction to this text).

The term *global health governance* was coined in the late 1990s to describe the diversity of governance mechanisms that go beyond intergovernmental health cooperation (Dodgson et al., 2002). In large part, these new institutional arrangements have been prompted by the "real world" of accelerating and intensifying globalization. These impacts are challenging the capacity of states, acting alone or collectively, to protect and promote the health of their domestic populations. As Aginam (2005, p. 58) writes:

> Transboundary disease spread now constitutes a global crisis that requires the pooling of efforts and resources by nation states in a multilateral context. In no other sphere of global relations is the global village metaphor more practical . . . [in view of] the permeation of national boundaries by disease pathogens and the consequent vulnerability of populations within those boundaries to microbial threats.

The forms of health governance that have emerged in response to these challenges are distinct in several ways. First, given the scale and scope of cross-border and transborder health risks faced, GHG is characterized by *collective action less focused on national borders*. International health governance is built on practices such as screening at ports of entry, quarantine measures, and national disease surveillance and reporting. The global geographies of health and disease, which create new spatial distributions and dynamics of health determinants and outcomes, require different forms of governance. Disease surveillance, monitoring, and reporting systems, which do not rely wholly on data collected by and about states, now provide valuable supplements to the International Health Regulations (IHRs), which were revised in 2005 (see the *Public Health Infrastructure* chapter for more information on the IHR). The Global Public Health Intelligence Network (GPHIN), created in 1997 by Health Canada, analyzes 20,000 online news reports worldwide in 9 languages each day, to alert WHO of potential outbreak risks (Dion, AbdelMalik, & Mawudeku, 2015). The concept of *planetary health*, based on an understanding that human health depends on the wise stewardship of natural systems, raises concerns about "present systems of governance" as "inadequate to address the threats to planetary health" (Whitmee et al., 2015, p. 1974) (see the *Environmental and Occupational Health* chapter for more on planetary health).

Second, GHG is characterized by a *growth in the number, and degree of influence, of non-state actors*. Non-state actors are organizations operating outside of the state or governmental sphere. They can be either for-profit or not-for-profit entities, depending on how they earn and use their financial resources. In general, *for-profit organizations* (e.g., private businesses, industry associations, consultancy firms) seek to maximize their financial returns, and use the earned proceeds to benefit private owners and shareholders. In contrast, *not-for-profit organizations* (e.g., patient groups, charities, religious groups, healthcare cooperatives) seek

financial returns to serve shared or public interests. The boundaries dividing these two categories are not entirely clear cut, with ongoing debates concerning the vested interests served by some non-state actors such as trade unions, think tanks, and private foundations.

What is notable, in global health and other spheres of global governance, is that the number of non-state actors has grown more rapidly than the number of governmental actors. Using data from the *Yearbook of International Organizations*, Weiss et al. (2013) describes a dramatically changing landscape over the 20th and early 21st centuries (**FIGURE 21-1**). Since 1950, more than 33,000 international governmental organizations (IGOs) and international nongovernmental organizations (INGOs) have been founded—more than one per day. INGOs grew especially rapidly during the final decades of the 20th century, to the point that they outnumbered IGOs by 9.5 to 1 by 2000.

As well as becoming more numerous, some non-state actors now bring substantial resources to global health and, by extension, can command a louder voice in GHG. The most prominent example is the Bill & Melinda Gates Foundation (BMGF), which was launched in 2000. By 2016, private foundations tracked by the Institute for Health Metrics and Evaluation provided $2.3 billion to global health development, of which 78.5% as disbursed by the BMGF (Institute for Health Metrics and Evaluation, 2017). This undeniable generosity brings vital new funding for a variety of health needs, and has been especially welcomed amid flagging aid commitments by donor governments since the late 2000s. As the relative size of funding by private foundations has grown, however, important questions about their governance have been raised. Whereas international (governmental) organizations are accountable to member states, in principle, and in turn the citizens of those member states, concerns have arisen about the public accountability of private foundations. The BMGF, for example, is headed by an Executive Leadership Team that "oversees all of the foundation's efforts" (BMGF, n.d.). As Harman (2016, p. 350) points out, this arrangement has led to concerns about the "sources of its money, how and where it spends its money, its partnerships, and the 'Bill Chill' effect of the foundation on global health institutions." The last concern includes the tendency to support, and therefore skew global health initiatives toward, disease-focused and technology-based approaches. Birn (2014), for example, goes further by arguing that its "pervasive influence . . . [is of] grave concern both to democratic global health governance and to scientific independence." In response, efforts are being made to improve the transparency and accountability of all non-state actors, especially those as influential as the BMGF (Youde, 2013).

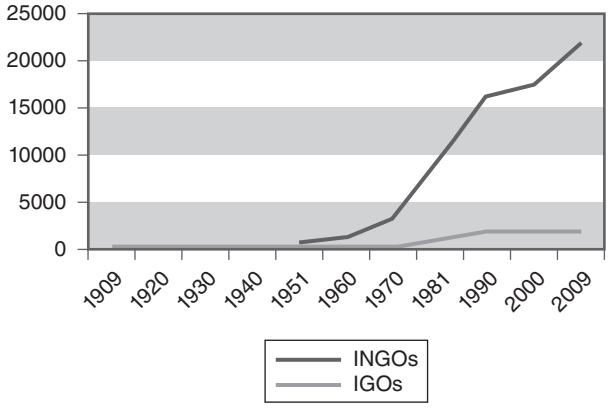


FIGURE 21-1 Number of intergovernmental and international nongovernmental organizations, 1909–2009.

Reproduced from Weiss, T., Seyle, D., & Coolidge, K. (2013). *The rise of non-state actors in global governance opportunities and limitations* [Discussion paper]. Broomfield, CO: One Earth Future Foundation. Retrieved from http://acuns.org/wp-content/uploads/2013/11/gg-weiss.pdf

The way that non-state actors are governed and their role in governing global health relate to the wider range of functions they now perform in GHG. As Haas (2004) describes, global governance comprises a range of functions, with state and non-state actors performing these functions either formally and directly, or informally and indirectly (**TABLE 21-1**). In GHG, traditionally non-state actors have played a supplementary role where government institutions have been weak or nonexistent, where there are gaps in funding and resources, or where neglected issues or constituencies require advocacy. However, a review of the functions performed by civil society organizations (CSOs) in relation to four GHG instruments—the International Code on the Marketing of Breastmilk Substitutes, the Framework Convention on Tobacco Control (FCTC), the International Health Regulations, and Codex Alimentarius—noted that they now play a critical watchdog role, ensuring that formally mandated governmental institutions fulfill their responsibilities appropriately, and keep a watchful eye on corporate actors exerting undue influence or engaging in health-harming activities (Lee, 2010). These functions have been essential to the International Code on the Marketing of Breastmilk Substitutes, where ongoing campaigning and advocacy by CSOs have been critical to achieving policy attention. CSOs play an even more significant role in the negotiation and implementation of the FCTC, fulfilling an unusually wide range of functions traditionally dominated by governments. By comparison, CSOs have been less involved in formal proceedings of the Codex Alimentarius Commission or the revision process for the IHR (2005). In regard to the latter, there is now formal recognition of the

TABLE 21-1 Matrix of Global Governance Functions

Function	Formal/Direct	Informal Indirect
Issue linkage	■ Intergovernmental negotiations ■ New information provided by epistemic communities ■ Through financial mechanisms (GEF) ■ IOs (GEO/WEO)	■ Scientists ■ Business/industry
Agenda setting	■ IOs and member states ■ Scientists	■ NGOs ■ Media ■ Scientists
Developing usable knowledge	■ Scientists	■ Scientists ■ NGOs ■ Business/industry
Monitoring	■ IOs ■ Commitees nominated by MEA secretariat ■ MEA signatory governments	■ NGOs (particularly in developing countries) ■ Scientists
Rule making	■ Negotiations by governments ■ NGOs (principled standards)	■ Business/industry (de facto standards) ■ NGOs (principled standards)
Norm development	■ Epistemic communities	■ NGOs (equity & environmental preservation) ■ Business/industry (efficiency)
Policy verification	■ Governments	■ NGOs ■ IOs
Enforcement	■ (Hard) Law ■ WTO and MEA rules	■ NGO campaigns
Capacity building (tech transfer)	■ Official technical assistance (national and local government) ■ Business/industry ■ Science community (education/ training)	■ Business/industry (joint venture)
Capacity building (organizational skills)	■ IOs ■ NGs ■ Scientific community (education/training)	■ Business/industry
Promote vertical linkage	■ IOs ■ National and local governments	■ NGO ■ Scientific community
Financing	■ Government (ODA) ■ Regional development banks ■ Multilateral bodies	■ Business/industry

importance of CSOs in effective disease surveillance, as these organizations may play a supplementary role when state institutions cannot or will not report to WHO promptly.

Moreover, new combinations of state and non-state actors, in a myriad of partnerships, alliances, coalitions, networks, and joint ventures, now populate the field of health cooperation. These institutional arrangements for GHG reflect the search for collective action that can overcome the limitations of state authority. In a few cases, they also mark a shift in the role of non-state actors from being the subject of governance (i.e., law takers) to being participants in governance (i.e., law makers) (Ryngaert, 2016). The board of the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM), for example, is "designed to incorporate leading stakeholders in an inclusive and effective way"; consequently, nongovernmental organizations (NGOs), communities affected by the three diseases, the private sector, and private foundations are represented as voting members of the board (GFATM, 2017). In addition, "a broad range of stakeholders, moving beyond the usual participants to reach those who are not normally involved in Global Fund processes" contribute "critical input, suggestions and views about the development of the strategy that guides the Global Fund's mission" through a Partnership Forum (GFATM, 2017). Similarly, the decision-making body of UNITAID (Executive Board), an international organization that raises and grants funds to prevent, diagnose, and treat the same three diseases, includes two representatives of relevant civil society networks (NGOs and communities living with HIV/AIDS, malaria, or tuberculosis [TB]) and one representative of private foundations (UNITAID, 2017).

Third, GHG is characterized by *binding and nonbinding instruments*. Binding instruments, in the form of treaties (sometimes known as conventions), are written agreements concluded by two or more states, and governed by international law. Signatories to binding treaties are legally required to comply with the measures contained within them and, in return, are themselves protected by their provisions. Nonbinding instruments come in diverse forms (e.g., declarations, principles, resolutions, voluntary codes, communiqués) and, given the primacy afforded to the sovereign authority of states, tend to be far more numerous (Taylor, 2013). Although these types of instruments cannot compel compliance by state and non-state actors, their power to shape collective action in global health can stem from their political or moral weight (Kates & Katz, 2010).

In global health cooperation, there are relatively few binding instruments, each of which has required substantial diplomatic efforts to achieve. **TABLE 21-2** provides examples of binding and nonbinding instruments related to GHG. In addition to the treaties leading to the creation of WHO and other international health organizations, WHO member states have agreed to the IHR and FCTC as binding treaties. Professor Lawrence Gostin and others have advocated for the adoption of a binding Framework Convention on Global Health to underpin GHG. As Gostin et al. (2012, p. 790) write:

> A framework convention would establish a health financing framework with clear obligations, and would create an accountability regime with robust standards, monitoring, and enforcement. It would advance health justice through engaging marginalized and underserved populations in making and evaluating policies and through comprehensive strategies and targeted interventions designed to overcome the barriers that prevent these populations from enjoying the conditions required for good health. Governments would be held to high standards of good governance, namely inclusive participation, transparency, honesty, accountability and stewardship. The framework convention would empower people to claim their right to health.

A number of treaties concluded by other international organizations have health-related provisions, such as the International UNICEF Convention on the Rights of the Child (1989) and the World Trade Organization's (WTO) Doha Declaration on the TRIPS Agreement and Public Health (1994). More common are the large number of nonbinding instruments that encourage state and non-state actors to adopt agreed standards, norms, and practices. **EXHIBIT 21-1** describes the WHO Global Code of Practice on the International Recruitment of Health Personnel (2010), which seeks to address the serious problem of health worker migration causing "brain drain" from LMICs to high-income countries. Overall, whether the global health community seeks to achieve a binding or nonbinding instrument depends on the perceived importance and priority of an issue, which in turn shapes the instrument's political feasibility. The decision by a government to become party to a binding instrument is taken after careful consideration of the country's foreign policy as a whole, and not just its health policy. There is generally a tradeoff when agreeing to be bound by international law, between

TABLE 21-2 Examples of Binding and Nonbinding Intergovernmental Agreements Related to Global Health Governance

Binding Instruments	Nonbinding Instruments
UN Convention Relating to the Status of Refugees (1967)	WHO/UNICEF Declaration of Alma Ata on Primary Health Care (1978)
Convention on the Elimination of All Forms of Discrimination Against Women (1981)	WHO Code on the International Marketing of Breastmilk Substitutes (1981)
UNICEF Convention on the Rights of the Child (1989)	UN Principles for the Protection of Persons with Mental Illness and for the Improvement of Mental Health Care (1991)
WTO Doha Declaration on the TRIPS Agreement and Public Health (1994)	ILO Code of Practice Management of Alcohol- and Drug-Related Issues in the Workplace (1999)
Cartagena Protocol on Biosafety to the UN Convention on Biological Diversity (2003)	UN Millennium Development Goals (2000)
Stockholm Convention on Persistent Organic Pollutants (2004)	UN Declaration of Commitment on HIV/AIDS (2001)
WHO Framework Convention on Tobacco Control (2005)	FAO Guidelines on the Right to Food (2004)
WHO FCTC Protocol to Eliminate Trade in Illicit Tobacco Products (2012)*	OECD Paris Declaration on Aid Effectiveness (2005)
WHO Revised International Health Regulations (2005)	WHO Global Code of Practice on the International Recruitment of Health Personnel (2010)
UN Convention on the Rights of Persons with Disabilities (2008)	WHO Global Recommendations on Physical Activity for Health (2010)
	UN Political Declaration of the High-Level Meeting of the General Assembly on the Prevention and Control of Non-communicable Diseases (2012)
	UN Sustainable Development Goals (2015)
	Shanghai Declaration on Promoting Health in the 2030 Agenda for Sustainable Development (2016)
	UN Political Declaration on Antimicrobial Resistance (2016)

Abbreviations: FAO = Food and Agriculture Organization; ILO = International Labour Organization; OECD = Organisation for Economic Co-operation and Development; UN = United Nations; WHO = World Health Organization; WTO = World Trade Organization.

* The agreement was signed on November 12, 2012, and will come into force when 41 states party to the FCTC have ratified and acceded to the protocol.

a lessened ability to act unilaterally, and the benefits accrued when other countries are bound in the same way. Negotiating binding instruments, therefore, requires governments to accept a shared interest in concluding such an agreement. Nonbinding instruments can raise attention to an issue, create agreement on new norms and practices, and thus be effective means of GHG. In some cases, nonbinding instruments can progress to the negotiation of binding agreements as perspectives change within the global health community.

