Epidemiology

An Introduction

SECOND EDITION

KENNETH J. ROTHMAN

OXFORD
UNIVERSITY PRESS

Epidemiology

An Introduction

SECOND EDITION ■

KENNETH J. ROTHMAN

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press, Inc., publishes works that further
Oxford University's objective of excellence
in research, scholarship, and education.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016
www.oup.com

Oxford is a registered trademark of Oxford University Press

Library of Congress Cataloging-in-Publication Data
Rothman, Kenneth J.
Epidemiology : an introduction / Kenneth J. Rothman. — 2nd ed.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-0-19-975455-7 (pbk. : alk. paper)
I. Title.
[DNLM: 1. Epidemiologic Methods. WA 950]
614.4—dc23
2011037114

9
Printed in the United States of America
on acid-free paper

To Emily, Margaret, and Samantha

PREFACE TO SECOND EDITION

Some observers appear to believe that epidemiology is no more than the application of statistical methods to the problems of disease occurrence and causation. But epidemiology is much more than applied statistics. It is a scientific discipline with roots in biology, logic, and the philosophy of science. For epidemiologists, statistical methods serve as an important tool but not a foundation. My aim in this book is to present a simple overview of the concepts that are the underpinnings of epidemiology, so that a coherent picture of epidemiologic thinking emerges for the student. The emphasis is not on statistics, formulas, or computation but on epidemiologic principles and concepts.

For some, epidemiology is too simple to warrant serious attention, and for others it is too convoluted to understand. I hope to persuade the reader that neither view is correct. The first chapter illustrates that epidemiology is more than just applying "common sense," unless one has uncommonly good common sense. Although it is unusual to begin an epidemiology text with a presentation of confounding, I believe that the problem of confounding exemplifies why we need to understand epidemiologic principles lest we fall victim to fallacious inferences. At the other extreme, those who believe that epidemiology is too complicated might think differently if they had a unifying set of ideas that extended across the many separate topics within epidemiology. My goal in this book has been to provide that unifying set of ideas.

The second chapter, new to this second edition, adds a historical perspective for the reader who is new to epidemiology, in the form of capsule profiles of pioneers in epidemiology and public health. The chapter illustrates the deep historical roots of epidemiology, highlighting important contributions from trailblazers such as Avicenna, Graunt, Nightingale, and Lane-Claypon. Chapter 3 lays a conceptual foundation by elucidating a general model of causation and discussing the process of causal inference. All too often, these concepts are skipped over in scientific education. Nevertheless, for epidemiologists they are bedrock concerns that belong in any introduction to the field. Chapter 4 continues with a description of the basic epidemiologic measures, and Chapter 5 covers the main study types. An important thread for the student is the emphasis on measurement and how to reduce or describe measurement error. Chapter 6, also new to the

second edition, presents a capsule summary of the key ideas in infectious disease epidemiology.

The next two chapters deal with measurement error. Systematic error, or bias, is treated first, in Chapter 7, and random error in Chapter 8. Chapter 9 introduces the basic analytic methods for estimating epidemiologic effects; these methods are extended in Chapter 10 to stratified data. Chapters 11 and 12 address the more advanced topics of interaction and regression modeling. These are subjects to be explored in detail in more advanced courses, but their presentation here in elementary terms lays the groundwork for advanced study. It also draws a boundary between the epidemiologic approach to these topics and nonepidemiologic approaches that can take the analysis in the wrong direction. The final chapter deals with clinical epidemiology, a branch of epidemiology that continues to grow in scope and importance.

Epidemiologic concepts are evolving, as any comparison of this volume with earlier texts will reveal. To complement the book, the publisher has graciously agreed to host a Website that will support reader participation in discussing, extending, and revising points presented in the book. The Website will post contributed answers to the questions raised at the end of each chapter in the text. Interested readers can find this site at www.oup.com/us/epi/.

Along the way I have received invaluable feedback from many students and colleagues. I am especially grateful to Kristin Anderson, Georgette Baghdady, Dan Brooks, Robert Green, Sander Greenland, Tizzy Hatch, Bettie Nelson, Ya-Fen Purvis, Igor Schillevoort, Bahi Takkouche, and Noel Weiss. Cristina Cann provided unflagging and generous encouragement; now that she is no longer here to continue advising me, she is sorely missed. Katarina Bälter deserves special mention for her careful reading of the manuscript and patient, helpful criticisms. Finally, I am indebted to my colleague Janet Lang, who gently prodded me at the right time and was an inspiration throughout.

