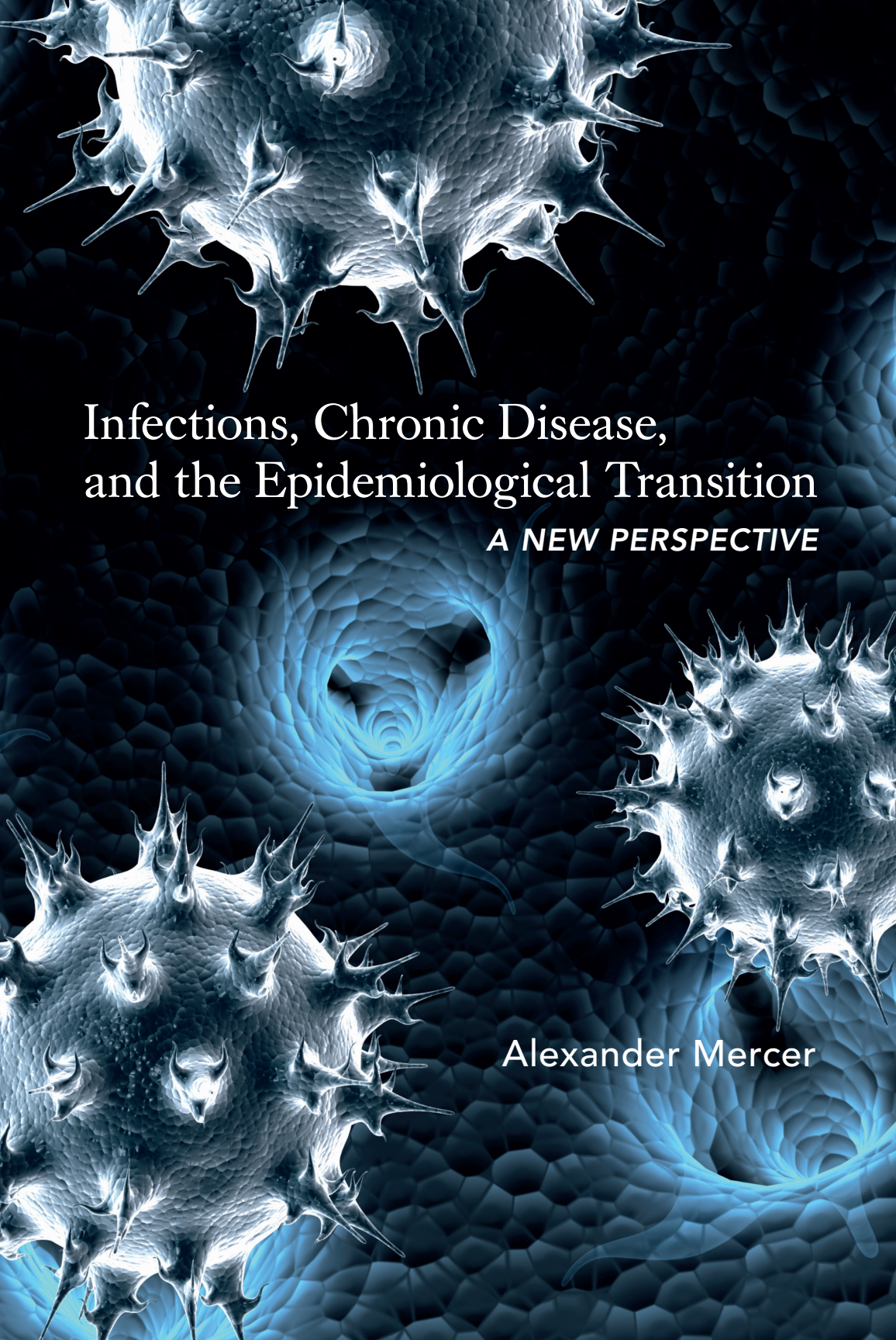

Infections, Chronic Disease, and the Epidemiological Transition

A NEW PERSPECTIVE

Alexander Mercer

Infections, Chronic Disease, and the Epidemiological Transition

Rochester Studies in Medical History

Senior Editor: Theodore M. Brown
Professor of History and Preventive Medicine
University of Rochester

Additional Titles of Interest

The Birth Control Clinic in a Marketplace World
Rose Holz

Bacteriology in British India: Laboratory Medicine and the Tropics
Pratik Chakrabarti

Barefoot Doctors and Western Medicine in China
Xiaoping Fang

Beriberi in Modern Japan: The Making of a National Disease
Alexander R. Bay

The Lobotomy Letters
Mical Raz

Plague and Public Health in Early Modern Seville
Kristy Wilson Bowers

Medicine and the Workhouse
Edited by Jonathan Reinarz and Leonard Schwarz

Stress, Shock, and Adaptation in the Twentieth Century
Edited by David Cantor and Edmund Ramsden

Female Circumcision and Clitoridectomy in the United States: A History of a Medical Treatment
Sarah B. Rodriguez

The Spanish Influenza Pandemic of 1918–1919: Perspectives from the Iberian Peninsula and the Americas
Edited by María-Isabel Porras-Gallo and Ryan A. Davis

A complete list of titles in the Rochester Studies in Medical History series may be found on our website, www.urpress.com.

Infections, Chronic Disease, and the Epidemiological Transition

A New Perspective

Alexander Mercer

University of Rochester Press

First published 2014

University of Rochester Press
668 Mt. Hope Avenue, Rochester, NY 14620, USA
www.urpress.com
and Boydell & Brewer Limited
PO Box 9, Woodbridge, Suffolk IP12 3DF, UK
www.boydellandbrewer.com

ISBN-13: 978-1-58046-508-3
ISSN: 1526-2715

Library of Congress Cataloging-in-Publication Data

Mercer, Alex, author.
Infections, chronic disease, and the epidemiological transition : a new perspective / Alexander Mercer.
p. ; cm. — (Rochester studies in medical history, ISSN 1526-2715 ; vol. 31)
Includes bibliographical references and index.
ISBN 978-1-58046-508-3 (hardcover : alk. paper)
I. Title. II. Series: Rochester studies in medical history. 1526-2715
[DNLM: 1. Communicable Diseases—epidemiology. 2. Disease Transmission, Infectious. 3. Mortality—trends. WA 110]
RA643
616.9—dc23

2014027696

A catalogue record for this title is available from the British Library.

This publication is printed on acid-free paper.
Printed in the United States of America.

For my parents

Contents

	Preface	ix
	Acknowledgments	xiii
	Introduction	1
1	Background	8
2	Theoretical Framework, Data, and Study Outline: The Concept of Epidemiological Transition	19
3	A New Infectious Disease Environment	34
4	Mortality Decline, Food, and Population Growth: "Standard of Living" and Nutrition	48
5	Smallpox	58
6	Typhus, Typhoid, Cholera, Diarrhea, and Dysentery	77
7	Infant Mortality	90
8	Child Mortality	101
9	Tuberculosis	118
10	Respiratory Diseases	131
11	Cardiovascular Disease	141
12	Cancer	163
13	Other Chronic Diseases	181
14	Epidemiological Transition: A New Perspective	203
	Appendixes	223
	Notes	241
	Bibliography	277
	Index	321

Preface

Much has been written since the 1970s about the causes of the decline in mortality from infectious diseases, but far less about the emergence of chronic diseases. When Abdel Omran first referred to the change in predominant causes of death as an "epidemiologic transition," academic studies tended to focus more on issues raised by Thomas McKeown in his influential book *The Modern Rise of Population.* McKeown's emphasis on improvements in the "standard of living" and nutrition as key determinants of increased life expectancy aroused considerable controversy, while Omran's "theory" of transition remained a basically descriptive framework for analysis of changes in disease and mortality. The epidemiological transition is ongoing globally, and constitutes a fundamental change in the human condition; as such, the phenomenon has both contemporary relevance and historical importance.