In summary, GHG concerns the agreed rules for interaction and procedures for decision making to address collective health needs and deliver shared health solutions in a global context. These rules and procedures are negotiated and upheld by state and non-state actors, serving formal and informal roles, and are embodied in binding and nonbinding instruments, resulting in direct and indirect impacts on global health. Importantly, in the early 21st century, a transition from *international* to *global* health governance has begun. The design and adoption of a wholly new "architecture" of GHG is unlikely given the complex political, technical, and practical challenges posed. Instead, the world is likely to be governed by diverse and complex institutional arrangements varying across geography, constituency, and issue-area. It is from this starting point that this chapter now turns to describing the existing institutional arrangements commonly referred to as GHG.

EXHIBIT 21-1 WHO Global Code of Practice on the International Recruitment of Health Personnel

According to the *World Social Protection Report 2014/15,* countries need at least 41.1 health workers per 10,000 people to be able to provide essential health care to their own populations. For many low-income countries, however, the available personnel fall far short of this ratio. For example, in countries such as Haiti, Niger, Senegal, and Sierra Leone, there are 5 or fewer health workers available per 10,000 people, compared to 269 in Finland. In 2014, it was estimated that there was a total shortfall of 10.3 million health workers worldwide (International Labour Organization [ILO], 2015). This situation has arisen for many reasons, but one key factor is certainly health worker migration.

Global patterns of health worker migration are complex and dynamic, producing a mixture of costs and benefits at the individual and population levels. The effects are also felt differently across the diverse source and destination countries. In general, the pattern of flow is from low-income countries to higher-income countries. The problem is most acute in sub-Saharan Africa, which has 24% of the global disease burden but only 3% of the health workforce (Taylor, Hwenda, Larsen, & Daulaire, 2011). This "brain drain" ultimately weakens the health systems capacity in the source countries that bear a disproportionate share of the costs. While destination countries benefit from such migration by filling gaps in human resources and strengthening their capacity to deliver health services, source countries lose valuable human capital and, in turn, face gaps in their health services.

Health worker migration has been the subject of longstanding discussions within the global health community. The issue requires collective action across countries because of the cross-border dynamics (i.e., the "pull" factors attracting health workers to migrate stem from outside of affected countries) and the shared experience of brain drain (i.e., 57 low-income countries face critical shortages in their healthcare workforces [WHO, 2006]). Moreover, health worker migration is arguably a planetary problem. As Taylor et al. (2011) write:

> These shortages . . . contribute substantially to the weakness of health systems and obstruct the achievement of public health goals such as reductions in maternal and child mortality; they also hinder implementation of such international legal agreements as the WHO 2005 International Health Regulations, which aim to protect populations from the international spread of disease and to enhance public health security. Strong health systems are essential to improving health outcomes in underserved populations. Without effective health systems employing enough skilled, motivated, well-supported, and adequately supervised health workers, it's unlikely that vaccines and other important medical interventions can improve global health outcomes in any sustainable way. *

Recognizing that countries cannot address this problem alone, the global health community has developed a variety of strategies to improve the governance of health worker migration. Alongside strategies to increase the supply of health workers, deploy available resources more effectively (e.g., task-shifting), and change the incentive structures supporting migration, codes of practice have been adopted to regulate the active recruitment of workers from low-income settings. These governance instruments include the Kampala Declaration adopted at the First Global Forum on Human Resources for Health, the World Medical Association Statement on Ethical Guidelines for the International Recruitment of Physicians (2003), and a large number of bilateral agreements (Pagett & Padarath, 2007). For the most part, these voluntary codes have had limited scope and ultimately minimal effects in stemming the outward flow of healthcare workers (Mackey & Liang, 2013).

In May 2010, the WHO Global Code of Practice on the International Recruitment of Health Personnel was adopted as a framework for cooperation and platform for continuing dialogue on the critical problem of health worker migration. Siyam et al. (2013, p. 820) describe the process of negotiation as "vigorous" and displaying "maturity and a favourable evolution in global health diplomacy." The Global Code is based on the agreed principle, embedded in the WHO Constitution, that all people have a right to the highest attainable standard of health. At the same time, it recognized that health workers have the right to migrate from one country to another in search of employment. As a nonbinding instrument, the Global Code calls upon member states to voluntarily implement and report on measures that uphold these principles.

While a voluntary code might seem to be a limited advance from previous efforts to address international recruiting of healthcare personnel (Tankwanchi, Vermund, & Perkins 2014), some have argued that the Global Code's approach signals a shift in the debate based on recognition of health worker migration as a social problem requiring global solidarity. The adoption of the Global Code by all 193 member states also reflects universal support among all countries for collectively addressing the issue more effectively. WHO describes the code as a "groundbreaking instrument [that] marks the first time that WHO member states have used the constitutional authority of the Organization to develop a code in thirty years" (WHO, n.d.).

With the Global Code now in place, attention has shifted to promoting its implementation, which has so far been limited. The need for more robust reporting by member states on their progress in the implementation of the code has been identified as an unresolved issue (Siyam et al., 2013).

*Reproduced from Taylor, A. L., Hwenda, L., Larsen, B., & Daulaire, N. (2011). Stemming the brain drain—A WHO Global Code of Practice on International Recruitment of Health Personnel. *The New England Journal of Medicine, 265*, 2348–2351.

Mapping Existing Institutional Arrangements Governing Global Health

The complex and diverse institutional arrangements that currently constitute GHG have evolved over time into a mixture of international organizations, mechanisms, and instruments. This evolution has not been guided by an overarching plan, but rather has been characterized by a *laissez-faire* approach to institutional building, largely driven by powerful state and non-state actors. Moreover, differing opinions among these powerful actors regarding which global health goals should be given priority, and how best to achieve them, have led to the formation of numerous initiatives focused on specific diseases, populations, and geographic locales. While there are interconnections and complementarities among some of these initiatives, there are also gaps, overlaps, and even working at cross purposes. In this context, mapping these numerous institutional arrangements, and the ways in which they have evolved over time, is not a straightforward task.

One way of mapping GHG is to understand it as a process of cumulatively adding new actors over time. A familiar starting point is to locate WHO as the beginning, and at the center, as the "directing and coordinating authority on international health work" (WHO, 1946). The period from the end of World War II to the late 1990s saw WHO's program of work expand in breadth and depth, alongside the establishment of other United Nations bodies concerned with health, including the UN Children's Fund (UNICEF), UN Development Program (UNDP), and UN Population Program. In the 1980s, the World Bank began to take an active role in health as part of its development financing role. In the 1990s, a broader range of other sectors began to be involved, including trade and investment, finance, customs and excise, education, and public works (water and sanitation). Globalization was broadening the factors affecting health determinants and outcomes, and health was impacting on other policy spheres. Health became more prominent on the agendas of the Organisation for Economic Co-operation and Development (OECD), the Group of Eight (G8), and the Group of 20 (G20). New global health partnerships were formed soon after, involving state and non-state actors, alongside emerging economies such as Brazil, Russia, India, China, and South Africa (i.e., the BRICS countries). Thus, GHG is seen as evolving from WHO's foundational role, with succeeding layers adding different actors and institutional arrangements over time (**FIGURE 21-2**).

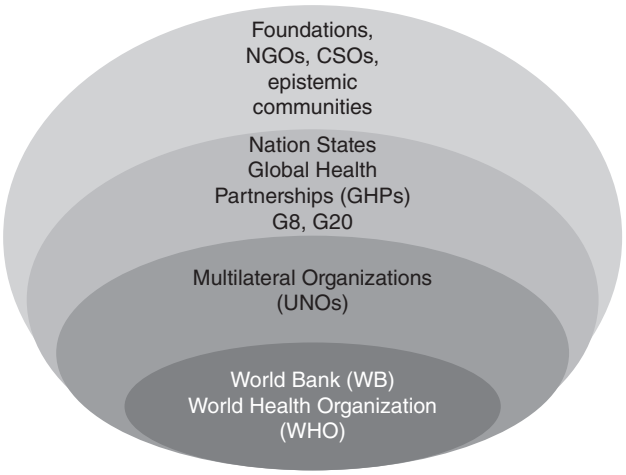


FIGURE 21-2 Prominent global health governance actors in the 21st century.

Used with permission from Okello, A., Vandersmissen, A., & Welburn, S. (2014). One health into action: Integrating global health governance with national priorities in a globalized world. In K. Zomsstag, E. Schelling, M. Whittaker, M. Tanner, & D. Waltner-Toews, (Eds.), *The theory and practice of integrated health approaches* (pp. 283–303). Wallingford, UK: CABI.

A second way of mapping GHG is to understand the changing roles and relationships played by the state, market, and civil society. State actors consist of governmental bodies, led by ministries or departments, service providers, and other agencies that serve the public interest. In GHG, this concept embraces national-level bodies concerned with health security (e.g., centers for disease control, public health agencies), health development (e.g., U.S. Agency for International Development, German Technical Cooperation Agency), and foreign policy (e.g., Global Affairs Canada), and along with international-level bodies with membership primarily composed of governmental actors (e.g., WHO, World Bank, UNICEF).

The distinctiveness of the state sphere, and therefore the central role and authority of states configured into intergovernmental organizations, is seen as changing with the advent of GHG. Since the 1990s, GHG can be mapped as an increasing overlap of the three spheres under shared institutional arrangements for the purpose of governing global health. As discussed in the previous section, non-state actors traditionally operate separately from the governmental (state) sphere, and can be divided into for-profit (market) and not-for-profit (civil society) entities. Both types of non-state actors have grown substantially in number and influence since the late 20th century. GHG can be described as bringing together the three spheres, in different configurations, to realize new resources, comparative advantages, and, ultimately, greater effectiveness (**FIGURE 21-3**). The formation of

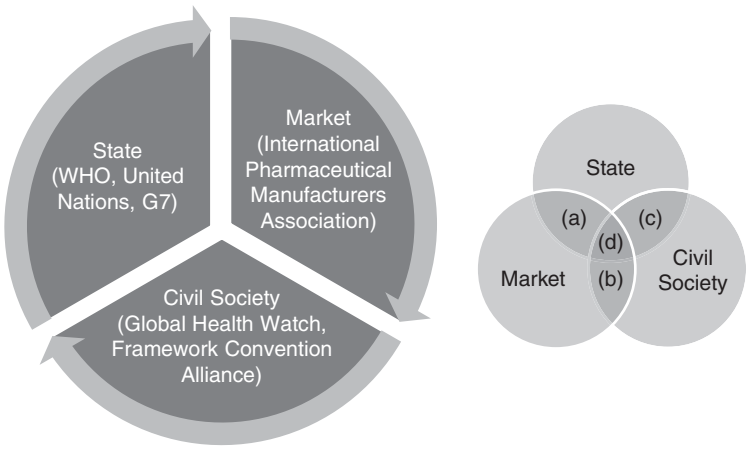


FIGURE 21-3 Transition from distinct to overlapping spheres of state, market, and civil society in global health governance.

dozens of global public–private partnerships between governments and the pharmaceutical sector, such as the Innovative Medicines Initiative between the European Commission and the European Federation of Pharmaceutical Industries and Associations, can be located in section (a) of Figure 21-3. Charitable foundations, which some argue should be located in either the market or civil society sector, might be more accurately fall between the two spheres into (b). The collaboration by Médecins sans Frontières with WHO to provide logistics for sleeping sickness (human African trypanosomiasis) drug distribution in endemic countries, is an example of section (c) of Figure 21-3. Finally, section (d) is where all three spheres come together, as exemplified by the GFATM, the Global Alliance for Vaccines and Immunization (GAVI), and UNITAID.

A third way of mapping institutional arrangements for GHG is by specific function. As described in Table 21-1, global governance fulfills a variety of functions (e.g., issue linkage, monitoring, financing) to which state and non-state actors may contribute formally or informally, with direct or indirect effects (Haas, 2004). Historically, WHO was tasked with fulfilling a broad range of these functions, including taking a lead role in agenda setting. By the 1990s, growing demands on that organization's capacities, and severe limitations of an essentially frozen budget, alongside internal challenges posed by a bureaucratic structure and weak leadership, elicited debate about the appropriate scope of WHO's mandate (Godlee, 1994). At the heart of the question of "who should be doing what in health" among UN bodies (Lee, Collinson, Walt, & Gilson, 1996) was defining the core or essential functions to be retained by WHO. Despite reforms to address some of the structural and procedural problems hindering WHO's capacity, the organization faced a growing number of "work-arounds" by dissatisfied donors who supported the creation of new initiatives. The result was a supply-driven expansion of global health initiatives, some located within WHO, but most governing their own operations.

By the early 2000s, GHG had become a varied collection of institutional arrangements that were fulfilling different functions, but not acting collectively as a whole. **FIGURE 21-4** illustrates the many state and non-state actors involved in the financing of global health (mobilization, allocation, and implementation) during this era. In this environment, despite its designation as the "directing and coordinating authority," WHO was now seen as one potential partner in a crowded GHG landscape.

While debate continues about WHO's mandate, other players have shifted their attention to the broader question of what should be the core functions of GHG more generally. For example, the concept of global public goods takes a residual view of GHG functions, considering those areas of concern to arise from "market failures" or to be undersupplied without collective action (Smith & MacKellar, 2007). Similarly, the Oslo–*The Lancet* Commission on Global Health Governance has suggested how different functions might be distributed across a restructured GHG system (Ottersen et al., 2014).

In sum, what is presently referred to as GHG is, in actuality, a less than perfect patchwork of institutional arrangements that have evolved, in *laissez-faire* fashion, from a combination of supply-driven donor dissatisfaction and demand-driven health needs arising from globalization. Truly *global* governance in the health sphere is so far limited in scope, authority, and function. It would thus be misleading to refer to GHG

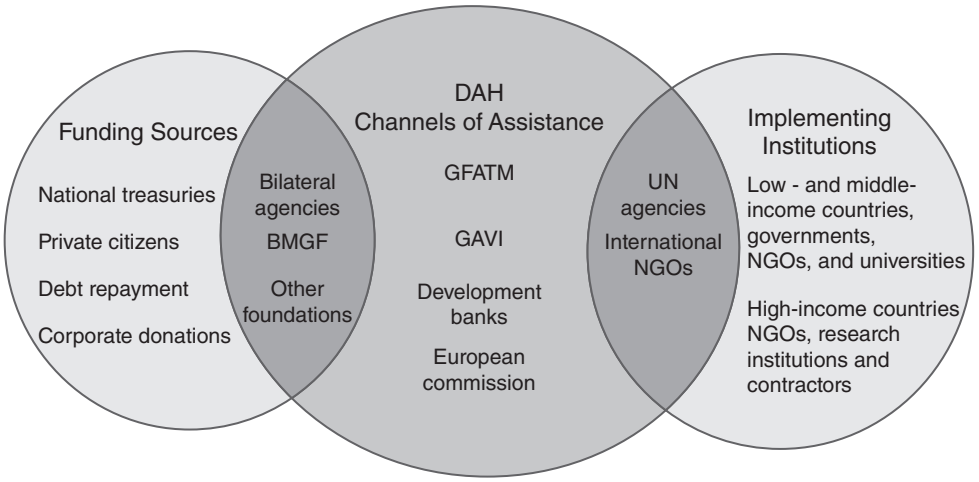


FIGURE 21-4 Overlapping roles of state and non-state actors in financing of global health governance.