Kenneth J. Rothman
Newton, Massachusetts
July, 2011

CONTENTS

1

Introduction to Epidemiologic Thinking

This book presents the basic concepts and methods of epidemiology, which is the core science of public health. Epidemiology has been defined as "the study of the distribution and determinants of disease frequency,"[1] or even more simply as "the study of the occurrence of illness."[2]

The principles of epidemiologic research appear deceptively simple. This appearance seduces some people into believing that anyone can master epidemiology just by applying some common sense. In a way that view is correct, but it is nevertheless an oversimplification. The problem is that the kind of common sense that is required may be elusive without training in epidemiologic concepts and methods. As a means of introducing epidemiologic thinking, I shall illustrate this point with some examples of the fundamental epidemiologic concept known as *confounding*. Confounding is described in greater detail in Chapter 7. Here I aim only to demonstrate the kind of problem that epidemiology deals with routinely. I will also point out some basic fallacies, the kind that can be found in the newspapers on a regular basis and that occur commonly among those who are not well versed in epidemiologic thinking.

Common sense tells us that residents of Sweden, where the standard of living is generally high, should have lower death rates than residents of Panama, where poverty and more limited health care take their toll. Surprisingly, however, a greater proportion of Swedish residents than Panamanian residents die each year. This fact belies common sense. The explanation lies in the age distribution of the populations of Sweden and Panama. Figure 1–1 shows the population pyramids of the two countries. A *population pyramid* displays the age distribution of a population graphically. The population pyramid for Panama tapers dramatically from younger to older age groups, reflecting the fact that most Panamanians are in the younger age categories. In contrast, the population pyramid of Sweden is more rectangular, with roughly the same number of people in each of the age categories up to about age 60 years and some tapering above that age. As these graphs

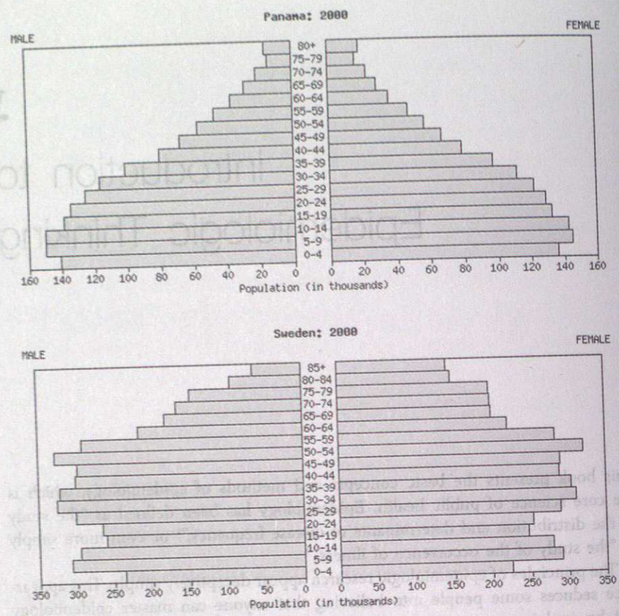

Figure 1–1 Age distribution of the populations of Panama and Sweden (population pyramids).
SOURCE: U.S. Census Bureau, International Data Base.

make clear, Swedes tend to be older than Panamanians. For people of the same age in the two countries, the death rate among Swedes is indeed lower than that of Panamanians, but in both places older people die at a greater rate than younger people. Because Sweden has a population that is on the average older than that of Panama, a greater proportion of all Swedes die in a given year, despite the lower death rates within age categories in Sweden compared with Panama.

This situation illustrates what epidemiologists call *confounding*. In this example, age differences between the countries are confounding the differences in death rates. Confounding occurs commonly in epidemiologic comparisons. Consider the following mortality data, summarized from a study that looked at smoking habits of residents of Whickham, England, in the period 1972–1974 and then tracked the survival over the next 20 years of those who were interviewed.[3–5] Among 1314 women in the survey, almost half were smokers. Oddly, proportionately fewer of the smokers died during the ensuing 20 years than nonsmokers. The data are reproduced in Table 1–1.