One objective of this book is to review evidence from local, national, and international studies indicating that preventive measures and changes in behavior played a far greater role in mortality decline in England and globally than McKeown or Omran suggested. However, the aim is to go beyond the controversy over the relative importance of "human intervention" and "standard of living" as determinants of increased life expectancy. Assessment of changes in disease and mortality in England since the 18th century and their possible causes is fundamental to an understanding of how contemporary patterns of disease emerged. A very long-term perspective is particularly important as it encompasses the development of a preventive approach to disease in response to devastating epidemics of plague, smallpox, and typhus that laid the foundation for later health interventions. Records of cause-specific mortality available for England, with supporting data from other Western countries, form the basis of this review of the epidemiological transition and illustrate its main features—acute infectious diseases being superseded by chronic diseases as the predominant causes of death, a shift in age at death from childhood to older adult ages, increased life expectancy, and a larger, "aging" population.

A central theme of the book is the interrelated nature of diseases, particularly the role of microbial infection in chronic diseases, which indicates there is another dimension to epidemiological transition. The original concept oversimplified the process, viewing it as the replacement of one group of distinct diseases (infectious) by another ("degenerative" or "noncommunicable"). This fails

to take into account the many interrelationships between infectious diseases, between infections and chronic diseases, and between the disorders underlying chronic diseases. For a more complete understanding of how chronic diseases came to predominate, it is essential to consider these relationships and the consequences of microbial infections that people encounter over the course of their lives. The need for a new conceptualization of epidemiological transition is indicated by statistical evidence of association between diseases and the findings of many clinical, observational, and epidemiological studies linking specific infections with chronic diseases. Common microorganisms, acute and chronic infections, related immune responses, and inflammatory processes are increasingly being recognized as contributing to the multifactorial etiology of chronic disease. The evidence considered here relates to various cancers, heart disease, stroke, peptic ulcer, diabetes, diseases of the liver and kidneys, and many other common chronic disorders with suspected infectious etiology.

The epidemiological-demographic perspective of this study complements an evolutionary-ecological perspective on contemporary disease patterns developed by microbiologists. Humans did not evolve for "modern lifestyles," but were affected by exposure to new microorganisms and infectious diseases as their way of life changed in earlier times. It is important to view the emergence of chronic diseases in the historical context of dramatic changes in the infectious disease environment. With urbanization and industrialization in the modern era, increasing exposure to environmental and lifestyle-related risk factors has exacerbated the biophysiological effects of infection, contributing to the development of chronic disorders. A consideration of some of the evidence of microbial involvement in chronic diseases in relation to historical trends in cause-specific mortality can contribute to an understanding of how contemporary patterns of disease and mortality have come about.

In the last two decades, there has been a growing awareness among epidemiologists that a multidisciplinary approach is required for a more complete understanding of variation and change in disease and mortality patterns. In the past, historical research tended to be fragmented, with studies by demographers, historians, geographers, economists, and medical specialists focusing on different aspects and phases of change, particular diseases, and issues of interest within disciplinary boundaries. For instance, my earlier monograph, *Disease, Mortality and Population in Transition*, focused on infectious disease mortality decline in England in the 18th and 19th centuries; selections from that monograph have been modified for inclusion here. Studies of mortality decline have generally not compared trends for infectious diseases and chronic diseases, while the epidemiological transition has not been comprehensively reviewed using data from one particular country. This book aims to address these issues and highlight a new perspective on the epidemiological transition reflecting the interrelated nature of diseases, particularly links between infection and chronic disease. Infectious etiology has now been established for a sufficient number of chronic diseases, accounting for a significant proportion of all morbidity and mortality, to indicate the need for a new conceptualization of epidemiological

transition. This view would focus more on etiology and pathogenesis than previous adaptations of the theory, which have focused on phases of transition, and would reflect the increasingly blurred distinction between infectious diseases and chronic diseases. The new perspective is consistent with the view of epidemiologists who suggest that future research should be guided by a new disease paradigm encompassing pathogenesis as well as risk factors for disease linked with lifestyle and socioeconomic circumstances. As the global phenomenon of epidemiological transition continues, and the growing burden of chronic disease presents major challenges for public health and health care services worldwide, the topic has much contemporary relevance to the "health transition" in both low-income and high-income countries. The final chapter highlights possible implications for public health policy, particularly the importance of prioritizing disease prevention, which arguably made the greatest contribution in the past to improvements in longevity, health, and the quality of life.

Acknowledgments

I am grateful to friends who have listened to and queried ideas that I hope are now more clearly articulated here. I would especially like to thank Theodore Brown and the University of Rochester Press for the opportunity to contribute to this series, and anonymous publisher's reviewers whose comments and suggestions were both helpful and challenging. I also greatly appreciate the unstinting support of the editorial team that has made publication possible, and the many suggestions for improvement of the manuscript during copyediting.

Introduction

The decline in mortality from infectious diseases has been one of the most significant changes in human history leading to unprecedented increases in life expectancy. There is an expectation in industrialized countries that almost all children will grow up to adulthood and that most people will live at least eighty years despite the high incidence of chronic diseases. Heart disease, stroke, cancers, diabetes mellitus, ulcers, diseases of the kidneys and liver, and other chronic diseases cause much debility, long-term sickness, and "premature" death, which to some extent detracts from the sense of progress in health that has come with the control of infectious diseases. The epidemic diseases that caused high mortality in the 18th and 19th centuries generally resulted in acute short-term sickness, which was often fatal, particularly in the early years of life, while the major chronic infectious disease, tuberculosis, caused high mortality among adults. Infectious diseases restricted average life expectancy to 30–40 years in England and other European countries. A change in the predominant causes of morbidity and mortality from acute infectious diseases and tuberculosis to other chronic diseases among adults, concomitant with increasing life expectancy and a higher proportion of older people in the population, characterizes the epidemiological transition. This term distinguishes the process from Abdel Omran's original concept, "epidemiologic transition," which included phases and models of change.[1] Inadequate definition of disease categories, lack of consideration of disease etiology, and other aspects of Omran's theory have been criticized, and many refinements have been made. Nevertheless, the basic concept of epidemiological transition is widely accepted as a framework for analysis of long-term changes in disease and mortality.

A major objective of this study is to review the emergence of chronic diseases, the now predominant causes of death in industrialized countries, in the context of dramatic changes in the infectious disease environment. The study focuses on the epidemiological transition that occurred in England from the 18th century onward. This is distinct from an earlier change in the predominant infectious diseases when humans first began to live in settled communities during the "Neolithic Revolution," and a further phase of the global transition in which new infectious diseases have emerged and old ones have reemerged as a threat in recent decades, mostly as a result of human action.[2] The new perspective in Omran's "epidemiologic transition" was his focus on a shift from one group

of causes of death, infectious diseases, to another group of causes considered to be distinctly different, "degenerative and man-made diseases." Omran later clarified that he included mental illness, drug dependency, radiation injury, occupational hazards, and traffic accidents.[3] Subsequently, the term "noncommunicable" has been widely used for chronic diseases and external causes of death (accidents, suicide, and violence). Greater knowledge about etiology made the term "degenerative disease" inappropriate for many chronic diseases, and use of the label "noncommunicable" is in many cases based on an assumption about noninfectious etiology.