Reproduced from Institute for Health Metrics and Evaluation. (2009). *Financing global health 2009: Tracking development assistance for health*. Seattle, WA: IHME.

as a "system" or "architecture" in its present form, as this would suggest an agreed strategy or vision to guide its construction or evolution. Rather, what we have is a partial transitioning from international to global health governance, for selected areas of collective action such as AIDS, TB, and malaria, and to prevent and control major communicable disease outbreaks (**EXHIBIT 21-2**). For most other issue areas, national and international health governance remains the basis for collective action.

The Role of Global Health Diplomacy

The building of effective GHG relies on the conduct of global health diplomacy. *Diplomacy* is the art or practice of conducting international relations through negotiating alliances, treaties, and other agreements. It is concerned with dialogue "designed to identify common interests and areas of conflict between the parties" (Evans & Newham, 1992). This discussion, in turn, is guided by a country's *foreign policy*, the activity whereby state actors act, react, and interact at the boundary between the internal (domestic) and external (foreign) environments. Thus, while foreign policy is "the substance, aims and attitudes of a state's relations with others," diplomacy is "one of the instruments employed to put these into effect" (Reynolds, 1980, p. 36). The term "new diplomacy" describes shifts in foreign policy that challenge how diplomatic practice is carried out. More recently, it has been used to describe the global context, diverse actors, and innovative processes shaping contemporary international relations. Another key feature of new diplomacy has been a shift in the participants, from highly trained officials within foreign affairs ministries to a broader range of political actors (Heine, 2006).

Importantly, new diplomacy has been shaped by non-state actors. While diplomats formally interact with their counterparts, their constituencies are increasingly broad based. Bayne and Woolcock (2007) argue that globalization, and the growing influence of non-state actors, is transforming foreign policy in the 21st century. As Davenport (2002, p. 19) writes, "thousands of nongovernmental organizations have come on stage in recent years, driving their own issues to the top of the diplomatic agenda." Solana (2009) argues, therefore, that "diplomacy is about more than mobilizing states. We need to find ways to harness the expertise and resources of nongovernmental organizations and companies and energize individuals towards shared goals." This diversification of a diplomat's constituencies is illustrated by the "explosion of multilateral negotiations aimed at addressing the new global environmental issues" (Benedick, 1998, p. 4). While the diplomat's formal role is to serve as an official representative of a state, new diplomacy requires engagement with wider constituencies. Thus, it requires what Riordan (2002) calls a "radical restructuring of diplomatic services, replacing hierarchical with networked structures, and the creation of new forms of interaction." New diplomacy, for example, seeks to harness new technologies, altering negotiations through enhanced consultation, coordination, information gathering, and analytical capabilities.

More broadly, social/cultural activities are seen by public diplomacy advocates as important means of promoting understanding and influence. Governments are not alone in the use of such technologies. Non-state actors see new technologies, including social networking, as means for exerting "soft power," especially when they lack military and economic resources. Consequently, there has been growing

EXHIBIT 21-2 The Global Governance of Pandemic Influenza

Influenza is a respiratory infection characterized by the rapid onset of symptoms such as fever, chills, body ache, and unproductive cough. Influenza can cause severe health problems for the elderly and the young, as well as for people with weakened immune systems. Of the three types of influenza viruses (A, B, and C), type A is of greatest public health concern. Type A viruses infect both human and various animal species, which expands the pathogens' reservoir. These viruses are unstable because they regularly mutate in minor ways (antigenic drift), but periodically mutate in major ways (antigenic shift). Genetic reassortments occur when a human influenza virus and zoonotic influenza virus coinfect a human or animal (e.g., pig, bird) host, exchange RNA, and produce a novel virus. Depending on its characteristics, the novel virus could trigger human or animal epidemics or wider pandemics.

Governments have long recognized that influenza requires collective action given the enduring risks posed by the disease and the substantial human and social impacts caused by outbreaks. For these reasons, WHO member states have agreed that global influenza governance must fulfill four core functions: surveillance and reporting, preventive measures, effective response, and public communication.

First, *surveillance and reporting* provides vital knowledge of which influenza strains are circulating, thereby enabling the planning and implementation of interventions such as vaccines. This function is carried out by the WHO Global Influenza Surveillance and Response System (GISRS), established in 1951 and operating in 83 countries. The GISRS consists of 112 national influenza centers, 4 Collaborating Centers for Reference and Research (Australia, Japan, United Kingdom, and United States), and the Collaborating Center for Studies on the Ecology of Influenza in Animals (United States). In addition, the GISRS operates FluNet, a global web-based tool for influenza virologic surveillance that was launched in 1997. The virologic data entered into FluNet (e.g., number of influenza viruses detected by subtype) are critical for tracking the movement of viruses globally and interpreting the epidemiologic data. FluNet recommends, twice annually, the content of the influenza vaccine for the coming season, and serves as an early alert mechanism for the emergence of a virus with pandemic potential. FluNet also forms part of the WHO Global Outbreak Alert and Response Network (GOARN), established in 2000, to combat the worldwide spread of disease.

Second, an important function of effective global influenza governance is to *undertake preventive measures that reduce the disease burden on populations*. Given that influenza viruses, including H5N1, are endemic in a growing number of countries, and contact between human and animal populations remains difficult to control, most experts agree that no form of governance can prevent the emergence of new influenza strains. However, public health agencies, in cooperation with animal health experts and the agricultural sector, can control the spread of the virus and mitigate its impact through rapid diagnosis, isolation of cases, and vaccination (Gerdil, 2003). Among healthy adults, influenza vaccine provides protection from infection, even when the circulating viruses do not exactly match the vaccine viruses. Unfortunately, supply limitations and the cost of vaccines put them out of reach of most populations worldwide. Other measures, such as good husbandry practices, can be supported through guidelines published by WHO and World Animal Health Organization.

Third, influenza can still cause significant disease burdens even when the preceding measures are implemented, so it requires *effective public health responses*. Seasonal epidemics are estimated to result in 3 to 5 million cases of severe illness each year, and 250,000 to 500,000 deaths annually. The morbidity and mortality for pandemic influenza vary depending on the virulence of the virus involved; for example, mortality from this cause was approximately 575,400 deaths in 2009 but 50 million to 100 million deaths in 1918–1919. The infrequency of influenza pandemics, combined with the potentially major severe public health impacts of such events, pose challenges for preparedness planning. The role of GHG, in this respect, focuses on enhancing the surge capacity of national-level health systems to cope with increased demands for diagnostic and clinical treatment, notably control of infection and administration of antiviral drugs. This effort is guided by the Pandemic Influenza Preparedness (PIP) Framework adopted by WHO member states in 2011 following growing concerns about the risks from highly pathogenic avian influenza (HPAI). More broadly, influenza governance seeks to maintain the functioning of societies, including public services and economic activity.

Finally, *public communication* is a key function of global influenza governance. Accurate and timely information is essential for ensuring an appropriate perception of risk among the public. This task is especially challenging during influenza pandemic because when and where it begins, which specific strain emerges, and how severe the health impact is can be unknown. An unduly high level of risk perception can prompt irrational behavior, which then adversely affects collective interests. The private stockpiling of antiviral drugs, for example, contributes to shortages in world supply. Maintaining a high level of alert can lead to "risk fatigue" and public skepticism when a pandemic does not materialize. At the same time, failure to raise sufficient awareness and concern may create problems in securing needed resources or public compliance with response measures.

(continues)

EXHIBIT 21-2 The Global Governance of Pandemic Influenza *(continued)*

Today, surveillance and reporting is the most developed function under global influenza governance, although data gaps remain for many LMICs. Web-based monitoring and reporting systems, integrating both state and non-state sources, and new technologies (e.g., web trawling) have further enhanced the delivery of this function. Global influenza governance remains most vulnerable in relation to the three core functions of preventive measures, response, and public communication. Despite recognition of the need to implement the PIP Framework, much of the effort to prepare for pandemic influenza remains focused at the national level. Pharmacologic interventions are, at present, available in only some countries. In addition, the stockpiling of antiviral medications and vaccines by individual countries causes shortages and increased prices worldwide. This practice is especially problematic if such drugs are needed, for example, to contain an outbreak during the initial stages. By undermining GHG capacity to ring-fence an outbreak, governments can actually reduce the health security of their domestic populations. As Coker and Mounier-Jack (2006, p. 889) argue, "The next pandemic will test notions of global solidarity."

Excerpted from Lee, K., & Fidler, D. (2007). Avian and pandemic influenza: Progress and problems for global governance. *Global Public Health 2*(3): 215.

recognition of the importance of *netpolitik* and *media-politik* (in contrast to *realpolitik*), a "new style of diplomacy that seeks to exploit the powerful capabilities of the Internet to shape politics, culture, values, and personal identity" (Bollier, 2003, p. 2).

Global change, in turn, is creating new problems ripe for collective action (e.g., climate change, population migration, economic instability, disease pandemics) through foreign policy and, consequently, diplomats. The new diplomacy agenda, then, includes a broader range of issue-areas deemed relevant to foreign policy (McInnes & Lee, 2006). This has given rise to new specialty areas of diplomacy such as *resource diplomacy* (the cultivation of relations with resource-rich countries), *Internet diplomacy* (negotiation of rules governing the Internet), and *disaster diplomacy* (understanding and addressing risks in a complex global system).

It is within this context that growing attention has been paid to global health diplomacy. Much of the GHD literature is overtly normative in advocating the use of global health activities to further foreign policy. For example, the U.K. government's *Health Is Global Strategy* seeks to "use health as an agent for good in foreign policy, recognizing that improving the health of the world's population can make a strong contribution towards promoting a low-carbon, high-growth global economy" (United Kingdom Department of Health, 2008). Similarly, Suleman et al. (2014, p. 1) propose that "[t]houghtful Health Diplomacy . . . has the potential to bridge the perceived divides between Western and predominantly Muslim nations." In the United States, GHD is often cast as a counterpart to military-backed "hard power" in the search for "what other tools are available to advance U.S. interests in the world" ("Health Diplomacy," 2007, p. B4).

In summary, GHD can be defined as "policy-shaping processes through which States, intergovernmental organizations, and non-state actors negotiate responses to health challenges or utilize health concepts or mechanisms in policy-shaping and negotiation strategies to achieve other political, economic, or social objectives" (Smith, Lee, & Fidler, 2009, p. 1). Such processes recognize that the changing roles and responsibilities of the increasingly diverse public and private actors concerned with global health require a range of approaches to collective action. Of significance is the absence of an overarching authority to adopt and enforce legally binding measures. Yet the world's experiences with severe acute respiratory syndrome (SARS) and H1N1 influenza, for example, have reemphasized the shared nature of global health challenges. To date, the focus has largely centered on negotiating global governance mechanisms to deal with acute public health threats—notably, infectious diseases outbreaks. However, greater attention needs to be given to how state and non-state actors at different policy levels can work more effectively together to address other global health issues, such as strengthening of health systems, health worker migration, access to medicines, and prevention and control of the rapid increase in chronic diseases.

Over the past six decades, WHO has been at the forefront of global health diplomacy as the United Nations' specialized agency for health. The World Health Assembly (WHA) has served as a valued forum for debating issues and encouraging consensus among member states. The post–Cold War era, however, has seen the rise of the G8 (reduced to G7 with suspension of Russia since 2014) countries as a core influence in international relations. Since the 2000s, the rising influence of the BRICS countries, along

with other "emerging economies," has been a notable development on the world stage. The most remarkable change among these countries is the transformation of China from an aid recipient to a leading aid donor over the past three decades. While there are limited official data sources on Chinese global health funding, a study by Grepin et al. (2014) reports that health has increased as a development aid priority for China. This change in perspective is reflected in the steady increase in the number of health, population, water, and sanitation projects supported in Africa since 2000, with China ranking among the top 10 bilateral global health donors to that continent. Beyond bilateral aid, China has been an active participant in the annual BRICS Health Ministers meetings that have held since 2011. These meetings include six roundtables and two Ministerial Forums on China–Africa Health Development, involving 40 African countries as well as representatives from the African Union, WHO, and UNAIDS (Röhren, 2017). In January 2017, China and WHO signed a memorandum of understanding on health issues related to the so-called Belt and Road Initiative, a monumental development strategy focused on strengthening land and sea trade routes between Europe and China. The agreement deals with health emergencies and with the certification of Chinese medicines and vaccines. This rapid growth of China's role in global health, in large part driven by the country's foreign and economic policy goals, has led to observations that China will soon overtake the United States as "the next leader in global health" (Dahl, 2017).

During the 21st century, it is likely that further geopolitical shifts will lead emerging economies, individually and collectively, to have greater prominence in GHD. The New Development Bank, formed by BRICS countries in 2014, intends to pursue the objective of "financing infrastructure and sustainable development projects in BRICS and other emerging economies and developing countries, complementing the efforts of multilateral and regional financial institutions toward global growth and development" (New Development Bank, n.d.). The G20 (the world's 20 leading industrial countries and emerging economies) is also poised to take a more active global health role. Its Berlin Declaration, issued in May 2017, recognizes health as "one of the most valuable resources for building a sustainable future" (Group of 20, 2017). The declaration, which follows the first G20 Health Ministers Meeting, focuses on antimicrobial resistance and health systems strengthening.

Along with state-based diplomacy, as discussed earlier in this chapter, global health actors embrace a broad range of non-state actors. For example, watching the G20 closely has been a coalition of civil society organizations known as the C20, which aims to facilitate "a structured and sustained exchange of critical reflection and political perspectives amongst civil society in G20 countries and beyond on the G20 agenda" (C20, n.d.). The diversity of actors concerned with global health is extended further by the influence of other sectors, including trade, security, environment, migration, and agriculture, upon the health sector. This is illustrated by the negotiation process to move from the MDGs to SDGs (**EXHIBIT 21.3**). Together, the greater diversity of actors and the complexity of issues gives the role of GHD even greater importance (Kickbusch & Kökény, 2013). Far greater understanding is needed concerning how GHD can best be conducted, which skills are needed for effective negotiation, and which ends GHD should seek to achieve. In response to calls for more information on these topics, there has been a blossoming of GHD literature and training programs focused on the skills and processes needed for the negotiation of global health-related agreements (Kickbusch & Kökény, 2013). As former WHO Director Margaret Chan (2007) noted, health negotiators need to assert themselves more effectively in new policy arenas:

> [D]ue to the impact of globalization on public health, a new type of diplomacy is necessary to navigate the changing landscape of international affairs and politics. The emergence of cross-border disease, bio-terrorism, shifting geopolitical environments, and the linkages between health, trade, intellectual property, and human rights, present stakeholders with a complex matrix of technical and relational challenges.

Key Tasks for Strengthening Global Health Governance

Debates over the past two decades about the present weaknesses of GHG have often focused on dissatisfaction with specific international organizations' willingness and ability to fulfill their mandates, notably WHO. A substantial amount of intellectual and practical energy has been expended searching for effective options for reforming, rebuilding, or even replacing WHO. From concerns during the mid-1990s about weak leadership, lack of nimbleness, and poor leadership, to the crisis of confidence following the Ebola virus outbreak of 2013–2016, there are clear frustrations with the process of "fixing" WHO. Even the reform recommendations prompted by the Ebola

EXHIBIT 21-3 Negotiating the Role of Health Within the UN Sustainable Development Goals

The 2030 Agenda for Sustainable Development was adopted by the UN General Assembly in September 2015 to replace the Millennium Development Goals (MDGs) as the central UN platform for achieving "integrated and indivisible" development across three dimensions: social, environmental, and economic. Encompassing 17 Sustainable Development Goals (SDGs) and 169 targets, which require national, regional, and local efforts across all sectors of society, the purpose of the SDG agenda is to serve as a broad and inclusive framework for ending poverty worldwide.