Only 24% of the women who were smokers at the time of the initial survey died during the 20-year follow-up period. In contrast, 31% of those who were

Table 1–1 Risk of Death in a 20-Year Period Among Women in Whickham, England, According to Smoking Status at the Beginning of the Period*

	Smoking		
Vital Status	**Yes**	**No**	**Total**
Dead	139	230	369
Alive	443	502	945
Total	582	732	1314
Risk (Dead/Total)	0.24	0.31	0.28

*Data from Vanderpump et al.[5]

nonsmokers died during the follow-up period. Does this difference indicate that women who were smokers fared better than women who were not smokers? Not necessarily. One difficulty that many researchers quickly spot is that the smoking information was obtained only once, at the start of the follow-up period. Smoking habits for some women will have changed during the follow-up. Could those changes explain the results that appear to confer an advantage on the smokers? It is theoretically possible that all or many of the smokers quit soon after the survey and that many of the nonsmokers started smoking. Although this scenario is possible, it is implausible; without evidence for these changes in smoking behavior, this implausible scenario is not a reasonable criticism of the study findings. A more realistic explanation for the unusual finding becomes clear if we examine the data within age categories, as shown in Table 1–2. (The risks for each age group are calculated by dividing the number who died in each smoking group by the total number of those dead or alive.)

Table 1–1 combines all of the age categories listed in Table 1–2 into a single table, which is called the *crude* data. It is obtained by adding together, or collapsing, the data for each age category in Table 1–2. The more detailed display of the same data in Table 1–2 is called an *age-specific* display, or a display *stratified* by age. The age-specific data show that in the youngest and oldest age categories there was little difference between smokers and nonsmokers in risk of death. Few died among those in the youngest age categories, regardless of whether they were smokers or not, whereas among the oldest women, almost everyone died during the 20 years of follow-up. For women in the middle age categories, however, there was a consistently greater risk of death among smokers than nonsmokers, a pattern contrary to the impression gained from the crude data in Table 1–1.

Why did the nonsmokers have a higher risk of death in the study population as a whole? The reason is evident in Table 1–2: A much greater proportion of the nonsmoking women were in the highest age categories, the categories that contributed a proportionately greater number of deaths. The difference in age distribution between smokers and nonsmokers reflects the fact that, for most people, lifelong smoking habits are determined early in life. During the decades preceding the study in Whickham, there was a trend for increasing proportions of young women to become smokers. The oldest women in the Whickham study grew up during a period when few women became smokers, and they tended to remain nonsmokers for the duration of their lives. As time went by, a greater proportion

Table 1–2 Risk of Death in a 20-Year Period Among Women in Whickham, England, According to Smoking Status at the Beginning of the Period, by Age

Age	Vital Status	Smoking: Yes	Smoking: No	Total
18–24	Dead	2	1	3
	Alive	53	61	114
	Risk	0.04	0.02	0.03
25–34	Dead	3	5	8
	Alive	121	152	273
	Risk	0.02	0.03	0.03
35–44	Dead	14	7	21
	Alive	95	114	209
	Risk	0.13	0.06	0.09
45–54	Dead	27	12	39
	Alive	103	66	169
	Risk	0.21	0.15	0.19
55–64	Dead	51	40	91
	Alive	64	81	145
	Risk	0.44	0.33	0.39
65–74	Dead	29	101	130
	Alive	7	28	35
	Risk	0.81	0.78	0.79
75+	Dead	13	64	77
	Alive	0	0	0
	Risk	1.00	1.00	1.00

of women who were passing through their teenage or young adult years became smokers. The result was strikingly different age distributions for the female smokers and nonsmokers of Whickham. Were this difference in the age distribution ignored, one might conclude erroneously that smoking was not related to a higher risk of death. In fact, smoking is related to a higher risk of death, but confounding by age has obscured this relation in the crude data of Table 1–1. In Chapter 10, I will return to these data and show how to calculate the effect of smoking on the risk of death after removal of the age confounding.

Confounding is a problem that pervades many epidemiologic studies, but it is by no means the only issue that bedevils epidemiologic inferences. One day, readers of the *Boston Globe* opened the paper to find a feature story about orchestra conductors. The point of the article was that conducting an orchestra is salubrious, as evinced by the fact that so many well-known orchestra conductors have lived to be extremely old. Common sense suggests that if the people in an occupation tend to live long lives, the occupation must be good for health. Unfortunately, what appeared to be common sense for the author of the article is not very sensible from an epidemiologic point of view. The long-lived conductors who were cited in the article were mentioned because they lived to be old. Citing selected examples in this way constitutes *anecdotal* information, which can be extremely misleading. For all we know, the reporter searched specifically for examples of

elderly conductors and overlooked other conductors who might have died at an earlier age. Most epidemiologists would not even classify anecdotal information as epidemiologic data at all.