This study takes into account interrelationships between infections and chronic diseases, and between the underlying disorders and biophysiological processes, which indicate that a concept of epidemiological transition based on a simple dichotomy of disease categories is too simplistic. A growing body of evidence indicates that microorganisms are associated with many chronic disorders previously regarded as having no infectious etiology and being noncommunicable, and a causal role has been confirmed in some cases. Given that some chronic diseases are known to have infectious etiology, and at least some microbial involvement is suspected in many others, the concept of a shift in the predominant causes of morbidity and mortality should take into account the increasingly blurred distinction between infectious and chronic disease. The possibility of infectious etiology being confirmed for more chronic diseases in the future suggests that the epidemiological transition might come to be viewed in terms of the balance between microbial and other causes of disease. Categories of disease that allow for what is known and not known about disease etiology are outlined in chapter 2, which provides a theoretical framework for this study of changes over time in disease and mortality.

Future research into the relationship between infectious disease and chronic disorders will contribute to a greater understanding of the ongoing global process of epidemiological transition. Increasing research interest in this field does not signify a tendency toward microbiological reductionism, as evidence that microorganisms, infection, and immune responses have a causal role in the pathogenesis of some chronic diseases complements evidence implicating various risk factors. Multifactorial causation of chronic diseases is now an accepted concept, encompassing biophysiological, lifestyle, behavioral, social, cultural, economic, and environmental factors. The mechanisms and pathways of causation for different diseases are often interconnected, and interrelationships and synergies between the underlying disorders can contribute to an increased risk of dying "prematurely." Interrelationships and synergies also existed between the infectious diseases that were major killers in the past, and these exacerbated the severity of sickness and increased the risk of dying. Sociodemographic factors also affect the severity of acute infectious diseases that contributed to high child mortality, and beneficial changes occurred toward the end of the 19th century. Early in the 20th century, it was recognized that more children were surviving to reach adulthood, although lifetime effects of childhood infections were suspected when age-specific adult death rates were found to vary in relation to

year of birth.[4] Subsequently, many studies have found more direct evidence linking particular chronic diseases with acute infectious diseases in childhood, chronic infections, and infections contracted in adult life.

A life course perspective on the development of chronic diseases complements a historical perspective on trends over time.[5] In addition, an evolutionary-ecological perspective, using the concept of epidemiological transition as a framework, also contributes to a more comprehensive understanding of the changing relationship between disease patterns, the environment, and the consequences of human action.[6] Long before many of the well-established risk factors for chronic disease became prevalent, humans were exposed to infections and developed immune responses that can affect health throughout life. Even though many of the associations between microorganisms and chronic diseases are still under investigation and criteria for establishing causality can be difficult to apply, it is conceivable that rapid expansion of this area of research and advances in technologies for identifying pathogens could lead to a radical change in the conceptualization of disease. It has even been suggested that as more causal links between infections and chronic disease are confirmed, a "new germ theory" of disease could emerge.[7] A new epidemiological paradigm for research might encompass pathogenesis and the many levels of causation, including genetic, immunological, biophysiological, sociocultural, and economic, as well as risk factors linked with lifestyle, behavior, and the environment.[8] These aspects of disease etiology need to be taken into account when considering the emergence of the chronic diseases that have replaced acute infectious diseases and tuberculosis as the predominant causes of morbidity and mortality.

The change in predominant diseases, the shift from child to adult mortality, and the increase in life expectancy in the epidemiological transition are integral to the process of demographic transition, which encompasses changes in mortality, fertility, and population. Although different versions of demographic transition theory have been proposed since the 1940s, the basic concept is widely accepted. Mortality decline has usually preceded fertility decline, leading to more rapid population growth.[9] Societies that have undergone a process of "modernization" since the 18th century have experienced a transition from high fertility and high mortality to low fertility and low mortality. It has been questioned whether the concept of demographic transition constitutes a theory, a historical or predictive model, a generalization, a mere description, or just a useful framework for analysis of data.[10] Similar issues surround the concept of epidemiological transition, although it has focused attention on important global changes in disease and mortality. The relationship between declining mortality and fertility is complex, with improved child survival potentially having a negative impact on fertility because conception is less likely during breastfeeding, while prevention of debilitating infectious disease among young adults reduces a constraint on fertility. Of more concern here is the consequence of longer survival and lower fertility in terms of "population aging," an increasing proportion of older people that has implications for morbidity and the provision of health

and social services. In this respect, a greater understanding of disease etiology and epidemiological transition is integral to the broader process of health transition. The term was originally used to distinguish cultural, social, and behavioral determinants of health from the effects of medical interventions and changes in people's economic standard of living. Health transition theory developed in the 1990s, encompassing health, morbidity, and the organized health-care response, as well as changes in mortality.[11]

Most studies of historical change in disease patterns have focused on mortality because of the availability of quantitative data, which are generally lacking for morbidity. The studies by Riley of sickness data for workingmen in England from the 19th century and studies by Fogel and Costa of records for Union Army veterans in the United States are notable exceptions. Fogel reported that a high proportion (13–22%) of males aged 16–39 years were rejected by the army in 1861 as they were found to have chronic health conditions.[12] Among those recruited, chronic diseases were more likely to develop later in life following infectious disease while in the army: acute respiratory infections followed by chronic respiratory problems and arrhythmias; measles by chronic respiratory problems and valvular heart disease; typhoid fever by valvular heart disease; rheumatic fever by heart disease, arrhythmias, congestive heart failure, and joint and back problems; diarrheal disease by arteriosclerosis; and malaria followed by joint problems.[13] Fogel and Costa analyzed data on height, body mass index, and morbidity, which indicated that Union Army veterans aged 65 years and over in 1910 were much more likely to have heart, musculoskeletal, respiratory, and digestive diseases than veterans aged 65 and over in a national survey in 1985–88. They attributed the difference to early life experience as well as health interventions and changes in lifestyle, as severity of infectious disease, undernourishment, and other types of stress in utero and in childhood may adversely affect the development of vital organs.[14] Fogel found that the onset of chronic disease had been delayed over the course of the 20th century. Compared with white males in 1895–1910, white males in 1983–92 were much more likely to be free of chronic diseases at ages 60–64 years, while arthritis, heart diseases, respiratory diseases, and cancer started 8–11 years later among men aged 65 years and over in 1983–92. He concluded that improvements in health in the second half of the 20th century were partly due to improved physiology among later cohorts who had benefited from improved technologies for food production, public health, personal hygiene, and medical interventions introduced earlier in the century.[15]

Riley found that when case fatality and mortality declined in England at the end of the 19th century, the average length of episodes of sickness among a sample of workingmen increased. This could be due to episodes of sickness being relatively short in the past because they either were not severe or quickly resulted in death, and to survival of more people predisposed to chronic sickness later in life by infectious diseases in childhood.[16] The aggregate sickness time has now become critical for health-care provision, and the challenge for health policy and programs is to promote lifestyles and working conditions conducive to

longer lives spent in good health.[17] Riley also found that the profile of sickness among workingmen was substantially different from the cause-of-death profile, with only respiratory disease and accidents featuring significantly in both. Major chronic diseases causing deaths in England at the end of the 19th century such as heart disease, cancer, and nephritis were not among the causes of sickness recorded at that time.[18] Despite the limitations of mortality data, they provide a more complete record of the emergence and growing predominance of chronic diseases. This study is based on the wealth of local and national historical mortality data available for England and Wales, with data from other Western countries used for comparison on particular aspects of disease and mortality change.