The SDGs differ from the MDGs in two important respects. First, the UN Secretary General's High-Level Panel on Sustainable Development highlights the importance of addressing the determinants of development and well-being within a "universal framework" in which goals apply to every country and across all sectors of society. While the MDGs applied to developing countries, the SDGs are framed to address poverty alleviation in all countries. Second, the SDGs place sustainability at the heart of the development agenda, recognizing the need to address the complex links between development and the environment. Of particular concern is the need to move away from a narrow set of quantitative goals and targets and toward a broader range of more "integrated and indivisible" goals and targets.

Before consensus could be achieved on these differences, however, negotiators had to navigate a broad range of sector-specific interests. This included the health sector, which enjoyed particular prominence under the MDGs. Three out of eight MDGs focused directly on health-related issues: reducing child mortality (MDG 4), improving maternal health (MDG 5), and combating HIV/AIDS, malaria, and other diseases (MDG 6). Three others were indirectly related to health: eradicating extreme poverty and hunger (MDG 1), ensuring environmental sustainability (MDG 7), and developing a global partnership for development (MDG 8). This emphasis arose because, at the time the MDGs were negotiated during the late 1990s, major donor agencies heavily favored support for addressing global health needs. During this period, increased funding was directed to global health, with the sector receiving substantial new resources and political support (Nunes, Lee, & O'Riordan, 2016).

A detailed comparison of the negotiation of the MDGs and SDGs is difficult given the "absence of contemporaneous, empirical documentation of the agenda-setting and decision-making process resulting in the MDGs" (Brolan & Hill, 2016, p. 514). Nevertheless, it appears that the processes were different. The MDGs arose from the Millennium Summit, described as the "largest gathering of world leaders in history" (UN Millennium Project, 2006). According to McArthur (2014, p. 7), the text of the UN Millennium Declaration was the product of an "intergovernmental drafting process . . . led by John Ruggie, Kofi Annan's advisor and assistant secretary-general. Mark Malloch Brown, the UNDP administrator. . . . In the end, world leaders signed off on a historic Millennium Declaration text that distilled a wide-ranging global agenda down to a relatively pithy framework of global priorities. . . . The text was ultimately drafted in a UN conference room, but its contents had roots in the outcomes from the 1990s conferences."

While the SDGs were supported by almost 200 UN member states, they were the product of a process facilitated by a Sustainable Development Solutions Network (SDSN), overseen by a Leadership Council, operating "under the auspices of the UN Secretary-General . . . [to mobilize] global scientific and technological expertise to promote practical solutions for sustainable development." The SDSN formed 12 Thematic Groups (TGs), including the Thematic Group on Health for All, whose members comprised leading scientists, engineers, academics, and practitioners from civil society and the business community (SDSN, 2014). As Wagner (2015) describes:

> Several elements combined to deliver a different negotiation process on the SDGs. A critical change was the fact that North–South differences were not as pronounced. The SDG negotiation process was conducted in a manner that reduced delegation rigidity, both of individual member states and within coalitions. Based on the Rio+20 instructions for how the SDGs would be developed, negotiations were conducted by an "Open Working Group" [OWG] in which 70 countries developed a sharing arrangement for the designated 30 "seats" for participants. The sharing arrangement broke up traditional coalitions, and facilitated discussions in which seat partners sought to identify what they shared in common with each other's position, rather than to strategize over how to elevate their different positions.

This new negotiating format led to a shift away from the prominence previously given to the health sector. During the negotiations, some argued that the MDGs had been too focused on quantifiable targets, which were closely associated with health targets such as reducing child mortality rates; those targets then became goals in themselves, rather than achieving overall development. As Hill et al. (2014, p. 1) describe, the "global context and framing of the new agenda is substantially different, and health advocates cannot automatically assume the same prominence." Instead, health needs to be reframed within a social sustainability paradigm, be universally relevant to all countries, and be asserted across all themes, rather than being identified as stand-alone goals and targets. On this basis, they suggest that "health advocates need to overtly explore what global governance structures will be needed to finance and implement these universal Sustainable Development Goals" (p. 1).

The end result of the negotiation process was that, among the 17 SDGs, only one is directly a health goal (SDG 3) including universal health coverage. For some, this outcome signals a downgrading of health on the global agenda. For others, the broader approach represents an improvement over the MDG targets, but questions remain about its achievability without a clear commitment to the right to health, collective responsibility (beyond individual societies) to uphold this commitment, and accountability mechanisms to ensure the delivery of the right to health as an entitlement (Go4Health, 2013).

outbreak (**EXHIBIT 21-4**)—an event in which WHO's conduct faced unprecedented criticism—have once again led to disappointment that real change remains elusive (Horton, 2015).

WHO is not alone in drawing reformers' attention: Other international organizations have also faced criticism. For example, the impacts on health equity and status in LMICs of the World Bank's market-driven policies and conditionalities have raised major concerns since the early 1990s (Birn & Dmitrienko, 2005). In 2011, the GFATM faced criticism regarding its governance and financial oversight, confirmed by findings of a high-level independent review panel. The review, chaired by former U.S. Health Secretary Michael Leavitt and former President of Botswana Festus Mogae, acknowledged the important work and achievements of the Global Fund, but noted that inadequate fiduciary controls resulted in misappropriation of funds in recipient countries (Boseley, 2011).

Beyond the governance of individual organizations, many challenges stem from the way that existing institutional arrangements together serve to govern global health. As discussed in this chapter, the need to strengthen collective action has been driven by impacts wrought by globalization (e.g., technological change, increased capital flows, intensifying population mobility) that existing forms of governance cannot address in an effective manner. Can existing bodies, individually and collectively, govern in ways that meet the health needs of a globalizing world, or is there need to design a new "system"? The latter task, of achieving a better "system" of GHG, must resolve a number of challenges.

Which Normative Frameworks Should GHG Seek to Uphold?

Any redesign of GHG requires agreement among those governed about the shared values and ethics underlying collective action and institution building. The normative frameworks shaping global affairs are, of course, diverse and sometimes at odds. Nevertheless, a substantial part of the global health community has moved toward the adoption of some values as the starting point for GHG, beginning with the basic right to health. A commitment to addressing health inequities, within and across countries, and to building universal health coverage, also forms important parts of the SDGs. What does fairness and justice look like in GHG?

Which Core Functions Should GHG Fulfill?

Among the many functions that can be performed by global governance (Table 21-1), there is a need for agreement on which functions should form part of GHG, and how they should be carried out. Today, health governance is exercised at many different institutional levels: local/subnational (e.g., district health authority), national (e.g., ministry of health), regional (e.g., Pan American Health Organization), international (e.g., WHO), and global (GFATM). Should GHG be considered residually, as comprising only those functions that cannot be performed by other levels of health governance? Or should GHG be considered holistically, as comprising all levels of health governance scaffolded in an interconnected and complementary way? For both visions of GHG, it is clear that building stronger global institutions is not a zero-sum game, whereby new institutional arrangements necessarily seek to replace existing ones. Rather, by focusing on the functions to be performed, effective GHG could be viewed as an enabling structure built upon strong institutions from the ground up.

What Should Be the Relative Roles of State and Non-state (For-Profit and Not-for-Profit) Actors?

This chapter has described how GHG increasingly involves state (e.g., ministry of health) and non-state (e.g., International Federation of Pharmaceutical Manufacturers Association) actors, sometimes operating separately, but increasing working in combination (e.g., Project Last Mile, Malaria for Medicines Venture). Pragmatism has been a major factor in the proliferation of global public–private partnerships. However, their roles have also raised questions about whose interests are served by GHG. Can GHG serve

EXHIBIT 21-4 Ten Recommendations for Strengthening Global Health Governance After the 2013–2016 Ebola Virus Outbreak

Recommendation 1: Develop a global strategy to invest in, monitor, and sustain national core capacities.
WHO should convene governments and other major stakeholders within 6 months to begin developing a clear global strategy to ensure that governments invest domestically in building core capacities and to mobilize adequate external support to supplement efforts in poorer countries.

Recommendation 2: Strengthen incentives for early reporting of outbreaks and science-based justifications for trade and travel restrictions.
Political leaders, governments, and international organizations must strengthen the set of incentives and disincentives so that governments report disease outbreaks early. Among these should be stronger disincentives for implementing trade and travel restrictions without a scientific or public health basis.

Recommendation 3: Create a unified WHO Center for Emergency Preparedness and Response with clear responsibility, adequate capacity, and strong lines of accountability.
High-level political leaders must clearly designate who is responsible for responding when disease outbreaks outstrip national capacities, invest in the capacity to respond, and ensure accountability for fulfillment of these responsibilities.

Recommendation 4: Broaden responsibility for emergency declarations to a transparent, politically protected Standing Emergency Committee.
Member states should amend the International Health Regulations to broaden responsibility for declaring a public health emergency of international concern. The Director-General convenes, and is advised by, an ad hoc Emergency Committee constituted from a list of independent experts; however, authority and responsibility to declare a public health emergency of international concern rest exclusively with the Director-General. We recommend the creation of a Standing Emergency Committee that meets regularly, with the mandate to declare a public health emergency of international concern by a majority vote of its members. The emergency declaration should trigger other actions, such as financial disbursements by development banks, emergency data-sharing and specimen-sharing rules, and emergency regulatory procedures for new drugs, vaccines, and diagnostics (recommendations 6 and 7). The Director-General should chair, communicate, and explain the Standing Emergency Committee's decisions. Following an open call for nominations, the Director-General would appoint the first members; thereafter, the Standing Emergency Committee itself would periodically vote in new members to preserve its independent character. Minutes and votes of Standing Emergency Committee members should be published immediately following each meeting for the sake of transparency, to build external confidence, reduce political interference, and strengthen the committee's hand against resistant states. Similarly to other institutions responsible for technically complex, yet politically consequential decisions, such as central banks or drug regulatory authorities, the Standing Emergency Committee must be protected from political pressure that might interfere with its judgment.

Recommendation 5: Institutionalize accountability by creating an independent Accountability Commission for Disease Outbreak Prevention and Response (Accountability Commission).
The UN Secretary General should create an Accountability Commission as an independent body composed of civil society, academia, and independent experts doing real-time and retrospective system-wide assessment of global responses to major disease outbreaks. The Accountability Commission would track and analyze the contributions and results achieved by national governments, donors, UN agencies, international and national nongovernmental organizations, and the private sector. All major actors would be expected to share information promptly with the Accountability Commission about financial, in-kind, or operational contributions; the Accountability Commission should publish the names of organizations unwilling to share such information. The Accountability Commission would assess aid effectiveness, including funds committed, paid, disbursed, and spent; both short-term and long-term accomplishments achieved with those funds; and the timeliness, effectiveness, cultural appropriateness, and equity of the response for intended beneficiaries. The Accountability Commission should liaise directly with and provide a forum for representatives of communities directly affected by outbreaks. Finally, it should monitor efforts to build and sustain national core capacities.

Recommendation 6: Develop a framework of rules to enable, govern, and ensure access to the benefits of research.
Before the 2016 World Health Assembly, WHO should convene governments, the scientific research community, industry and nongovernmental organizations to begin developing a framework of norms and rules for research relevant to disease outbreaks. The framework's goal would be to provide guidance on three interrelated issues: (a) access to data and samples to enable and accelerate research; (b) improved ethical standards; and (c) equitable access to the benefits of research such as priority and affordable access to newly developed technologies and treatments.

Recommendation 7: Establish a global facility to finance, accelerate, and prioritize research and development.
The UN Secretary General and the WHO Director-General should convene in 2016 a high-level summit of public, private, and not-for-profit research funders to establish a global financing facility for research and development for health technology relevant for major disease outbreaks. The facility would support manufacturing, research, and development for drugs, vaccines, diagnostics, and other nonpharmaceutical supplies (such as personal protective equipment) where the commercial market does not offer appropriate incentives. For known pathogens, the facility could invest in bringing candidate drugs, vaccines, technology platforms, and other relevant products through proof of concept, phase 1, and phase 2 testing in humans, so that they are ready for wider testing, manufacturing, and distribution when an outbreak strikes. During an outbreak the facility would rapidly mobilize finance for priority research and development projects, such as diagnostics for novel pathogens.

Recommendation 8: Sustain high-level political attention through a Global Health Committee of the Security Council.
In recognition of health as an essential facet of human and national security, the UN Security Council should establish a Global Health Committee consisting of government representatives. The Committee's main goal would be to expedite and elevate political attention to health issues posing a serious risk to international peace and security and provide a prominent arena to mobilize political leadership. Specifically, the Committee would monitor and publish an annual report on progress in building a strong and effective global health security system, taking into account analyses from the Accountability Commission and WHO. The Committee would also address alleged noncompliance with International Health Regulation provisions on trade and travel measures. The Committee would not declare public health emergencies of international concern. That decision would remain technically driven and under the authority of WHO. The Committee would not be able to veto WHO decisions or reports of the Accountability Commission. Rather, the Committee's main role would be as an arena for high-level attention to health threats and a forum for problems not adequately resolved by the WHO governing bodies.

Recommendation 9: A new deal for a more focused, appropriately financed WHO.
To rebuild trust, respect, and confidence within the international community, WHO should maintain its broad definition of health, but substantially scale back its expansive range of activities to focus on core functions. The scope of WHO's work would thus continue to embrace the full range of health issues, but its functions should be far more circumscribed. We restrict our analysis to core functions in infectious disease outbreaks. However, there remains the need to define WHO's core functions in other key areas of work, such as noncommunicable diseases, injuries, environmental health, health systems, and social determinants of health. For this purpose, the January 2016 Executive Board should launch a fundamental review of the organization's constitution and mandate to define its core functions. This review should identify and hand over noncore activities to other actors, thereby streamlining WHO's activities. It should also examine which core functions are not being fulfilled or adequately funded.

Recommendation 10: Good governance of WHO through decisive, time-bound reform, and assertive leadership.
Restoring credibility demands that WHO institutionalizes accountability mechanisms, strengthens and clarifies how it works with other actors, and fosters strong leadership. The January 2016 Executive Board should launch a process to implement four new policies for WHO to meet basic principles of good governance: establish a freedom of information policy, with appropriate safeguards; create a permanent Inspector General's office to monitor overall performance of the organization and its entities, reporting to the Executive Board; conclude continuing work on the Framework of Engagement with Non-State Actors to better govern the way WHO interacts with civil society, academia, foundations, and the private sector; and revise human resources policies to attract or retain well-qualified staff, including for leadership positions, while letting go of chronic underperformers.