Furthermore, the reporter's observation had problems that went beyond the reliance on anecdotes instead of a formal evaluation. Suppose that the reporter had identified all orchestra conductors who worked in the United States during the past 100 years and studied their longevity. This approach avoids relying on hand-picked examples, but it still suffers from an important problem that leads to an incorrect answer. The problem is that orchestra conductors are not born as orchestra conductors. They become conductors at a point in their careers when they may have already attained a respectable age. If we start with a group of people who are 40 years old, on the average they are likely to survive to an older age than the typical person who was just born. Why? Because they have a 40-year head start; if they had died before age 40, they could not have been part of a group in which everyone is 40 years old. To determine whether conducting an orchestra is beneficial to health, one should compare the risk of death among orchestra conductors with the risk of death among other people who have attained the same age as the conductors. Simply noting the average age at death of the conductors gives the wrong answer, even if all orchestra conductors in a population are studied.

Here is another example that makes this point clearly. Suppose that we study two groups of people over a period of 1 year, and we look at the average age at death among those who die during that year. Suppose that in group A the average age at death is 4 years and in group B it is 38 years. Can we say that being a member of group A is riskier than being a member of group B? We cannot, for the same reason that the age at death of orchestra conductors was misleading. Suppose that group A comprises nursery school students and group B comprises firefighters. It would be no surprise that the average age at death of people who are currently firefighters is 38 years or that the average age at death of people who are currently nursery school students is 4 years. Still, we suspect that being a firefighter is riskier than being a nursery school student and that these data on the average age at death do not address the issue of which of these groups faces a greater risk of death. When one looks at the average age at death, one looks only at those who actually die and ignores all those who survive. Consequently, the average age at death does not reflect the risk of death but only expresses a characteristic of those who die.

In a study of factory workers, an investigator inferred that the factory work was dangerous because the average age at onset of a particular kind of cancer was lower in these workers than among the general population. But just as for the nursery students and firefighters, if these workers were young, the cancers that occurred among them would have to be occurring in young people. Furthermore, the age at onset of a disease does not take into account what proportion of people get the disease. It may have been that the number of workers who developed the cancer was no greater, or was even smaller, than the number expected to do so based on the risk of cancer in the general population. Looking at the age at which cancer develops among those who get cancer cannot address the question of risk for cancer.

These examples reflect the fallacy of comparing the average age at which death or disease strikes, rather than comparing the risk of death between groups of the

same age. In later chapters, I will explore the proper way to make epidemiologic comparisons. The point of these examples is to illustrate that what may appear to be a commonsense approach to a simple problem can be overtly wrong until we educate our common sense to appreciate better the nature of the problem. Any sensible person can understand epidemiology, but without considering the principles outlined in this book, even a sensible person using what appears to be common sense is apt to go astray. By mastering a few fundamental epidemiologic principles, it is possible to refine our common sense to avoid these traps.

QUESTIONS

1. Age is a variable that is often responsible for confounding in epidemiology, in part because the occurrence of many diseases changes with age. The change in disease risk with age is often referred to as the effect of age. Does it make sense to think of age as having an effect on disease risk, or is it more sensible to think that the effect of age is itself confounded by other factors?

2. More people in Los Angeles die from cardiovascular disease each year than do people in San Francisco. What is the most important explanation for this difference? What additional factors would you consider to explain the difference in the number of deaths?

3. In Table 1–2, which age group would you say shows the greatest effect of smoking on the risk of death during the 20-year interval? How have you defined "greatest effect"? What other way could you have defined it? Does your answer depend on which definition you use?

4. On a piece of graph paper, use the data in Table 1–2 to plot the 20-year risk of death against age. Put age on the horizontal axis and the 20-year risk of death on the vertical axis. Describe the shape of the curve. What biological forces account for the shape?

5. A physician who was interested in jazz studied the age at death of jazz musicians, whom he identified from an encyclopedia of jazz. He found that the average age at death of the jazz musicians was about the same as that in the general population. He concluded that this finding refuted the prevailing wisdom that jazz musicians tended to live dissolute lives and therefore experienced greater mortality than other people. Explain his error.

6. A researcher determined that being left-handed was dangerous, because he found that the average age at death of left-handers was lower than that of right-handers. Was he correct? Why or why not?