It is now widely accepted that the long-term decline in mortality in England began in the mid-18th century, with subsequent downturns at the beginning of the 19th century and in the 1870s, as in much of Europe.[19] Within this overall mortality transition, there was a resurgence in the national death rate between the 1830s and 1870s when the population of industrial towns was increasing rapidly. Many people who migrated to the towns to escape rural poverty and seek new livelihoods experienced squalid living conditions, poverty, and disease in densely populated "slum" areas. Death rates were extremely high, particularly among children, with up to half of them dying before the age of 5 years. With so many very short lives, average life expectancy was only about 30 years in major industrial cities such as Manchester and Liverpool.[20] From the 1870s, the decline in the national death rate resumed and became more rapid and continuous from the beginning of the 20th century. A major study of aggregated data from several hundred parishes in England by Wrigley and Schofield provided stronger evidence of the earlier phase of mortality decline that occurred before national registration. This confirmed the major contribution to more rapid population growth made by mortality decline, and suggested that the level of fertility made a greater contribution than previously thought.[21] The estimates for England as a whole indicated a decline in the death rate from the 1740s, and an increase in the birthrate throughout the 18th century, contrary to earlier indications from small-scale studies.[22] The new national estimates led to a revision of the theory of demographic transition regarding the relative contribution of changes in mortality and fertility to population growth.

Demographic transition and epidemiological transition are closely related processes in an ongoing global phenomenon. The lag between mortality decline and fertility decline has generated unprecedented population growth. Although demographic balance has been restored in industrialized countries, high fertility rates persist in many African and Middle Eastern countries, despite considerable improvements in child survival. Population growth rates still exceed the peak rates experienced in England in the 19th century, and economic opportunities are grossly inadequate to meet the needs of a growing number of young adults. Cultural, social, economic, and political factors inhibit the process of fertility transition, and rapid population growth is projected to continue for many years. In contrast, remarkable reductions in fertility have been achieved in many low-income countries in Africa and Asia as a result of reproductive health and family

planning programs, enrollment of girls in school, postponement of marriage, employment opportunities for women, and their empowerment to make decisions about childbearing and contraception. Migration from rural areas into the towns has exposed people to new ideas and values, including limitation of family size. At the same time, migration into the towns has led to rapid growth in urban population size and density, with adverse consequences for living conditions and health. As in Europe in the past, people in the poorer areas of expanding towns and cities in developing countries are exposed to infectious diseases that can transmit more rapidly in crowded conditions. The infectious diseases include many of those that caused high mortality in England in the past, such as diarrheal diseases, respiratory diseases, tuberculosis, and childhood epidemic diseases, as well as warm-climate diseases and previously unknown infectious diseases that have emerged recently.

The main focus in this study is the transition in cause of death in England from acute infectious diseases and tuberculosis to other chronic diseases, the decline in child mortality, increased survival to older ages at which chronic diseases are more likely to occur, longer lives, and a shift toward predominantly adult deaths. A primary objective is to reassess the major changes in disease and mortality since the 17th century and possible explanations. Much of the decline in mortality between the mid-18th and mid-19th century involved acute infectious diseases, although a major component after this was the declining death rate from tuberculosis. While there are acute forms,[23] it is typically regarded as a chronic infectious disease as most clinical cases are due to reactivation of a latent infection or reinfection among those already tuberculin sensitive from an earlier infection. Other chronic diseases among adults became increasingly predominant as mortality from acute infectious diseases and tuberculosis declined.

The historical change in disease and mortality patterns experienced in England during rapid urbanization has parallels in developing countries today. Many of the infectious diseases that caused high mortality in England in the past still contribute to unacceptably high child mortality in low-income countries. In addition, diseases prevalent in warm climates, particularly malaria, continue to cause much sickness and high mortality in both urban and rural areas. Infectious diseases continue to be a major cause of death, accounting directly for about one-quarter of deaths globally, and contributing to chronic conditions and complications that account for many more deaths.[24] The pandemic of HIV/AIDS has led to reductions in average life expectancy in some parts of Africa, although the disease has caused relatively few deaths in England.[25] Elsewhere, child death rates and life expectancy have generally improved, but death rates from chronic diseases have reached levels found in industrialized countries. Although the composition of the disease burden in developing countries is different from that in England in the past, and new infectious diseases are emerging and old ones reemerging, a reassessment of this transition for which there is particularly well-documented evidence can contribute to a better understanding of the ongoing process of global epidemiological transition.

Historical changes in disease and mortality in England can be seen as broadly typical of the European experience of epidemiological transition, and the main diseases recorded in the 18th century that contributed to mortality decline were similar throughout the Continent. There is wide consensus that long-term mortality decline began first in England and northwestern Europe, with an early phase characterized by diminished spikes in mortality caused by epidemic diseases.[26] Although there were different levels of mortality and different turning points in disease-specific trends, there were many common features, such as the disappearance of plague by the early 18th century, the decline of typhus, and the control of smallpox by the end of the 19th century. As tuberculosis mortality among adults declined, the other chronic diseases that came to predominate as causes of death in England were the same as those in other European countries, although the relative contribution of particular diseases to overall mortality varies. Before considering the concept of epidemiological transition further and outlining the theoretical framework for this study, it is useful to review some of the issues addressed in studies of infectious disease mortality decline in England and the more recent mortality transition in developing countries.

Chapter One

Background

The decline in mortality from the mid-18th century in England and other European countries and the more rapid growth in population that occurred up to the mid-20th century have been the subject of investigation and debate among historians, geographers, economists, demographers, and medical specialists. Attention has focused on the relative importance of different factors, such as treatment, care of the sick, measures to control and prevent infectious diseases, living conditions, "standard of living," nutrition, and changes in the nature of some diseases. Historians and economists have considered the decline in mortality in the context of agricultural and industrial revolutions leading to improvements in food supplies and expansion of economic activity. Studies of the relationship between food supply and death rates in the period leading up to the mortality transition have indicated that many epidemic diseases caused spikes in mortality that were independent of the effects of famine and poor harvests.[1] In earlier premodern times, changes in the infectious disease environment were linked with great transformations in the way of life of humans, particularly the development of agriculture and adaptation to living in settled communities.[2] Eventually urbanization and large-scale migrations led to new ways of living in very different environments, while trade and changes in population size and density brought exposure to new pathogens, different infectious diseases, and changes in the immunological status of individuals and populations.