Extracted from Moon, S., Sridhar, D., Pate, M., Jha, A., Clinton, C., Delauney, S., . . . Piot, P. (2015). Will Ebola change the game? Ten essential reforms before the next pandemic. The report of the Harvard-LSHTM Independent Panel on the Global Response to Ebola. *The Lancet, 386*: 2204–2221.

public interests with the participation of both state and non-state actors? Which institutional arrangements are needed to ensure appropriate "checks and balances" to protect public interests in GHG? The adoption of the WHO Framework for Engagement with Non-State Actors (FENSA) in May 2016 followed a long and "difficult gestation" (Buse & Hawkes, 2016) amid concerns regarding the organization's relationship with the private for-profit sector (**EXHIBIT 21-5**). While FENSA "recognizes the risks of potential conflicts of interest from engagement with non-state actors, including undue influence in setting or applying policies, norms, and standards," and "WHO and its governing body have taken an important step in democratizing the invite list to the policy table and establishing the dining etiquette," (p. 5) Buse and Hawkes (2016, p. 5) argue that this was a lost opportunity for the organization to "leverage its mandate and authority to address the larger issue of governing the activities of industry."

EXHIBIT 21-5 WHO Framework of Engagement with Non-State Actors

Introduction

1. The overarching framework of engagement with non-state actors and the WHO policy and operational procedures on management of engagement with non-state actors apply to all engagements with non-state actors at all levels of the Organization, whereas the four specific policies and operational procedures on engagement are limited in application to, respectively, nongovernmental organizations, private sector entities, philanthropic foundations and academic institutions.

Engagement: Rationale, Principles, Benefits, and Risks

Rationale

2. WHO is the directing and coordinating authority in global health in line with its constitutional mandate. The global health landscape has become more complex in many respects; among other things, there has been an increase in the number of players including non-state actors. WHO engages with non-state actors in view of their significant role in global health for the advancement and promotion of public health and to encourage non-state actors to use their own activities to protect and promote public health.
3. The functions of WHO, as set out in Article 2 of its Constitution, include: to act as the directing and coordinating authority on international health work; to establish and maintain effective collaboration with diverse organizations; and to promote cooperation among scientific and professional groups which contribute to the advancement of health. The Constitution further mandates the Health Assembly or the Executive Board, and the Director-General, to enter into specific engagements with other organizations. WHO shall, in relation to non-state actors, act in conformity with its Constitution and resolutions and decisions of the Health Assembly, and bearing in mind those of the United Nations General Assembly or the Economic and Social Council of the United Nations, if applicable.
4. WHO's engagement with non-state actors supports implementation of the Organization's policies and recommendations as decided by the governing bodies, as well as the application of WHO's technical norms and standards. Such an effective engagement with non-state actors at global, regional, and country levels, also calls for due diligence and transparency measures applicable to non-state actors under this framework. In order to be able to strengthen its engagement with non-state actors for the benefit and interest of global public health, WHO needs simultaneously to strengthen its management of the associated potential risks. This requires a robust framework that enables engagement and serves also as an instrument to identify the risks, balancing them against the expected benefits, while protecting and preserving WHO's integrity, reputation, and public health mandate.

Principles

5. WHO's engagement with non-state actors is guided by the following overarching principles.
   Any engagement must:
   a. demonstrate a clear benefit to public health;
   b. conform with WHO's Constitution, mandate, and General Programme of Work;
   c. respect the intergovernmental nature of WHO and the decision-making authority of member states as set out in the WHO's Constitution;
   d. support and enhance, without compromising, the scientific and evidence-based approach that underpins WHO's work;
   e. protect WHO from any undue influence, in particular on the processes in setting and applying policies, norms, and standards;
   f. not compromise WHO's integrity, independence, credibility, and reputation;
   g. be effectively managed, including by, where possible, avoiding conflict of interest and other forms of risks to WHO;
   h. be conducted on the basis of transparency, openness, inclusiveness, accountability, integrity, and mutual respect.

How Should GHG Intersect with, or Be Integrated with, Other Spheres of Governance?

Health determinants and outcomes, in a globalizing world, are affected by a broad range of factors. As a consequence, effective GHG cannot be limited to the health sector, but rather must operate within a larger ecosystem of global governance. How should GHG engage with non-health sectors (e.g., trade, environment, law enforcement, education, transportation)? How can this be achieved at all levels of health governance to achieve policy coherence (Silberschmidt & Zeltner, 2013)? The formulation of the SDGs has adopted this approach and offer opportunities to embed GHG more deeply in other sectors. Similarly, *The Lancet*–University of Oslo Commission on Global Governance for Health proposed a UN Multistakeholder Platform on Global Governance for Health that would "engage governments, intergovernmental organizations (in the areas of finance, trade, labor, food, environment, human rights, migration, and peace and security), and non-state actors including civil society, academic experts, and business" (Ottersen et al., 2014, p. 630). The Platform (**FIGURE 21-5**) would address the problem of what the Commission calls "institutional stickiness," whereby it becomes "difficult to reform institutions to evolve with the times" because power becomes entrenched. The Commission has argued that the Platform would enable "more inclusive, better integrated, and more coherent policy dialogue across institutions and arenas."

How Should Decisions Be Made Within a System of GHG?

Decision-making processes are central to governance—that is, who participates and how decisions are made. International health governance gives formal power and authority for decision making largely to state actors. This chapter describes how non-state actors have become increasingly involved in global health and, to a growing extent, in decision making, both formally and informally. Proponents argue that this "opening up" of participation enables a broader range of stakeholders to be represented in GHG decisions and, by extension, facilitates better decisions. Others, however, question whether more voices are necessarily the right voices. To what extent are the voices of big business and CSOs necessarily representative of local people, notably the poor and vulnerable? More voices also do not necessarily mean better decision making. Indeed, lessons from the Occupy Movement and the UN Climate Change Conference suggest that too many diverse voices coalescing around "amorphous structures" can be a hindrance to effective global change (Klein, 2014). Who should be enfranchised in decision making and how should they meaningfully participate in GHG? How can dissenting voices be

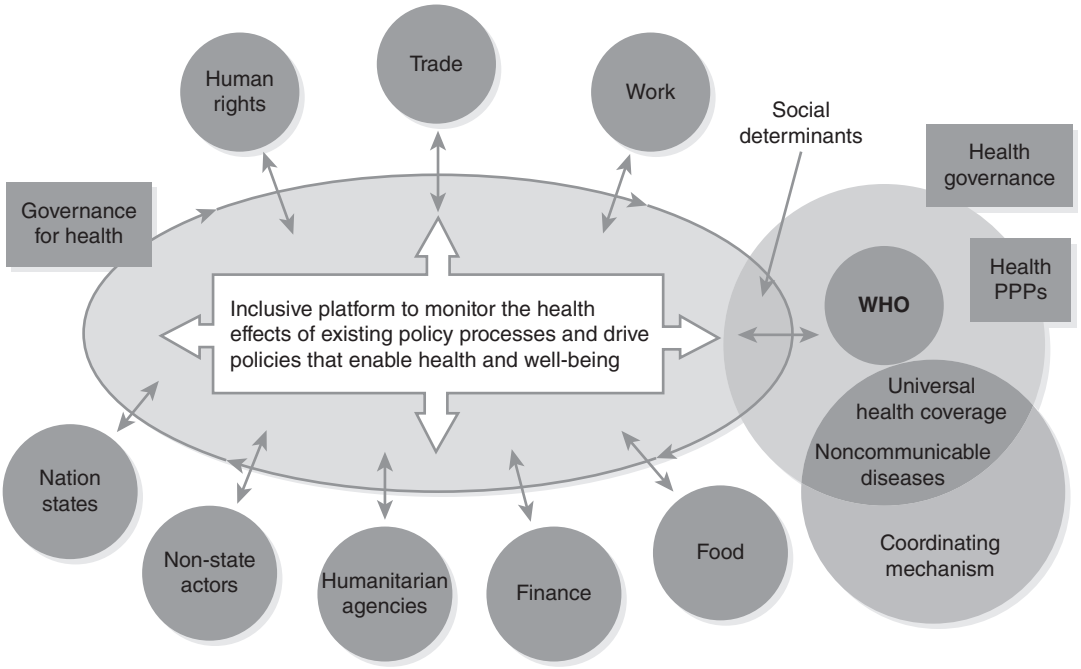


FIGURE 21-5 United Nations stakeholder platform.

Reprinted from Ottersen, O., Dasgupta, J., Blouin, C., Buss, P., Chongsuvivatwong, V., Frenk, J., . . . Scheel, I. (2014). The political origins of health inequity: prospects for change. Report of *The Lancet*-University of Oslo Commission on Global Governance for Health. *The Lancet, 383*, 630–667. Copyright © 2014, with permission from Elsevier.

heard but, most importantly, reconciled in ways that allow legitimate collective action?

Ultimately, the challenges faced by GHG cannot be reduced to technical or administrative fixes. If only a new vaccine could be developed, if only new technology could be applied more widely, if only a financial oversight committee could be formed to watch over the disbursing of available resources—these are all important achievements in GHG but will not deliver the kinds of changes needed. The key challenges in GHG are fundamentally political: They do not deal with interferences with scientific or rational action, but rather with building institutions that enable the agreed exercise of power for collective action. The proposal to create a Committee C (**EXHIBIT 21-6**), for example, would retain the primary role of WHO member states, while enabling "major stakeholders in global health—international agencies, philanthropic organizations, multinational health initiatives, and representatives from major civil-society groups, particularly those who legitimately represent the most vulnerable populations"—to feed more actively into state-led decision making (Silberschmidt, Matheson, & Kickbusch, 2008, p. 1485).

Supporting Institutional Innovation in Global Health Governance

As described in the previous section, while there is broad agreement that existing institutions face many problems, agreed ways of moving forward to address them are lacking. Proposals to strengthen GHG to date have ranged from the reform of selected international organizations (notably focused on WHO) to the design of an entirely new GHG architecture. On the one hand, numerous, and largely unsuccessful, efforts to "fix" WHO since the 1990s have produced "reform fatigue." On the other hand, the design of an entirely new system has not been supported by a clear and agreed vision of what it should look like (Gill & Benatar, 2016; Lee & Pang, 2015; Youde, 2012).

To get beyond this impasse, after so much debate and numerous initiatives, a potentially useful way forward is to draw on thinking about institutional innovation. Institutional innovation is defined by Raffaelli and Glynn (2015, p. 1) as "novel, useful, and legitimate change that disrupts the cognitive, normative or regulative mainstays of an organizational field." Institutional innovation can occur along a continuum ranging from incremental change within existing institutions (modification) to more radical change in the form of creating new institutions (disruptive innovation). Institutional innovation theory has received substantial attention in business studies in the context of explaining success and failure in companies' adaptation to fast-paced globalization. Thus, by the mid-2000s, institutional theory had shifted from a focus on explaining stability, permanence, and homogeneity, to increased attention to change and variation.

EXHIBIT 21-6 Proposal to Create a Committee C of the World Health Assembly

Currently the formal work of the WHA is focused on determining the policies of WHO itself and on making recommendations for member states. The assembly is the key annual event of global health, sometimes more because of the many informal and formal side-meetings than for its agenda. The consequence is that many major decisions in global health are taken unlinked to the formal business of the assembly. . . . The mechanism we propose is to consider the establishment of a committee C of the WHA. This committee would debate major health initiatives by other key players in the global-health arena. It would provide the opportunity for these organizations to present their plans and achievements to the delegates of the WHA and the nongovernmental organizations in official relations with WHO. It would also provide an opportunity to address coordination and common concerns of different partners in global health. Organizations wishing to make use of this mechanism would send their proposal to WHO's executive board, which would set the agenda for this committee as it does with the existing committees.

A committee C would need rules of procedures that give sufficient space to the other groups involved while fully respecting the role of governmental delegations. This double requirement could be met by proposing resolutions for adoption by the plenary of WHA as in committees A and B, but to explicitly welcome within such resolutions commitments independently taken by other partners that would be annexed to the resolution (panel). With such a procedure the other entities keep their full independence by autonomously adopting—according to their internal rules—their declarations to be annexed. There is an interaction between the member states and the other entities in committee C on the conditions and methods under which to annex a declaration and the core of the resolution. Member states keep their sovereignty in the final adoption of the resolution in the WHA plenary where other stakeholders cannot vote or intervene.

Institutional innovation has so far received limited attention in the study of public institutions, particularly international organizations. Nevertheless, some examples can be cited of institutional innovation occurring to support GHG in response to the disconnect between the forms of governance needed in a globalizing world and the institutional arrangements presently available (Smith & Lee, 2017).

One important example is innovative financing mechanisms. The two traditional sources of healthcare financing at the national level have largely been various forms of taxation and out-of-pocket payments. At the international level, financing of health governance functions has been through assessed contributions by governments according to their size and ability to pay (a form of taxation) and voluntary donations from individuals, states, and non-state actors. While funding for global health initiatives increased to unprecedented levels from the late 1990s until the global financial crisis beginning in 2008, available resources have fallen short of need. Where available funds through these two sources have not been sufficient, innovative financing mechanisms have been developed. Many alternative financing mechanisms have emerged to fill the funding gap, such as taxation of global transactions (e.g., air travel, foreign currency exchange) and crowdsourcing. Relatedly, efforts to reduce the costs of global health activities have led to important innovations such as forward purchasing, generic pricing, and guaranteed contracts.

Another area where institutional innovation is occurring is disease surveillance, monitoring, and reporting using web-based technologies. In some ways, strengthening existing capacities is likely to be most appropriate in this area of focus. A good example is the revised International Health Regulations (2005). To prevent smaller-scale outbreaks from becoming larger-scale emergencies, all WHO member states need to have a minimum level of core capacities to detect, report, and respond to such events rapidly. Governments agreed to develop these core capacities by 2012 under the revised IHR following the SARS outbreak of 2002–2003. A failure to meet this deadline by many countries led to an extension of the target time frame to 2014. Following the Ebola virus outbreak, the deadline was further extended to 2019 (Moon et al., 2015). As of 2016, and based on self-assessments, approximately half of WHO member states had yet to comply, exposing the world to similar risks in future (Lillywhite, 2016).

To support these efforts, the Global Health Security Agenda (GHSA) was launched in 2014 "to advance a world safe and secure from infectious disease threats, to bring together nations from all over the world to make new, concrete commitments, and to elevate global health security as a national leadership priority." The GHSA was subsequently endorsed by the G7, with the governments of Finland and Indonesia hosting commitment development meetings (see the *Public Health Infrastructure* chapter for more information on the GHSA). A 10-year strategic plan has also been set out by WHO to enable compliance by LMICs (WHO, 2016b). The overall lessons from the Ebola virus outbreak has been that noncompliance by LMICs needs to be addressed through financial support, rather than through changes to the IHR, which is considered "technically sound" (Lillywhite, 2016).

For other functions, however, innovations applying information and communication technologies (ICTs), open source learning, and big data offer opportunities to improve the capacity to identify, track, and even predict disease outbreaks of global concern. An early example of ICT-based innovation is ProMed (Program for Monitoring Emerging Diseases), an Internet-based reporting system established in 1994 to rapidly disseminate information on outbreaks of infectious diseases and acute exposures to toxins that affect human health, including those in animals and in plants grown for food or animal feed. ProMed currently reaches more than 70,000 subscribers in at least 185 countries. It is open to all sources of information, including media reports, official reports, online summaries, local observers, and ProMed subscribers. A team of expert human, plant, and animal disease moderators screen, review, and investigate reports before posting them to the network. Reports are then distributed by email to direct subscribers and posted immediately on the ProMed website (ProMed, n.d.).