7. What is the underlying problem in comparing the average age at death, or the average age at which a person gets a specific disease, between two populations? How should you avert this problem?

REFERENCES

1. MacMahon B, Pugh TF. *Epidemiology: Principles and Methods.* Boston: Little, Brown; 1970:137–198,175–184.
2. Gaylord Anderson, as cited in: Cole P. The evolving case-control study. *J Chron Dis.* 1979;32:15–27.
3. Appleton DR, French JM, Vanderpump MPJ. Ignoring a covariate: an example of Simpson's paradox. *Am Statistician.* 1996;50:340–341.
4. Turnbridge WMG, Evered DC, Hall R, et al. The spectrum of thyroid disease in a community. *Clin Endocrinol.* 1977;7:481–493.
5. Vanderpump MPJ, Turnbridge WMG, French JM, et al. The incidence of thyroid disorders in the community: a twenty-year follow-up of the Whickham survey. *Clin Endocrinol.* 1995;43:55–69.

2

Pioneers in Epidemiology and Public Health

Throughout human history, the major problems of health that men have faced have been concerned with community life, for instance, the control of transmissible disease, the control and improvement of the physical environment (sanitation), the provision of water and food of good quality and in sufficient supply, the provision of medical care, and the relief of disability and destitution. The relative emphasis placed on each of these problems has varied from time to time, but they are all closely related, and from them has come public health as we know it today.

GEORGE ROSEN, *A History of Public Health*, 1958[1]

ORIGINS OF PUBLIC HEALTH

Public health may be defined as the community effort to protect, maintain, and improve the health of a population by organized means, including preventive programs, hygiene, education, and other interventions. Preventive medicine and health care are important components of public health, but the reach of public health extends beyond medicine to clean water, sanitation, housing, sex education, and other areas that affect the health of communities. Among the fields that contribute to public health, epidemiology plays a crucial role, but many other disciplines are involved. Public-health efforts may involve contributions from fields as diverse as engineering, architecture, biology, social science, ecology, and economics.

The origin of public health dates from the first aggregation of small clans into larger, settled communities. Some basic needs, such as provision of potable water and disposal of bodily waste, were best addressed as community concerns. Public wells that provided clean water for drinking and aqueducts that transported water from mountain springs or creeks into towns, cities, and agricultural fields were community efforts that had obvious health benefits. Fountains, public baths,

and water delivery systems supplying individual homes were features of ancient civilizations in the Indus River valley, Mesopotamia, Persia, Troy, and China.

Treatment of drinking water in settling basins and purification with chemicals was practiced in many early civilizations. Sanitary removal of human and animal waste is a fundamental public-health concern, one that became more challenging as settlements increased in size. The construction of sewers to remove waste materials was also common in antiquity. In ancient Rome, water delivery and sanitation were sophisticated engineering feats. Roman aqueducts brought mountain water from sources as far away as 100 km or more to the city, using trenches, arched bridges, walls, tunnels, and siphons to traverse varied landscapes. The water supplied a few wealthy homes, public baths, and public fountains, from which it overflowed to flush the streets. Excess water collected into a remarkable sewer system, culminating in the *cloaca maxima* (literally, "greatest sewer"), which conveyed effluent from the city into the Tiber River, facilitating hygiene for the city, though at the expense of polluting the river. Parts of the cloaca maxima are still connected to the Roman sewage system in the 21st century. Water supply and sanitation remain essential concerns for large communities and rank as top priorities for public health today. Among those still afflicted by poverty, and with the planet facing continued population growth and mounting environmental concerns, the basic public-health needs of clean water and sanitation are unlikely to recede in importance.

Another basic public-health priority has been control of transmissible diseases. Historically, community efforts to prevent the spread of communicable disease were often hampered by inadequate understanding of effective means to control transmission. The microbial origin of many transmissible diseases became evident only during the latter part of the 19th century. Nevertheless, concern about disease transmission and measures to control it has deep roots. In the Middle Ages, the growth of cities in Europe was accompanied by epidemics of leprosy, plague, and other scourges that stirred frightened populations into drastic actions, not often successful. The Black Death was a pandemic of what is believed to be bubonic plague that swept Europe in the mid–14th century. Although its cause was not clear to medieval populations, who tended to see epidemics as divine wrath, it was considered transmissible, which led communities to bar entry to foreigners from plague-ridden areas and to isolate patients and all their contacts. During periods when plague threatened Venice, an important shipping port, all vehicles, goods, and travelers were isolated for a period of time. This procedure, today called *quarantine* (from the Italian phrase *quaranta dei*, meaning "40 days"), is still invoked occasionally. It was used to notable advantage during the outbreak of severe acute respiratory syndrome (SARS) in 2003[2] (see Chapter 6).