Industrial development and economic growth in Western countries eventually brought many improvements in living conditions, but the concept of "standard of living" is too general to provide a satisfactory explanation for the decline in mortality and increased life expectancy. It reflects the values of particular sections of society, cultures, and historical periods and masks factors specific to the decline in mortality from particular diseases. It is clearly important to distinguish the specific changes in people's lives that contributed to improved survival chances. In the limited sense of people's own economic circumstances and access to food, shelter, and other basic needs, the concept of "standard of living" can generate hypotheses that are testable with reference to specific indicators. However, generalizations have often been formulated that persisted as explanations of mortality decline despite the lack of supporting evidence. McKeown in particular was an influential exponent of the view that improvements in the "standard of living" and nutrition were the main cause of mortality decline in

England in the second half of the 19th century. In his early work, he and his colleagues concluded that about half of the decline in the overall death rate was the result of improvements in the "standard of living," particularly the effect of a putative improvement in nutrition on tuberculosis mortality.[3] He attributed a further quarter of the decline to sanitary reforms and hygiene measures affecting waterborne diseases and the remainder of the decline to a suggested change in the "character" of scarlet fever. A similar explanation was given by Omran in a study of historical data from New York, in which he referred to mortality decline and the change in the predominant causes of death as an "epidemiologic transition." He concluded that mortality decline was due to improvements in nutrition, "living standards," housing, personal hygiene, and a change in the "character" of scarlet fever.[4] More recently, Fogel analyzed anthropometric data and found that much of the decline in mortality in England between 1751 and 1975 was associated with changes in physiology. He interpreted these as an indication that levels of malnutrition had been reduced, although he recognized that malnutrition was also a consequence of high infectious disease rates.[5]

McKeown emphasized the role of nutrition in the decline of mortality from airborne infectious diseases, concluding that "the most acceptable explanation of the large reductions of mortality and growth of population which preceded advances in hygiene is an improvement in nutrition due to greater food supplies."[6] This view has been widely criticized, partly because of a lack of direct evidence of an improvement in diet, or in the economic standard of living of most people between the mid-18th and mid-19th century. Szreter pointed out that McKeown's thesis came to be accepted by many professionals concerned with public health, despite being largely speculation for which no supporting evidence was provided.[7] In fact, McKeown acknowledged reaching his conclusion that nutrition must have played a part because nutritional deprivation is known to contribute to the risk of dying from infectious diseases in developing countries, and he could offer no other explanation for the decline in airborne infectious disease death rates in England.[8]

Many studies have investigated the role of nutrition in resistance to disease and its effect on the efficiency of immune responses to viral infections.[9] Inadequate nutrition is clearly an important contributing factor in many child deaths from infectious disease in developing countries, and undernourishment undoubtedly contributed to high mortality in Europe in the past. However, there is a complex interrelationship between nutrition, infectious disease, and mortality risk, with severe infectious diseases causing malnutrition, interfering with the body's ability to convert food into energy and growth, and stunting physiological and intellectual development.[10] Evidence of improvements in physiology does not necessarily indicate improved intake of nutrients (diet), and nutritional status is a measure of net nutrition after the nutrient requirements of the body have been met.[11] While agricultural development and imports improved the supply of food between the mid-18th and mid-19th century, supplies may only have kept pace with population growth, while widespread poverty greatly restricted access to food. Studies of food consumption per capita and real

incomes in England in the 18th century have not indicated a consistent association with mortality decline, and the economic standard of living of many people may have deteriorated in the first quarter of the 19th century when mortality was lower and population growth more rapid. There is unlikely to have been much improvement in the "standard of living" of the new urban labor force in the period of rapid industrialization between the 1830s and 1870s, although the economic circumstances of many occupational groups probably improved in the last three decades of the century.[12]

As cities grew during the industrial revolution, health hazards increased with industrial pollution, particles from wood and coal fires in the air, poor disposal of animal and human waste, and streets contaminated with standing water, rubbish, and excrement. Many risk factors for chronic diseases became more prevalent and employment in hazardous occupations increased, contributing to longer episodes of sickness due to accidents and respiratory disease.[13] Exposure to pollution and hazardous substances in the air and food would have contributed cumulatively to chronic sickness, while severe infectious diseases in childhood contributed to chronic disease later in life (see chapters 9–12). Szreter emphasized the severe negative consequences for health of conditions in industrial towns that constituted a threat to the whole population and presented major public health challenges.[14] Material living standards were critically important for health, as they are in developing countries undergoing rapid urbanization today. Szreter suggested that the rapid industrial and economic growth in Britain between the 1830s and 1870s had socially divisive effects, including high rents for in-migrants from rural areas, residential segregation in the growing towns, and increasing economic inequality. He argued that addressing the problems in the increasingly crowded, poorer areas of industrial towns in England in the 19th century required political commitment, organization, and the application of resources;[15] and economic growth was only converted into improvements in health through political, social, and cultural influence. Municipal authorities in industrial towns eventually implemented appropriate measures in response to the public health problems and high mortality, although not without opposition from those who were expected to contribute financially; meanwhile, the wealthiest were able to move to more salubrious areas.

Although McKeown recognized the contribution made by improvements in sanitation, he made no reference to the political commitment and organization required to achieve reforms. Szreter suggested that his unsubstantiated speculation that an improvement in people's "standard of living" and nutrition was the main reason for the decline in infectious disease mortality was widely accepted by health professionals as if it were evidence based. Furthermore, it unwittingly lent support to advocates of noninterventionism and reduced public investment in health, education, and human development in the 1980s and 1990s.[16] Easterlin argued that an assumption that economic growth based on a free market can solve problems of health and development is belied by the historical experience of increasing life expectancy linked with the control of infectious diseases and public-sector intervention.[17] Fogel, in developing a theory of technological,

physiological, and economic change, emphasized that human beings have gained an unprecedented degree of control over their environment in the last 300 years, which has led to a dramatic increase of 50% in average body size and over 100% in average life expectancy.[18] He recognized that this "technophysio evolution" was not necessarily stable, and that both height and life expectancy had declined in some periods of economic growth in the past.[19] Like McKeown, he emphasized the association between chronic malnutrition and mortality in the 18th and 19th centuries. Much of his research focused on agricultural development, food consumption, and human physiology. He found a strong statistical association between mortality and changes in human height and body mass index, concluding that the decline in mortality could in large part be explained in terms of a reduction in chronic malnutrition.[20] Fogel emphasized that these anthropometric measures are indicators of net nutrition, the balance between nutrient intake (diet) and the body's nutrient requirements.[21] He recognized that infectious diseases have adverse effects on nutritional status, physiological development, and the risk of chronic disease later in life;[22] however, his explanation of long-term mortality decline did not take into account the control that human beings gained over infectious diseases.

The examination here of trends in death rates from infectious diseases and efforts to control them suggests that some studies have greatly underestimated the role in mortality decline of preventive measures, public health interventions, social organization, and sociodemographic changes. Smallpox in particular would not have been eliminated without widespread use of vaccination later supported by other preventive measures. The recession of major epidemic diseases such as typhus, typhoid, and cholera in the 19th century is unlikely to have occurred without preventive measures and public health interventions. Rational responses to the threat of feared epidemic diseases had developed over centuries, along with systems for monitoring the spread of disease based on the collection of mortality statistics. This provided a foundation for further improvements when the germ theory was accepted and knowledge of transmission developed at the end of the 19th century. The evidence reviewed here indicates continuity in the key role of preventive and public health measures in bringing changes in disease and mortality globally. Studies in both industrialized and developing countries in the 20th century have contributed enormously to an empirically based understanding of the global process of mortality decline and epidemiological transition.