The Global Public Health Intelligence Network operates in a similar way by identifying outbreaks from websites, news wires, and local and national newspapers retrieved through news aggregators. Global Pulse is a flagship initiative of the United Nations Secretary-General that is seeking to harness big data safely and responsibly as a public good. A good example of the innovative use of big data (extremely large data sets that require analysis computationally to reveal patterns, trends, and associations, especially related to human behavior and interactions) for global health is GermTracker (**EXHIBIT 21-7**).

Another example of institutional innovation has been representation in decision making. At the heart of GHG innovation is finding appropriate ways for decisions to be made collectively to identify and achieve shared goals. A longstanding concern with WHO has been the limited voice given to civil society organizations, which have become increasingly

EXHIBIT 21-7 GermTracker

Researchers at the University of Rochester showed last year how Twitter can be used to predict how likely it is for a Twitter user to become sick. They have now used Twitter to model how other factors—social status, exposure to pollution, interpersonal interaction, and others—influence health.

"If you want to know, down to the individual level, how many people are sick in a population, you would have to survey the population, which is costly and time-consuming," said Adam Sadilek, postdoctoral researcher at the University of Rochester. "Twitter and the technology we have developed allow us to do this passively, quickly, and inexpensively; we can listen in to what people are saying and mine this data to make predictions."

Many tweets are geo-tagged, which means they carry Global Positioning System (GPS) information that shows exactly where the user was when he or she tweeted. Collating all this information allows the researchers to map out, in space and in time, what people said in their tweets, as well as where they were and when they were there. By following thousands of users as they tweet and go about their lives, researchers also could estimate interactions between two users and between users and their environment.

Using tweets collected in New York City over a period of a month, the researchers looked at factors such as how often a person takes the subway, goes to the gym, or visits a particular restaurant; proximity to a pollution source; and online social status. They examined 70 factors in total. They then looked at whether these factors had a positive, negative, or neutral impact on the users' health.

Some of their results are perhaps not surprising. For example, pollution sources seem to have a negative effect on health. However, the study marked the first time that this impact has been extracted from the online behavior of a large online population. The paper also reveals a broader pattern, in which almost any activity that involves human contact leads to significantly increased health risks. For example, even people who regularly go to the gym get sick marginally more often than less active individuals. Nevertheless, people who merely talk about going to the gym, but actually never go (verified based on their GPS data), get sick significantly more often. Thus, interesting confounding factors can now be studied at scale.

The technology that Sadilek and his colleague Professor Henry Kautz developed has led to a web application called GermTracker. The application color-codes users (from red to green) according to their health by mining information from their tweets for 10 cities worldwide. Using the GPS data encoded in the tweets, the app can then place people on a map, which allows anyone using the application to see their distribution.

"This app can be used by people to make personal decisions about their health. For example, they might want to avoid a subway station if it's full of sick people," Sadilek suggested. "It could also be used in conjunction with other methods by governments or local authorities to try to understand outbursts of the flu."

During flu season, as the number of people with the flu across the United States increases, so do the number of people monitoring GermTracker. On some days in January, as many as 10,000 people visit http://fount.in where the app is hosted.

The model that Sadilek and his colleagues developed is based on machine learning. At the heart of their work is an algorithm that can distinguish between tweets that suggest the person tweeting is sick and those that do not.

"It's like teaching a baby a new language," Sadilek said. He explained that they first generated a training set of data—5,000 tweets that had been manually categorized and from which the algorithm can start to distinguish what words and phrases are associated with someone being sick. He added, "We need the algorithm to understand that someone who tweets 'I'm sick and have been in bed all day' should be characterized as sick, but 'I'm sick of driving around in this traffic' shouldn't be."

The application is also improving the algorithm. Every time users access the app and click on one of the colored dots that represent the tweeting users, they can see the specific tweet that led someone to be classified in a specific way. The application asks users to assess the tweet and say whether they agree with the classification. These data are fed back into the algorithm, which continues to learn from its mistakes.

University of Rochester. (2013). Using Twitter to predict the influence of lifestyle on health. http://www.rochester.edu/news/show.php?id=5532

important in global health activities. Historically, WHA has been the plenary body through which governments set priorities. Officially recognized non-state actors can be given observer status at the WHA, for example, but obtaining this status is a highly selective process controlled by states.

Overall, change has been too slow and too limited for many CSOs seeking to inject a more critical voice in WHO deliberations. The previously described limitations of WHO as an authoritative decision-making body raise the question of what kind of body is needed. Some have suggested an enhanced WHA through a committee C involving additional constituencies beyond member state delegations, such as CSOs, charitable foundations, and other non-state actors (Exhibit 21-6). The People's Health Assembly, a consortium of CSOs, was initiated by organizations that did not feel represented by existing institutional

arrangements, which focused on WHO and member states. The four *Global Health Watch* (2014) reports were published as an alternative to the *World Health Report* published by WHO—they sought to offer a critical voice. Each report was intended to serve as

> the definitive voice for an alternative discourse on health and health care. It covers a range of issues that currently impact on health, including the present political and economic architecture in a fast-changing and globalized world; a political assessment of the drive towards Universal Health Coverage; broader determinants of health, such as gender-based violence and access to water; stories of struggles, actions and change; and a scrutiny of a range of global institutions and processes. It integrates rigorous analysis, alternative proposals and stories of struggle and change to present a compelling case for a radical transformation of the way we approach actions and policies on health (Global Health Watch, 2014, Overview).

Conclusion

The health impacts of contemporary globalization are now widely documented, and are recognized as being more extensive and intensive than the impacts that occurred in earlier periods of social change. The resultant effects on health determinants and outcomes, affecting individuals and populations worldwide, pose fundamental challenges to how we govern health from the local level to the planetary level. "Global health" is now a widely used and accepted term, albeit one laden with imprecision and redundant meanings. A full transition in thinking and practice entails more than a change in nomenclature; it requires embracing a paradigmatic shift about health determinants and outcomes. An understanding of the distinct meaning of global health, in turn, leads to different expectations about collective active to govern the health impacts of globalization.

The transition from international to global health governance is also a work in progress, characterized by conceptual fuzziness and contested views about what should be achieved and how to achieve it. Numerous challenges remain to be addressed, but many opportunities are available to approach these challenges using institutional innovation. This is already beginning to happen, resulting in the strengthening of some GHG functions in selected issue areas. It is more likely that the global health community will continue to move incrementally in this way, rather than adopt a wholesale redesign of existing institutional arrangements.

Discussion Questions

1. Which core functions do you believe an effective system of global health governance should serve?
2. How might the concept of network governance be used to link together existing institutional arrangements for global health? Are there any governance gaps that need filling?
3. Which stakeholders should be involved in making decisions concerning global health? How should participation in decision making be governed?
4. Which functions might non-state actors perform in global health governance?

References

Aginam, O. (2005). *Global health governance, international law and public health in a divided world*. Toronto, Canada: University of Toronto Press.

Bayne, N., & Woolcock, S. (2007). *The new economic diplomacy: Decision-making and negotiation in international economic relations*. London, UK: Ashgate.

Beck, U. (2015). *What is globalization?* London, UK: Polity Press.

Benedick, R. E. (1998). Diplomacy for the environment. In American Institute for Contemporary German Studies, *Environmental diplomacy* [Conference report] (pp. 3–12). Washington, DC: Johns Hopkins University.

Birn, A. E. (2014). Philanthrocapitalism, past and present: The Rockefeller Foundation, the Gates Foundation, and the setting(s) of the international/global health agenda. *Hypothesis, 12*(1), e8. doi:10.5779/hypothesis. v12i1.229

Birn, A. E., & Dmitrienko, K. (2005). The World Bank: Global health or global harm? *American Journal of Public Health, 95*(7), 1091–1092.

Birn, A. E., Pillay, Y., & Holtz, T. (Eds.). (2017). *Textbook of global health* (4th ed.). Oxford, UK: Oxford University Press.

Bill & Melinda Gates Foundation (BMGF). (n.d.). Who we are, Executive Management Team. Retrieved from

https://www.gatesfoundation.org/Who-We-Are/General-Information/Leadership/Executive-Leadership-Team

Bollier, D. (2003). *The rise of Netpolitik: How the Internet is changing international politics and diplomacy*. Washington DC: Aspen Institute.

Boseley, S. (2011, September 19). Global Fund verdict could be devastating for the fight against AIDS. *Guardian*. Retrieved from https://www.theguardian.com/society/sarah-boseley-global-health/2011/sep/19/aids-malaria

Brolan, C., & Hill, P. (2016). Universal health coverage's evolving location in the post-2015 development agenda: Key informant perspectives within multilateral and related agencies during the first phase of post-2015 negotiations. *Health Policy and Planning, 31*(4), 514–526.

Brown, T., Cueto, M., & Fee, E. (2006). The World Health Organization and the transition from "international" to "global" public health. *American Journal of Public Health, 96,* 62–72.

Buse, K., & Hawkes, S. (2016). Sitting on the FENSA: WHO engagement with industry. *The Lancet, 388*, 446–447.

C20. (n.d.). About C20. Retrieved from https://civil-20.org/home/about-c20/

Chan, M. (2007, February 13). Health diplomacy in the 21st century. Retrieved from http://www.who.int/dg/speeches/2007/130207_norway/en/

Cohn, S. (2002). The Black Death: End of a paradigm. *American Historical Review, 107*(3), 703–738.

Coker, R., & Mounier-Jack, S. (2006). Pandemic influenza preparedness in the Asia-Pacific region. *The Lancet, 368*(9538), 886–889.

Dahl, J. (2017, April 26). How China is on course to unseat U.S. as the next leader in global health. *Forbes*. Retrieved from https://www.forbes.com/sites/jordyndahl/2017/04/26/how-china-is-on-course-to-unseat-u-s-as-the-next-leader-in-global-heath/#c8a7a2c467a

Davenport, D. (2002). The new diplomacy. *Policy Review, 116*, 17-30.

Dion, M., AbdelMalik, P., & Mawudeku, A. (2015). Big data and the Global Public Health Intelligence Network (GPHIN). *Canada Communicable Disease Report Monthly, 41*(9), 209–214.

Dodgson, R., Lee, K., & Drager, N. (2002). *Global health governance: A conceptual review*. Discussion Paper No. 1. Geneva, Switzerland: WHO Department for Health in Sustainable Development.

European Centre for Disease Prevention and Control. (n.d.). *Aedes aegypti:* Factsheet for experts. Retrieved from https://ecdc.europa.eu/en/disease-vectors/facts/mosquito-factsheets/aedes-aegypti

Evans, G., & Newnham, J. (1992). *The dictionary of world politics*. London, UK: Harvester Wheatsheaf.

Gerdil, C. (2003). The annual production cycle for influenza vaccine. *Vaccine*, *21*(16), 1776–1779.

Gill, S., & Benatar, S. (2016). Global health governance and global power: A critical commentary on *The Lancet*–University of Oslo Commission report. *International Journal of Health Services, 46*(2), 346–365.

Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM). (2017). Board: Overview. Retrieved from https://www.theglobalfund.org/en/board/

Global Health Watch. (2014). *Global Health Watch 4, An Alternative World Health Report*. London, UK: Zed Books.

Go4Health. (2013). Universal health coverage anchored in the right to health. *Bulletin of the World Health Organization, 91,* 2–2A. doi:10.2471/BLT.12.115808

Godlee, F. (1994). The World Health Organisation: WHO in crisis. *British Medical Journal, 309,* 1424.

Gostin, L., Friedman, E., Buse, K., Waris, A., Mulumba, M., Joel, M., . . . Sridhar, D. (2012). Towards a framework convention on global health. *Bulletin of the World Health Organization, 91,* 790–793. doi:10.2471/BLT.12.114447

Grepin, K., Fan, V., Shen, G., & Chen, L. (2014). China's role as a global health donor in Africa: What can we learn from studying under reported resource flows? *Globalization and Health, 10,* 84. doi:10.1186/s12992-014-0084-6

Group of 20. (2017). Berlin declaration of the G20 health ministers: Together today for a healthy tomorrow. Berlin: G20 Information Centre. Retrieved from http://www.g20.utoronto.ca/2017/170520-health-en.html

Haas, P. (2004). Addressing the global governance deficit. *Global Environmental Politics, 4*(4), 1–19.

Harman, S. (2016). The Bill & Melinda Gates Foundation and legitimacy in global health governance. *Global Governance, 22*(3), 349–368.

Health diplomacy: Rx for peace. (2007, August 26). *Washington Times*, pp. B4.

Heine, J. (2006). *On the manner of practising the new diplomacy*. Working Paper No. 22. Waterloo, Canada: Centre for International Governance and Innovation, University of Waterloo.

Hill, P., Buse, K., Brolan, C. E., & Ooms, G. (2014). How can health remain central post-2015 in a sustainable development paradigm? *Globalization and Health, 10,* 18.

Horton, R. (2015). Offline: A pervasive failure to learn the lessons of Ebola. *The Lancet, 386*(9998), 1024.

Huynen, M., Martens, P., & Hilderink, H. (2005). The health impacts of globalization: A conceptual framework. *Globalization and Health, 1,* e14.

Institute for Health Metrics and Evaluation. (2017). *Financing global health 2016*. Seattle, WA: University of Washington.

International Labour Organization (ILO). (2015). *World social protection report 2014/15*. Geneva, Switzerland: Author.

Kates, J., & Katz, R. (2010). *U.S. participation in international health treaties, commitments, partnerships, and other agreements*. Washington, DC: Henry J. Kaiser Family Foundation.

Kickbusch, I., & Kökény, M. (2013). Global health diplomacy: Five years on. *Bulletin of the World Health Organization, 91,* 159–159A. Retrieved from http://www.who.int/bulletin/volumes/91/3/13-118596/en/

Klein, N. (2014) *This changes everything: Capitalism versus the climate*. Toronto, Canada: Knopf Canada.

Koplan, J. P., Bond, T. C., Merson, M., Reddy, K. S., Rodriguez, M. H., Sewankambo, N., & Wasserheit, J. (2009). Towards a common definition of global health. *The Lancet, 373,* 1993–1995.

Labonte, R., & Schrecker, T. (2007). Globalization and social determinants of health: Introduction and methodological background. *Globalization and Health, 3,* e5.

Lee, K. (2003). *Globalization and health: An introduction*. London, UK: Palgrave Macmillan.

Lee, K. (2010). Civil society and intergovernmental organisations: The functions of global health governance. *Global Health Governance, 3*(2). 1–20. Retrieved from https://blogs.shu.edu/ghg/files/2011/11/Lee_Civil-Society-Organizations-and-the-Functions-of-Global-Health-Governance_Spring-2010.pdf

Lee, K., Collinson, S., Walt, G., & Gilson, L. (1996). Who should be doing what in international health: A confusion of mandates in the United Nations? *British Medical Journal, 312*, 302–307.

Lee, K., & Dodgson, R. (2000). Globalisation and cholera: Implications for global governance. *Global Governance, 6*(2), 213–236.

Lee, K., Hawkins, B., & Wiist, W. (2016). Researching corporations and global health governance. In K. Lee & B. Hawkins (Eds.), *Researching corporations and global health governance: An interdisciplinary guide* (pp. 1–14). London, UK: Rowman and Littlefield International.