Evidence began to emerge in the 1940s and 1950s that mortality was declining in developing countries as well as industrialized countries.[23] The importation of public health techniques from the West into many low-income countries resulted in death rates that were lower than in industrialized countries at a similar stage of economic development.[24] A study in South India, for example, found improvements in mortality were achieved using relatively simple health technologies when basic health services were provided and used by people who responded to advice with changes in health-related behavior.[25] In Sri Lanka the death rate declined from the early 1920s, apart from spikes in mortality during

outbreaks of malaria in 1935 and 1943–46. Infant mortality declined as numbers of health centers and visits to mother-and-child health clinics increased. When attendance doubled in 1937 and again in 1939, maternal mortality began to decline consistently.[26] Rapid increases in average life expectancy were also recorded in Venezuela between 1920 and 1960. A comparative analysis of mortality change in Latin American countries found slower mortality decline in the less developed countries before 1930, but similar rates of decline to those in more developed countries after this. Mortality decline had become more dependent on imported preventive and public health measures, which had a significant impact on death rates that was independent of economic development.[27] In Guyana the death rate declined between 1911 and 1940 despite economic stagnation, and again after World War II (1939–45) despite adverse economic circumstances.[28] In Chile the crude death rate declined from 31 to 8 per 1,000 between the early 1920s and 1970s, and a similar decline from 25 to 9 per 1,000 was recorded in Mexico.[29] Differential decline in infectious disease mortality between the regions of Mexico was attributed to differences in the application of public health measures.[30] Data from Cuba provided evidence of epidemiological transition, with a rapid decline in infectious disease mortality over the period 1901–53, and an increase in the proportion of deaths attributed to cardiovascular disease from 27% to 38%.[31]

In the second half of the 20th century rapid improvements in life expectancy were achieved in some developing countries in the absence of significant national wealth creation, as the governments prioritized basic health and social programs in Costa Rica, Sri Lanka, Kerala (India), Cuba, and China, in particular. Implementation was driven by ideology in China, by social and political will in the other countries,[32] and generally by a commitment to equity. Mass education and effective political participation fostered a more equitable distribution of resources and access to schools and health services in Kerala State.[33] Besides Kerala, Caldwell identified 11 countries that had health achievements well beyond those predicted by per capita incomes. He found that educational levels of women of maternal age, family planning, and male education were the factors most strongly associated with health improvement.[34] Several studies identified the particular importance of female education as a key determinant of improvement in child survival, an effect that was independent of people's income and other indicators of material well-being.[35] Even so, favorable social conditions, including greater independence for women and more egalitarian, radical, and democratic traditions, only began to foster improvements in health when appropriate technologies from the West became available. Average life expectancy in 1930 was just 30 years in Kerala and 38 years in Sri Lanka, but by the early 1980s it had increased dramatically to 66 years and 69 years, respectively, despite the lack of improvement in the average annual income of less than US$400 per capita.[36] This experience of mortality decline without much improvement in people's economic standard of living clearly does not imply that economic factors are unimportant as determinants of health. On the contrary, evidence indicates that poverty is a major contributor to high mortality in

developing countries, and wide differentials in health measures exist in relation to indicators of people's economic standard of living, such as per capita income or the amount required to purchase a nutritionally adequate diet.[37]

Deaths in childhood and during childbirth are considerably more common among the poor in developing countries, and both child and maternal mortality are more poverty-sensitive health indicators than average life expectancy. Data for 1990 indicate that 35% of deaths at ages under 5 years in developing countries occurred among the poorest 20% of the population, based on a ranking of countries by per capita income.[38] The effect of poverty is much greater at younger ages, and the death rate for children aged 0–14 years was nine times higher among the poorest 20% than among the richest 20%, while at ages 45 and over it was about twice as high.[39] This reflected a poor-rich ratio for mortality that was much greater for communicable diseases than for those assumed to be noncommunicable.[40] World Health Organization data from 46 developing countries allowed more accurate identification of the poorest in terms of the income required to purchase a nutritionally adequate diet. On average, child (0–4 years) mortality around 1990 was four to five times higher for girls and boys below this poverty line compared with those above it, while among adults aged 15–59 years it was about four times higher for men below the poverty line and twice as high for women.[41] Indicators of poverty based on nonfinancial resources also identify mortality differentials, particularly mother's educational level and father's occupation.[42]

The association of higher mortality in developing countries with poverty and lack of education clearly does not imply that higher life expectancy in industrialized countries was achieved through improvements in the economic standard of living and education,[43] and much evidence to the contrary is presented in later chapters. Differences in child survival by educational level of mothers in developing countries in the second half of the 20th century were greater than they were in Western countries in the 19th century. In some cultures, adoption of Western health technologies was inhibited by traditional belief systems and attitudes, although these influences could be overcome through parental education or when people moved to the towns.[44] Caldwell found that the influence of women's education on child survival in Nigeria was independent of economic circumstances and distance to health facilities.[45] Mother's education and literacy probably explain much of the differential in child survival in developing countries because relevant behaviors, such as seeking maternal and child health services, having children vaccinated, and maintaining hygiene in the home, were more likely to be practiced by women who had attended school.[46] Caldwell concluded that evidence-based national policies for developing countries would support investment in education, particularly schooling for girls, preventive health services, and basic curative care.[47]

Clearly, growth in national income can provide more resources for health and education services;[48] however, where the political will and social organization existed, countries have implemented basic health programs in adverse economic circumstances. In the 1950s and 1960s, governments in many low-income

countries provided health services based on simple health technologies and knowledge that had been successfully applied earlier in Western countries, and average life expectancy increased from 35 to 45 years. In a study of international mortality data, Preston found that the decline in death rates in developing countries was due mainly to preventive health interventions introduced from industrialized countries, together with some basic treatment measures developed since the 1940s.[49] Although gross domestic product per capita, a proxy for average income, was an important determinant of life expectancy particularly in poorer countries, he found that improvements in life expectancy in the 1950s and 1960s were not due to improvements in people's economic standard of living. Preventive interventions such as immunization, environmental sanitation, insect control, health education, and basic mother and child health services were largely responsible in developing countries, while immunization, sulfonamides, and antibiotics were the main causes of mortality decline in industrialized countries in that period.[50]

In the first three decades of the 20th century, average life expectancy increased regardless of the level of average per capita income, and it increased even more rapidly in the next three decades. Factors other than changes in the national income of countries accounted for 75–90% of the global increase in life expectancy over the whole period.[51] For almost all countries, preventive and public health interventions, simple technologies, and basic health services contributed more to the reduction in mortality than any rise in per capita incomes.[52] Preston suggested that the decline in mortality in the United States in the first few decades of the 20th century was based on application of the germ theory and knowledge about how germs spread that had indicated ways of reducing the risk of infection.[53] Caldwell emphasized the importance of people believing that the new knowledge was correct and accepting responsibility to act on it as far as they could, and also use health services that became available. He suggested that the development of education and secular rational traditions contributed to a growing belief that people could have some control over events affecting health.[54] Socioeconomic and cultural differences led to differences in the timing and course of the mortality transition between subgroups of the population, exacerbating already wide socioeconomic differentials in mortality. At the beginning of the 20th century, children of unskilled workers in England were about half as likely to survive to five years of age as children of parents with professional occupations.[55] Differentials in child mortality were not as great in the United States, but they increased as the professional classes became more aware of the dangers of infection and more rapidly adopted preventive measures such as hand-washing and food hygiene measures.[56]

Although the existence of microorganisms was not known before the bacteriological findings at the end of the 19th century, this did not preclude the development of effective preventive measures and public health interventions in England. The advances in life expectancy in the 20th century should be viewed in the context of a preventive approach to disease that had been developing over many centuries. As early as the plague pandemics of the 14th century, it

was recognized that the disease was associated with certain environments and movement of people from one area to another. Ideas developed about the relationship between disease and environmental conditions, and in the second half of the 17th century demographic and statistical methods were developed by John Graunt and William Petty for monitoring historical trends in mortality and identifying meaningful associations with living conditions.[57] Various preventive measures were adopted in Europe and Asia to protect populations from devastating outbreaks of plague, and in England the disease did not recur in the 18th century. The long-term perspective of this study encompasses the development of preventive approaches to disease in response to devastating epidemics of plague, smallpox, typhus, and cholera that laid the foundation for the later rapid improvements in health and life expectancy in the 20th century.

It was well known in the 18th century that diseases were linked with environmental conditions, and a new public spiritedness infused attempts to intervene and protect people.[58] Hygiene measures were adopted in army camps and hospitals, which were amenable to the disciplines required.[59] Measures to improve cleanliness and ventilation were introduced that were appropriate for reducing the risk of patients transmitting infection to each other, and typhus recovery rates improved in British naval hospitals after the 1740s.[60] Hygiene measures for preventing the spread of infection were more widely promoted among the civilian population through dispensaries and the "lying-in" movement. Although the impact of such measures on overall mortality in the 18th century may not have been great, they were nevertheless important in the development of a preventive approach. Dispensaries and voluntary hospitals in England provided care for casualties and patients with various diseases including respiratory conditions;[61] however, death rates in some major towns had begun to decline before hospitals opened in the 1770s.[62]

The need to care for the sick was more widely recognized by the end of the 19th century, and further efforts were made to organize health care.[63] Until that time, those suffering from infectious diseases were not hospitalized on a large scale and there was still no effective treatment for major killers such as smallpox, typhus, typhoid, and tuberculosis, which caused high mortality regardless of socioeconomic circumstances. Medicines and treatment could not have made any significant contribution to the decline in infectious disease mortality before the 1940s, as McKeown emphasized.[64] Antitoxin introduced for the treatment of diphtheria in the 1890s may have been the only effective treatment for a major infectious disease before sulfonamides and antibiotics. Salvarsan was used to treat syphilis from World War I (1914–18) until the 1940s when more effective treatment with penicillin became available, although records do not suggest the disease was a major cause of death in England.[65] Sulfapyridine for the treatment of respiratory disease and sulfonamides for whooping cough and puerperal fever were not available until the late 1930s, there were no effective drugs for treatment of tuberculosis until 1947 when streptomycin became available, and chloramphenicol for the treatment of typhoid was not available until 1950. In addition to antibiotics for infectious diseases, medical interventions for

chronic diseases are likely to have contributed significantly to the increase in life expectancy at adult ages in the second half of the 20th century. However, preventive measures had a major impact on chronic disease death rates, through benefits in adult life and as a result of improvements in health and physiological development in childhood, as discussed in chapters 9–12. Evidence presented in chapters 4–8 also indicates major reductions in mortality from acute infectious diseases and tuberculosis were achieved through preventive and public health measures long before effective treatments became available.

Smallpox was controlled in England over the course of the 19th century through vaccination and other preventive measures.[66] This success laid the foundation for global eradication and inspired immunization programs against other diseases, which reduced child mortality worldwide in the second half of the 20th century. McKeown greatly underestimated the contribution of smallpox vaccination and control measures against plague and typhus as a result of focusing on the second half of the 19th century.[67] By that time, plague had long since been prevented from entering the country again and typhus and smallpox mortality had been greatly reduced.[68] The decline in incidence of these three major infectious diseases had a fundamental effect on people's lives regardless of their economic circumstances or nutritional status. Even so, despite demonstration of effective preventive intervention, destructive diseases were still viewed by many as the work of God, and there was opposition to smallpox vaccination as an interference in "God's works."[69] In spite of this objection, vaccination was widely and effectively used long before the theory of germs, identification of microbial agents, or any concept of immune response provided a scientific explanation. Similarly, when cholera arrived in Britain in 1831 during a Europe-wide pandemic, it was greatly feared for the rapid wasting and horrifying death it caused, and there was little knowledge of how the disease spread or how to treat the sick. Nevertheless, control measures such as quarantine were implemented and routes of entry into the country were monitored in a preventive approach reminiscent of the response to the threat of plague in earlier times. The arrival of cholera also led to a public health movement aimed at preventing the disease, and evidence of association with sources of drinking water and suspected contamination eventually led to a theory of transmission and a scientific rationale for improving water and sewage systems, decades before the existence of germs became known and widely accepted.

The germ theory of disease causation became the subject of bitter and lasting controversy at the end of the 19th century, reminiscent of the controversy over the theory of evolution.[70] Many found these new scientific ideas incompatible with religious doctrine and the notion of an "all-wise God." The findings of bacteriological investigations based on growing pure cultures and demonstrating association with specific infectious diseases contributed to acceptance of the germ theory as a more scientific explanation for epidemics and patterns of disease.[71] This dispelled much of the mystery and irrationality that had surrounded infectious diseases, and specific microorganisms were established as the causal agents of diseases that had been major killers known for centuries. The

bacteriological discoveries lent support to the idea of isolating severe cases to prevent transmission, although for some infectious diseases there was a decline in case fatality rather than incidence.[72] To explain the decline of the major infectious diseases, it is important to consider turning points in death rates in relation to changes that could affect resistance such as diet, nutritional status, coinfections, sociodemographic changes that affect severity of infection, and preventive measures that affect exposure or the immune response to infection.

The new sciences of bacteriology, virology, and immunology contributed to the development of further preventive measures, and the national registration system introduced in England in 1837 provided mortality statistics with which to monitor the impact. Records of cause-specific mortality from England, with supporting data from other Western countries, form the basis of this study of the epidemiological transition and illustrate its main characteristics—acute infectious diseases and tuberculosis being superseded by other chronic diseases as the predominant causes of death, increased average life expectancy, an increasing proportion of older people, a shift in age at death from childhood to older adult ages, and an increase in population. With the lower death rates from infectious diseases in the first half of the 20th century, concern grew about chronic diseases such as heart disease, stroke, and cancers that were seen to be increasingly common causes of death among adults in middle and older age groups. New methods and techniques were developed to determine the geographic distribution of deaths from chronic diseases,[73] and from the 1950s epidemiology focused more on identifying and quantifying risk factors.