Lee, K., & Pang, T. (2014). WHO: Retirement or reinvention? *Public Health, 128*(2), 119–123.

Lee, K., & Smith, R. (2011). What is "global health diplomacy"? A conceptual review. *Global Health Governance, 5*(1), 1-12. Retrieved from https://blogs.shu.edu/ghg/files/2011/11/Lee-and-Smith_What-is-Global-Health-Diplomacy_Fall-2011.pdf

Lewis, M., & MacPherson, K. (Eds.). (2013). *Health transitions and the double disease burden in Asia and the Pacific*. London, UK: Routledge

Lillywhite, L. (2016, April 20). Expert comment: International Health Regulations must overcome perceptions of double standard. Chatham House. Retrieved from https://www.chathamhouse.org/expert/comment/international-health-regulations-must-overcome-perceptions-double-standard

Mackey, T., & Liang, B. (2013). Restructuring brain drain: Strengthening governance and financing for health worker migration. *Global Health Action, 6.* doi:10.3402/gha.v6i0.19923

Martens, P., Akin, S. M., Maud, H., & Mohsin, R. (2010). Is globalization healthy? A statistical indicator analysis of the impacts of globalization on health. *Globalization and Health, 6,* e16.

McArthur, J. (2014). The origins of the Millennium Development Goals. *SAIS Review, 34*(2), 5–24.

McInnes, C., & Lee, K. (2006). Health, foreign policy and security. *Review of International Studies, 32*(1), 5–23.

Moon, S., Sridhar, D., Pate, M., Jha, A., Clinton, C., Delauney, S., . . . Piot, P. (2015). Will Ebola change the game? Ten essential reforms before the next pandemic. The report of the Harvard–LSHTM Independent Panel on the Global Response to Ebola. *The Lancet*, *386*, 2204–2221.

New Development Bank. (n.d.). About us. Retrieved from http://www.ndb.int/about-us/essence/history/

Nunes, A. R., Lee, K., & O'Riordan, T. (2016). The importance of an integrating framework for achieving the Sustainable Development Goals: The example of health and well-being. *British Medical Journal Global Health, 1,* e000068.

Ottersen, O., Dasgupta, J., Blouin, C., Buss, P., Chongsuvivatwong, V., Frenk, J., . . . Scheel, I. (2014). The political origins of health inequity: Prospects for change. Report of *The Lancet*–University of Oslo Commission on Global Governance for Health. *The Lancet, 383*, 630–667.

Pagett, C., & Padarath, A. (2007). *A review of codes and protocols for the migration of health workers*. Cape Town, South Africa: Regional Network for Equity in Health in East and Southern Africa (EQUINET) & Health Systems Trust. Retrieved from http://www.who.int/workforcealliance/15.pdf?ua=1

ProMed. (n.d.). About ProMed-mail. Retrieved from http://www.promedmail.org/aboutus/

Raffaelli, R., & Glynn, M.A. (2015). Institutional innovation: Novel, useful and legitimate. In C. Shalley, M. Hitt, & J. Zhou (Eds.), *Oxford handbook of creativity, innovation and entrepreneurship*. Oxford, UK: Oxford University Press.

Reynolds, P. A. (1980). *An introduction to international relations* (2nd ed.). New York, NY: Longman.

Riordan, S. (2002). *The new diplomacy*. London, UK: Polity.

Röhren, C. (2017, April 22). Why China could be a game changer for global health. *Diplomat*. Retrieved from http://thediplomat.com/2017/04/why-china-could-be-a-game-changer-for-global-health

Rosenau, J. N. (1995). Governance in the twenty-first century. *Global Governance, 1*(1).

Ryngaert, C. (2016) *Non-state actor dynamics in international law: From law-takers to law-makers*. London, UK: Routledge.

Scholte, J. A. (2005). *Globalization: A critical introduction* (2nd ed.). London, UK: Palgrave Macmillan.

Scholte, J. A. (2008). Defining globalisation. *World Economy, 31*(11), 1471–1502. Retrieved from http://onlinelibrary.wiley.com.proxy.lib.sfu.ca/doi/10.1111/j.1467-9701.2007.01019.x/epdf

Silberschmidt, G., Matheson, D., & Kickbusch, I. (2008). Creating a committee C of the World Health Organisation. *The Lancet, 371,* 1483–1486.

Silberschmidt, G., & Zeltner, T. (2013). Global health begins at home: Policy coherence. In T. Novotny, I. Kickbusch, & M. Told (Eds.), *21st century global health diplomacy* (pp. 279–298). Singapore: World Scientific.

Siyam, A., Zurn, P., Ro, O., Gedik, G., Ronquillo, K., Co, C., . . . Dal Poz, M. (2013). Monitoring the implementation of the WHO Global Code of Practice on the International Recruitment of Health Personnel. *Bulletin of the World Health Organization, 91*(11), 816–823.

Smith, K., Goldberg, M., Rosenthal, S., Carlson, L., Chen, J., Chen, C., & Ramachandran, S. (2014). Global rise in human infectious disease outbreaks. *Journal of the Royal Society Interface, 11*(101).

Smith, R., & Lee, K. (2017). Global health governance: We need innovation, not renovation. *British Medical Journal Global Health, 2,* e000275.

Smith, R., Lee, K., & Fidler, D. (2009). *Global health diplomacy: A synthesis paper*. Geneva, Switzerland/New York, NY: World Health Organization & Rockefeller Foundation.

Smith, R., & MacKellar, L. (2007). Global public goods and the global health agenda: Problems, priorities and potential. *Globalization and Health, 3,* 9.

Solana, J. (2009, January 20). Five lessons in global diplomacy. *Financial Times*.

Stiglitz, J. (2007). *Making globalization work*. New York, NY: W. W. Norton & Company.

Suleman, M., Ali, R., & Kerr, D. (2014). Health diplomacy? A new approach to the Muslim world? *Globalization and Health, 10,* 50. doi:10.1186/1744-8603-10-50

Sustainable Development Solutions Network (SDSN). (2014). *Health in the framework of sustainable development: Technical report for the Post-2015 Development Agenda*. New York, NY: Author. Retrieved from http://unsdsn.org/wp-content/uploads/2014/02/Health-For-All-Report.pdf

Tankwanchi, A., Vermund, S., & Perkins, D. (2014). Has the WHO Global Code of Practice on the International Recruitment of Health Personnel been effective? *The Lancet Global Health, 2*(7), e390–e391.

Taylor, A. (2013). Global health law, concepts, issues, actors, instruments, fora and cases. In I. Kickbusch, G. Lister, M. Todd, & N. Drager (Eds.), *Global health diplomacy, concepts, issues, actors, instruments, fora and cases* (pp. 37–54). New York, NY: Springer.

Taylor, A., Hwenda, L., Larsen, B. I., & Daulaire, N. (2011). Stemming the brain drain: A WHO Global Code of Practice on

International Recruitment of Health Personnel. *New England Journal of Medicine, 365,* 2348–2351.

UN Millennium Project. (2006). What they are. Retrieved from http://www.unmillenniumproject.org/goals/

UNITAID. (2017). About us: Governance. Retrieved from https://unitaid.eu/about-us/governance/

University of Rochester. (2013). Tracking influenza via Twitter. Retrieved from https://www.youtube.com/watch?v=w_j0VuJ-J9k

United Kingdom Department of Health. (2008). *Strategy for global health 2008–2013*. London, UK: Author.

Wagner, L. M. (2015, September 17). The breakthrough secret to negotiating the SDGs. *Huffington Post*. Retrieved from http://www.huffingtonpost.com/lynn-m-wagner/the-breakthrough-secret-to-negotiating-the-sdgs_b_8148064.html

Watts, S. (2003). *Disease and medicine in world history*. London UK: Routledge.

Weiss, T., Seyle, D., & Coolidge, K. (2013). *The rise of non-state actors in global governance opportunities and limitations* [Discussion paper]. Broomfield, CO: One Earth Future Foundation. Retrieved from http://acuns.org/wp-content/uploads/2013/11/gg-weiss.pdf

Whitmee, S., Haines, A., Beyrer, C., Boltz, F., Capon, A., de Souza Dias, B., . . . Yach, D. (2015). Safeguarding human health in the Anthropocene epoch: Report of the Rockefeller Foundation–*The Lancet* Commission on planetary health. *The Lancet, 386,* 1973–2028.

World Health Organization (WHO). (1946). *Constitution of the World Health Organization*. New York, NY: United Nations.

World Health Organization (WHO). (2006). *World health report 2006: Working together for Health*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2009). *Global health risks, mortality and burden of disease attributable to selected health risks*. Geneva, Switzerland: Author.

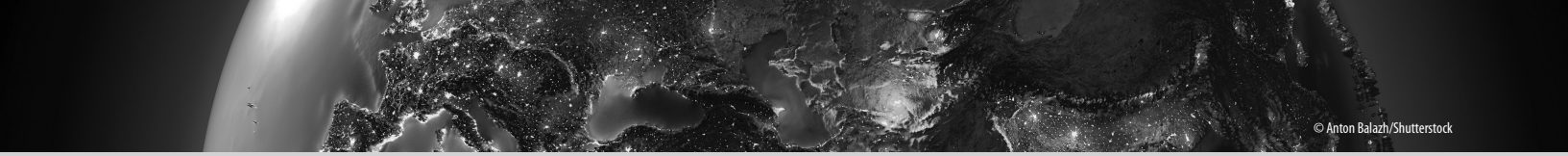
World Health Organization (WHO). (2016a, May 28). Framework of engagement with non-state actors. 69th World Health Assembly, WHA69.10. Retrieved from http://www.who.int/about/collaborations/non-state-actors/A69_R10-FENSA-en.pdf?ua=1

World Health Organization (WHO). (2016b). Implementation of the International Health Regulations (2005). Report of the Review Committee on the Role of the International Health Regulations (2005) in the Ebola Outbreak and Response. 69th World Health Assembly, Doc. A69/. Retrieved from http://apps.who.int/gb/ebwha/pdf_files/WHA69/A69_21-en.pdf

World Health Organization (WHO). (n.d.). *Managing health workforce migration: The Global Code of Practice*. Geneva, Switzerland: Human Resources for Health. Retrieved from http://www.who.int/hrh/migration/code/practice/en/

Youde, J. (2012). *Global health governance*. London, UK: Polity Press.

Youde, J. (2013). The Rockefeller and Gates Foundations in global health governance. *Global Society, 27*(2), 139–158.



Acronyms

ACC/SCN	Administrative Committee on Coordination/Sub-Committee on Nutrition (UN)
ACF	Action Contre la Faim
ACSD	Accelerated Child Survival and Development program (UNICEF)
ACT	artemisinin-based combination therapy
ADDO	accredited drug dispensing outlet
ADE	adverse drug event
ADHD	attention-deficit/hyperactivity disorder
ADLs	activities of daily living
ADR	adverse drug reaction
AFENET	African Field Epidemiology Network
AGA	appropriate weight for gestational age
AGP	alpha-1 acid glycoprotein
AI	Adequate Intake
AIDS	acquired immunodeficiency syndrome
ALRI	acute lower respiratory infection
ALT	alanine aminotransferase
AMC	advance market commitment
AMR	antimicrobial resistance
ANC	antenatal care
APA	American Psychiatric Association
ARI	acute respiratory infection
ART/ARV	antiretroviral therapy
ASEAN	Association of Southeast Asian Nations
ASEOWA	African Union Support for the Ebola Outbreak in West Africa
ASMR	age-specific mortality rates
ASP	antimicrobial stewardship program
AST	aspartate aminotransferase
AYUSH	Ayurveda, yoga and naturopathy, Unani, Siddha, homeopathy
AZT	Azidothymidine
BAC	blood alcohol concentration
BCC	behavioral change communication
BCG	bacillus Calmette-Guérin (vaccine)
BMGF	Bill and Melinda Gates Foundation
BMI	body mass index
BOP	balance of payments
BP	*British Pharmacopoeia*
BRICS	Brazil, Russia, India, China, and South Africa
CAPI	computer-assisted personal interviewing
CARA	Conscience of Antimicrobial Resistance Accountability (initiative)
CARIFTA	Caribbean Free Trade Agreement
CBH	community-based health
CBHI	community-based health insurance
CBR	community-based rehabilitation
CBRN	chemical, biological, radiological, and nuclear
CBT	cognitive-behavioral therapy
CCD	cultural concept of distress
CCM	country coordination mechanism
CCMD	Chinese Classification of Mental Disorders
CCT	conditional cash transfer
CDC	Centers for Disease Control and Prevention (U.S.)
CDR	case disability ratio
CE	complex emergency
CECA	Comprehensive Economic Cooperation Agreement (India–Singapore)
CEF	Contingency Fund for Emergencies (WHO)
CETA	Common Elements Treatment Approach
CFI	cultural formulation interview
CFR	case fatality ratio
CERF	Central Emergency Response Fund (UN)
CHD	coronary heart disease
CHE	complex humanitarian emergency
CHeSS	Country Health Systems Surveillance
CHF	Community Health Fund (Tanzania)
CHW	community health worker
CI	confidence interval
CIDT	Community Informant Detection Tool

CIT	crisis intervention team
CITEM	Center for International Trade Expositions and Missions (Philippines)
CKD	chronic kidney disease
CLAS	*Cultural and Linguistically Appropriate Services* guidelines
CMAM	community management of acute malnutrition
CMD	common mental disorder
CMR	crude mortality rate
CMS	central medical store
COPD	chronic obstructive pulmonary disease
CPAP	continuous positive airway pressure
CRA	comparative risk assessment
CRC	child-resistant container
CRP	C-reactive protein
CSDH	Commission on Social Determinants of Health (WHO)
CSO	civil society organization
CSR	corporate social responsibility
CVD	cardiovascular disease
cVDPV	circulating, vaccine-derived polio viruses
CVI	Childhood Vaccine Initiative
DAH	development assistance for health
DALE	disability-adjusted life expectancy
DALY	disability-adjusted life year
DBT	dialectical behavioral therapy
DCP	*Disease Control Priorities*, third edition
DDT	dichlorodiphenyltrichloroethane
DFATD	Department of Foreign Affairs, Trade and Development (Canada)
DFID	Department for International Development (U.K.)
DFLE	disability-free life expectancy
DHHS	Department of Health and Human Services (U.S.)
DHMT	district health management team
DHS	demographic and health survey; district health system
DOTS	directly observed treatment, short course
DPP	Diabetes Prevention Program
DPSEEA	driving forces, pressures, state of environment, exposures, effects, and actions (model)
DPT	diphtheria, pertussis, tetanus vaccine
DRC	Democratic Republic of Congo
DRR	disaster risk reduction
DSM-IV	*Diagnostic Statistical Manual of Mental Disorders, Fourth Edition* (American Psychiatric Association)
DSM-5	*Diagnostic Statistical Manual of Mental Disorders, Fifth Edition* (American Psychiatric Association)
DSS	demographic surveillance site/system
EAR	Estimated Average Requirement
ECD	early childhood development
ECDC	European Centres for Disease Control
ECOWAS	Economic Community of West African States
EDCTP	European and Developing Countries Clinical Trials Partnership
EHRA	environmental health risk assessment
EI	extractive industry
EIA	environmental impact assessment
EIS	Epidemic Intelligence Service (CDC)
ELISA	enzyme-linked immunosorbent assay
EMA	European Medicines Agency
EMERALD	Emerging Mental Health Systems in Low- and Middle-Income Settings
EMRO	Eastern Mediterranean Regional Office (WHO)
EMT	Emergency Medical Team
ENACT	Enhancing Assessment of Common Therapeutic Factors
ENMR	early neonatal mortality rate
EOC	emergency operations center
EPEC	enteropathogenic *Escherichia coli*
EPHF	essential public health functions
EPI	Expanded Program on Immunization
EPIDOS	European Patent Information and Documentation Systems
EPIET	European Programme for Intervention Epidemiology Training (ECDC)
EQ-5D	European Quality of Life with Five Domains
ERA	environmental risk assessment
ERC	Emergency Relief Coordinator (UN)
ERHI	employment-related health insurance
ESG	environmental, social, and governance (indicators)
ESKAPE	*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter*
ETS	environmental tobacco smoke
EU	European Union
FAO	Food and Agriculture Organization
FCTC	Framework Convention on Tobacco Control