In the second half of the 20th century, epidemiological studies investigated differentials in mortality in relation to behavior, lifestyle, socioeconomic circumstances, environment, and geographic location. Before the global HIV/AIDS pandemic from the 1980s and the growing problem of resistance to antibiotics, there had generally been a shift in scientific attention away from infectious diseases toward the chronic diseases. Nevertheless, research continued into the transmission and population dynamics of infectious diseases, partly because developing countries were at an earlier stage of epidemiological transition. Transmission theory developed and research findings greatly increased knowledge about factors affecting severity, leading to more appropriate measures for controlling acute infectious disease worldwide.[74] However, a unified approach such as that developing with the germ theory when mortality from acute infectious diseases and tuberculosis was high in Europe, became less sustainable as a mixed burden of infectious diseases and chronic diseases with very different etiologies emerged globally.[75]

In recent years, there has been a growing awareness among epidemiologists that a multidisciplinary approach is required for a more complete understanding of variation and change in disease and mortality patterns. In the past, historical research tended to be fragmented, with studies by historians, geographers, economists, demographers, and medical specialists focusing on different aspects and phases of change in mortality, particular diseases, or issues of of special interest within their discipline. Hitherto, a comprehensive overview

of the epidemiological transition has not been presented using data from one country, and the emergence of chronic diseases has not been assessed in the context of changes in the infectious disease environment. Through a synthesis of findings from research in different disciplines and analysis of trends in death rates from specific diseases, this study aims to provide a more complete account of the epidemiological transition, and a new perspective that reflects the interrelated nature of diseases, particularly links between infections and chronic diseases. The limitations of the original concept of "epidemiologic transition" and theoretical framework for this study are considered in the next chapter, together with issues relating to the data available, classification of causes of death, and disease categories.

Chapter Two

Theoretical Framework, Data, and Study Outline

The Concept of Epidemiological Transition

The original concept of "epidemiologic transition" outlined by Omran in 1971 focused attention on extremely important changes in cause of death.[1] The infectious diseases that had caused very high mortality among children and adults in the past had been replaced as the main causes of death in industrialized countries by chronic diseases among adults assumed to be noninfectious in origin. There was much criticism of his concept as a "theory," and it does appear simplistic in the light of subsequent findings about the etiology of chronic diseases. Patterns of disease, mortality, and life expectancy are more complex than a simple shift to predominantly chronic diseases as infectious disease mortality declined.[2] Nevertheless, despite criticisms and revisions of the phases of "epidemiologic transition," it has provided a framework for comparison of secular trends and profiles of disease in different populations. An objective of this study is to develop a perspective on historical changes in disease and mortality that goes beyond descriptive accounts of different phases of transition, takes into account disease etiology and the interrelationships between diseases, and contributes to an understanding of how contemporary patterns of disease have come about.

Omran's first formulation of his theory consisted of five propositions, which can be summarized as follows:

1. "Mortality is a fundamental factor in population dynamics."
2. "A long-term shift occurs in mortality and disease patterns whereby pandemics of infection are gradually displaced by degenerative and man-made diseases as the chief form of morbidity and primary cause of death."
3. "The most profound changes in health and disease patterns obtain among children and young women."

4. "The shifts in health and disease patterns that characterise the epidemiologic transition are closely associated with the demographic and socioeconomic transitions that constitute the modernisation complex."
5. There are three basic models of the "epidemiologic transition" corresponding to the experience of most Western countries, countries such as Japan where the transition was more rapid, and developing countries where the decline in mortality commenced more recently.[3]

The identification of a shift in disease patterns distinguished the theory from demographic transition theory, although there was more interest among demographers than epidemiologists until the 1990s. Omran's main concern may have been fertility and the dynamics of population change rather than chronic diseases, and he later added a fourth model in which rapid mortality decline was followed by fertility decline.[4] Some aspects of Omran's concept of epidemiological-demographic change are well accepted, such as increasing life expectancy, a shift in the age distribution of populations from younger to older ages, more rapid population growth, and changes in the cause-of-death structure concomitant with increasing average life expectancy. The key role of mortality decline in the demographic transition is also generally accepted, and with some exceptions mortality decline preceded fertility decline and may have precipitated it. The decline of child mortality made limitation of family size a less risky strategy, and its particularly important contribution to the increase in average life expectancy is considered in chapter 8. The change in the predominant causes of death from acute infectious diseases and tuberculosis to other chronic diseases is uncontroversial as the main feature of epidemiological transition, although exceptions to the different phases of transition, together with etiological considerations, indicate the need for a more empirically based and scientific theory of epidemiological-demographic change.

Omran identified three categories of determinants of "epidemiologic transition": (1) ecobiologic determinants, involving the "complex balance between disease agents, the level of hostility in the environment and the resistance of the host"; (2) socioeconomic, political, and cultural determinants, including "standards of living, health habits and hygiene and nutrition"; and (3) medical and public health determinants, including "public sanitation, immunization and the development of decisive therapies." In general terms, he attributed the decline in mortality and the change in disease patterns to factors linked with "modernisation" and socioeconomic change (proposition 4). He suggested that socioeconomic changes and improvements in "standards of living" were the important changes underlying the "epidemiologic transition" in Western populations, while public health and medical technologies were more important in the later transition being experienced in developing countries. As mentioned, the current study suggests more historical continuity with respect to the role of public health and preventive measures as the main cause of increased life expectancy before and after antibiotics were introduced. Omran's generalizations about the effect on mortality of socioeconomic change and improvement in "standards of living" in Western countries were not based on any specific indicators or even

more general evidence. Like McKeown, he provided no evidence of changes in the "standard of living" or nutrition in periods of mortality decline. It will also be shown that many differences in the trends and turning points in death rates for particular diseases belie such a generalization about the decline in airborne infectious disease mortality.

Omran's theory has been criticized as insufficiently epidemiological, and it lacked consideration of disease etiology and causality.[5] Reference to the chronic diseases that have replaced acute infectious diseases and tuberculosis as the main causes of death as "degenerative and man-made" has been criticized for the vaguely moralistic undertones and an assumption that pathogenesis was an age-related process of "degeneration."[6] The term "degenerative" may not be appropriate given underlying processes now known to be involved in heart disease, cerebrovascular disease, and cancer, for example. The concept of "man-made" disease, while reflecting the health consequences of human action and behavior, underestimates socioeconomic, cultural, and other determinants of health that can be difficult for individuals to control. While better-off people have the means to adopt new products and habits later discovered to be detrimental to health, they are likely to respond more quickly to advice, while the poor are more restricted by their social and economic circumstances.[7] Other characterizations of chronic disease such as "diseases of affluence" and "diseases of civilization" have no clear scientific basis,[8] and the relationship between economic standard of living and chronic disease mortality is complex. The more affluent countries have high death rates from these diseases, although even higher age-specific death rates are found in some low-income countries that have undergone little industrial development. This belies any simplistic notion of "diseases of the industrial way of life," while higher than average death rates from chronic disease among the lower socioeconomic groups within high-income countries indicate that the term "diseases of affluence" is not appropriate.[9]

There has also been criticism and considerable modification of Omran's phases of transition in disease patterns: "age of pestilence and famine," "age of receding pandemics," and "age of degenerative and man-made diseases." Reference to an "age of receding pandemics" is inappropriate as smallpox, tuberculosis, dysentery, and pneumonia were endemic in 19th-century England.[10] The global pandemic of influenza in 1918 was also a significant anomaly, as it probably caused more deaths than any pandemic in history during Omran's era of "degenerative and man-made diseases." The HIV/AIDS pandemic had yet to emerge at the time he was writing, but this clearly constitutes a major exception. The emphasis on a decline of infectious disease has been called into question as new diseases including HIV/AIDS and reemerging diseases such as tuberculosis have challenged the implicit assumption of inevitable progress with the control of infectious diseases leading to improved quality of life.[11] Historically, epidemic diseases have had their own dynamics,[12] and a resurgence of some acute infectious diseases of childhood, including whooping cough, measles, and streptococcus-A infection in the United States, for example, has raised further doubts about the sustainability of infectious disease control.[13]