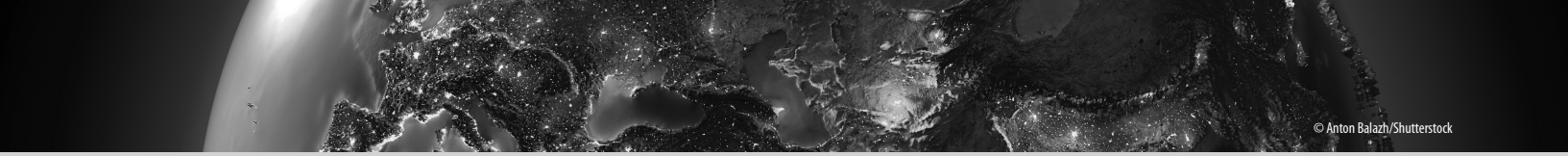
FDA Food and Drug Administration (U.S.)
FDC fixed-dose combination
FDI foreign direct investment
FELTP Field Epidemiology and Laboratory Training Program
FENSA Framework for Engagement with Non-State Actors
FETP Field Epidemiology Training Program
FMD Falsified Medicines Directive (EU)
FTA free trade agreement
G8 Group of Eight
G20 Group of 20
GAMDI Global Alliance for Medical Diagnostics Initiative
GATS General Agreement on Trade in Services (WTO)
GATT General Agreement on Tariffs and Trade
GAVI Global Alliance for Vaccines and Immunizations
GBD Global Burden of Disease
GBV gender-based violence
GCGMH Grand Challenges in Global Mental Health
GDI gross domestic income
GDM gestational diabetes mellitus
GDP gross domestic product
GFATM Global Fund to Fight AIDS, Tuberculosis and Malaria (Global Fund)
GHD global health diplomacy
GHG global health governance
GHI global health initiative
GHO Global Health Observatory
GHS Ghana Health Service
GHSA Global Health Security Agenda
GIS geographic information system
GISRS Global Influenza Surveillance and Response System (WHO)
GLASS Global AMR Surveillance System
GloPID-R Global Research Collaboration for Infectious Disease Preparedness
GMP Good Pharmaceutical Manufacturing Practice
GNI gross national income
GNP gross national product
GOARN Global Outbreak Alert and Response Network
GOBI growth monitoring, oral rehydration, breastfeeding, and immunization
GPHF Global Pharma Health Fund (Germany)
GPHIN Global Public Health Intelligence Network
GPPP global public–private partnership
GTS Global Technical Strategy for Malaria 2016–2030
HAI Health Action International
HALE health-adjusted life expectancy
HANCI Hunger and Nutrition Commitment Index
HARITA Horn of Africa Risk Transfer for Adaptation Program
HAV hepatitis A virus
HAZ height-for-age *z*-score
HBV hepatitis B virus
HCE healthcare expenditure
HCV hepatitis C virus
HDA Health Development Army (Ethiopia)
HDP health-damaging pollutant
HDV hepatitis D virus
HeaLY healthy life years
HEV hepatitis E virus
HEW health extension worker
HF hemorrhagic fever
HIA health impact assessment
HiAP Health in All Policies
Hib *Haemophilus influenzae* type b
HIC high-income country
HIPAA Health Insurance Portability and Accountability Act (U.S.)
HiTAP Thai Health Intervention and Technology Assessment Program
HIV human immunodeficiency virus
HMIS health management information system
HMN Health Metrics Network
HPAI highly pathogenic avian influenza
HPV human papillomavirus
HR hazard ratio; human resources
HSA health surveillance assistant
IADLs instrumental activities of daily living
IAH intersectoral action of health
IANPHI International Association of National Public Health Institutes
IARC International Agency for Research on Cancer
IASC Inter-Agency Standing Committee
IAVI International AIDS Vaccine Initiative
ICCC Innovative Care for Chronic Conditions
ICCIDD International Council for the Control of Iodine Deficiency Disorders

iCCM Integrated Community Case Management
ICD *International Classification of Diseases*
ICD-10 *International Classification of Diseases—Tenth Revision*
ICD-11 *International Classification of Diseases—Eleventh Revision*
ICER incremental cost-effectiveness ratio
ICESCR International Covenant on Economic, Social and Cultural Rights
ICF *International Classification of Functioning, Disability, and Health*
ICIDH *International Classification of Impairments, Disabilities, and Handicaps*
ICMI Integrated Management of Childhood Illness
ICPD International Conference on Population Development
ICRC International Committee of the Red Cross
IDD iodine-deficiency disorders
IDDO Infectious Diseases Data Observatory
IDP internally displaced person
IDRC International Development Research Centre
IDSR Integrated Disease Surveillance and Response (Pakistan)
IFA iron–folic acid (supplement)
IFPMA International Federation of Pharmaceutical Manufacturers & Associations
IGO international governmental organization
IHME Institute for Health Metrics and Evaluation
IHP International Health Partnership
IHR International Health Regulations
ILO International Labour Organization
IMAI Integrated Management of Adult Illness
IMCI Integrated Management of Childhood Illness
IMF International Monetary Fund
IMNCI Integrated Management of Newborn and Childhood Illness
IMPACT International Medical Products Anti-Counterfeiting Taskforce
IMR infant mortality rate
INCAP Instituto Nutricional de Central America y Panama
INFORMAS International Network for Food and Obesity/NCDs Research, Monitoring, and Action Support
INGO international nongovernmental organization
INRUD International Network for Rational Use of Drugs
IP intellectual property
IPC infection prevention and control
IPCC Intergovernmental Panel on Climate Change (UN)
IPR intellectual property rights
IPT interpersonal therapy
IPV injectable polio vaccine
IQ intelligence quotient
IS implementation strength
ISARIC International Severe Acute Respiratory and Emerging Infection Consortium
ISC International Sanitary Convention
ISH International Society for Hypertension
ITN insecticide-treated net
IUD intrauterine device
IUGR intrauterine growth restriction
IYC infants and young children
IYCF infant and young child feeding
JEE Joint External Evaluation
JPEPA Japan–Philippines Economic Partnership Agreement
LAZ length-for-age *z*-score
LBW low birth weight
LGBTI lesbian, gay, bisexual, transgender, and intersex
LI lifestyle index
LiST Lives Saved Tool
LMICs low- and middle-income countries
LNS lipid-based nutrient supplements
LPG liquefied petroleum gas
MAC *Mycobacterium avium* complex
MCE multi-country evaluation
MCH mean cell hemoglobin; maternal and child health
MCHC mean cell hemoglobin concentration
MCV mean cell volume
MDGs Millennium Development Goals
MDR-TB multidrug-resistant tuberculosis
MERS Middle East respiratory syndrome
MHIN Mental Health Innovation Network
MHSSS Mental Health Services Satisfaction Scale

MHW	mental health worker
MI	myocardial infarction
MICS	Multiple Indicator Cluster Surveys (UNICEF)
MIS	management information system; malaria indicator survey
MISP	minimum initial service package
MLEM	Model List of Essential Medicines (WHO)
MMR	maternal mortality ratio
MMV	Medicines for Malaria Venture
MNCH	maternal, newborn, and child health
MNS	mental, neurologic, and substance use disorders
MOH	ministry of health
MR	mortality rate
MSF	Médecins sans Frontières
MSM	men who have sex with men
MT	medical transcription
MUAC	mid-upper arm circumference
NAFTA	North American Free Trade Agreement
NAM	National Academy of Medicine (U.S.)
NATO	North Atlantic Treaty Organization
NCCIH	National Center for Complementary and Integrative Health (U.S.)
NCD	noncommunicable disease
NEP	National Evaluation Platform
NGO	nongovernmental organization
NHC	neighborhood health committee
NHIF	National Health Insurance Fund (Tanzania)
NHS	National Health Service (U.K.)
NICE	National Institute for Clinical Excellence (U.K.)
NID	National Immunization Day
NIH	National Institutes of Health (U.S.)
NIMH	National Institute for Mental Health (U.S.)
NMCP	National Malaria Control Program (Tanzania)
NME	new molecular entity
NMR	neonatal mortality rates
NMRA	national medicines regulatory authority
NPHW	non-physician healthcare worker
NTD	neglected tropical disease
OCHA	Office for the Coordination of Humanitarian Affairs (UN)
ODA	official development assistance
OECD	Organization for Economic Cooperation and Development
OOP	out of pocket
OOS	out of specification
OP	outpatient
OPV	oral polio vaccine
OR	odds ratio
ORS	oral rehydration solution
ORT	oral rehydration therapy
P4P	pay for performance
PAF	population-attributable factor
PAHO	Pan American Health Organization
PATH	Program for Appropriate Technology in Health
PCBs	polychlorinated biphenyls
PCIS	patient care information system
PCR	polymerase chain reaction
PDP	productive development partnership
PEF	Pandemic Emergency Financing Facility (World Bank)
PEM	protein-energy malnutrition
PEN	Package of Essential Medicines and Technologies for Non-Communicable Disease Interventions
PEP	post-exposure prophylaxis (for HIV)
PEPFAR	President's Emergency Plan for AIDS Relief
PEV	Expanded Program on Immunization (French)
PHC	primary health care
PHE	Public Health England
PHEIC	public health emergency of international concern
PHFI	Public Health Foundation of India
PHR	Physicians for Human Rights
PIEA	political and institutional economy analysis
PLHIV	people living with HIV
PLWHA	people living with HIV and AIDS
PMNCH	Partnership for Maternal, Newborn, and Child Health
POC	point of care
POLHN	Pacific Open Learning Health Net
PPI	Private Participation in Infrastructure
PPM	public–private mix
PPP	purchasing power parity; public–private partnership
PPTC	Pacific Paramedical Training Centre
PRA	participatory reflection and action; probabilistic risk assessment
PrEP	pre-exposure prophylaxis (for HIV)

PRIME	Program for Improving Mental Health Care
PRSP	Poverty Reduction Strategy Paper
PSAC	Policy Strategy Advisory Committee
PSR	pressure–state–response
PTSD	post-traumatic stress disorder
QA	quality assurance
QALY	quality-adjusted life year
QC	quality control
QWB	quality of well-being
RAE	retinol activity equivalent
RAP	Rapid Assessment Procedure
RARE	Rapid Assessment, Response, and Evaluation
RBM	Roll Back Malaria Partnership
RBP	retinol binding protein
RCT	randomized controlled trial
R&D	research and development
RDA	Recommended Dietary Allowance
RDoC	Research Domain Criteria (U.S.)
REDISSE	Regional Disease Surveillance Systems Enhancement (African program)
REQA	Regional External Quality Assessment
RHU	Refugee Health Unit (Somalia)
RR	relative risk
RSV	respiratory syncytial virus
RTA	regional trade agreement
RUTF	ready-to-use therapeutic food
RWJF	Robert Wood Johnson Foundation
SAARC	South Asian Association for Regional Cooperation
SAM	severe acute malnutrition
SARS	severe acute respiratory syndrome
SD	standard deviation
SDGs	Sustainable Development Goals
SDH	social determinants of health
SDHE	social determinants of health equity
SEA-EU-NET	Southeast Asia–Europe Joint Funding Scheme for Research and Innovation
SFP	supplementary feeding program
SGA	small for gestational age
SIP	sector investment program
SMART	specific, measurable, attainable, relevant, and time bound (objectives)
SPS	Agreement on the Application of Sanitary and Phytosanitary Measures (WTO)
SSB	sugar-sweetened beverage
STEM	science/technology/engineering/mathematics
STEPS	STEPwise approach to Surveillance
STI	sexually transmitted infection
SWAp	sector-wide approach
SWOT	strengths, weaknesses, opportunities, and threats
TAC	Treatment Action Campaign
TB	tuberculosis
TBT	Agreement on Technical Barriers to Trade (WTO)
TCM	traditional Chinese medicine
TDR	WHO-based Special Programme for Research and Training in Tropical Diseases
TEA	total exposure assessment
TEHIP	Tanzania Essential Health Interventions Project
TFC	transnational food company
TFP	therapeutic feeding program
TFR	total fertility rate
THET	Tropical Health and Education Trust
TIA	trade and investment agreement
TNC	transnational corporation
TPP	Trans-Pacific Partnership
TRAP	traffic-related air pollution
TRIPS	Trade-Related Intellectual Property Rights (WTO)
TRL	Transport Research Laboratory
TSH	thyroid-stimulating hormone
TTIP	Transatlantic Trade and Investment Partnership
UCT	unconditional cash transfer
UDHR	Universal Declaration of Human Rights
UHC	universal health coverage
UI	urinary iodine
UK PHRST	U.K. Public Health Rapid Support Team
UN	United Nations
UNAIDS	United Nations Programme on HIV/AIDS
UNAMID	United Nations–African Union Mission in Darfur
UNCED	United Nations Conference on Environment and Development
UNCHR	United Nations Commission on Human Rights

UNCPRD	United Nations' Convention on the Rights of Persons with Disability
UNCTAD	United Nations Conference on Trade and Development
UNDP	United Nations Development Program
UNEP	United Nations Environment Program
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNFPA	United Nations Population Fund
UNGC	United Nations Global Compact
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund
UNISDR	United Nations Office for Disaster Risk Reduction
USAID	U.S. Agency for International Development
USP	*United States Pharmacopoeia*
UV	ultraviolet
UVR	ultraviolet radiation
VA	verbal autopsy
VAT	value-added tax
VVM	vaccine vial monitor
WASH	water, sanitation, and hygiene
WAZ	weight-for-age *z*-score
WCH	women's and children's health
WEIRD	Western, educated, industrialized, rich, and democratic
WFP	World Food Program
WHA	World Health Assembly
WHO	World Health Organization
WHZ	weight-for-height *z*-score
WIPO	World Intellectual Property Organization
WLZ	Weight-for-length *z*-score
WPRO	Western Pacific Regional Office (WHO)
WTO	World Trade Organization
XDR-TB	extremely drug-resistant tuberculosis
YLD	years of life lived with disability
YLL	years of life lost
YP-CDN	Young Professionals Chronic Disease Network



Index

Note: *e* denotes exhibits; *f* denotes figures; *n* denotes footnotes; *t* denotes tables.

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y

Z