FROM HIPPOCRATES TO SNOW AND THE MODERN DAY

Although public health has its origins in antiquity and epidemiology is considered by many to be the fundamental science of public health, epidemiology did not develop noticeably as a scientific discipline until the 20th century. Nonetheless, certain signal contributions to public health and epidemiology represented

long-lasting, important increments to our outlook and understanding. Here we review highlights from some of the key contributors to the foundation of public health.

Hippocrates (~460–370 BC)

Hippocrates was a Greek physician who had a profound influence on the practice of medicine as well as on public health. Scholars believe that many of the surviving writings that are nominally attributed to Hippocrates were authored by others, probably his students. Nevertheless, these writings were undoubtedly influenced by Hippocrates and reflect a revolutionary approach to health and disease that Hippocrates helped to bring about. The prevailing view in the time of Hippocrates was that disease was the result of demonic possession or divine displeasure. Hippocrates turned attention to earthly causes. Greek physicians in Hippocrates' day were itinerant. Hippocrates advised these physicians to consider what environmental factors in each community might affect locally occurring diseases. For example, in *Airs, Waters, and Places,* the influence of the environment on disease occurrence was stressed for the first time in a scholarly treatise[3]:

> Whoever wishes to investigate medicine properly, should proceed thus: in the first place to consider the seasons of the year, and what effects each of them produces for they are not at all alike, but differ much from themselves in regard to their changes. Then the winds, the hot and the cold, especially such as are common to all countries, and then such as are peculiar to each locality. We must also consider the qualities of the waters, for as they differ from one another in taste and weight, so also do they differ much in their qualities. In the same manner, when one comes into a city to which he is a stranger, he ought to consider its situation, how it lies as to the winds and the rising of the sun; for its influence is not the same whether it lies to the north or the south, to the rising or to the setting sun. These things one ought to consider most attentively, and concerning the waters which the inhabitants use, whether they be marshy and soft, or hard, and running from elevated and rocky situations, and then if saltish and unfit for cooking; and the ground, whether it be naked and deficient in water, or wooded and well watered, and whether it lies in a hollow, confined situation, or is elevated and cold; and the mode in which the inhabitants live, and what are their pursuits, whether they are fond of drinking and eating to excess, and given to indolence, or are fond of exercise and labor, and not given to excess in eating and drinking…
>
> And in particular, as the season and the year advances, he can tell what epidemic diseases will attack the city, either in summer or in winter, and what each individual will be in danger of experiencing from the change of regimen. For knowing the changes of the seasons, the risings and settings of the stars, how each of them takes place, he will be able to know beforehand what sort of a year is going to ensue. Having made these investigations, and knowing beforehand the seasons, such a one must be acquainted with each particular, and must succeed in the preservation of health, and be by no means unsuccessful in the practice of his art.

Many of the specific theories espoused in the writings that are attributed to Hippocrates would seem strange to modern readers. He believed it important to study astrology, with each astrological sign being associated with a part of the body. He also embraced the theory of humors, which held that when one of the four humors (black bile, phlegm, yellow bile, and blood) was out of balance, disease resulted. Although such theories are not in accord with modern views of disease, Hippocrates fostered a sea change away from mysticism and religion toward observation and reason as means of understanding the causes of disease.

Avicenna (Ibn Sina) (980–1037)

Avicenna, or Ibn Sina, a Persian philosopher, scientist, and physician, lived during the time that Europe was plodding through the Dark Ages, with science in full retreat. But in the Islamic world, it was a golden age of knowledge, to which Avicenna was a major contributor. A prolific genius, he wrote on many topics, foremost among them medicine and health. He is not identified specifically with public health, but, like Hippocrates, his contributions to the understanding of disease causation and his emphasis on empirical evidence had an enormous influence on both medicine and public health. His 14-volume *Canon of Medicine* is perhaps the most renowned textbook of medicine ever written. It was translated into Latin in 1187 and had a long-lasting influence in both the East and the West. In Europe, it was the primary medical text for centuries and was still in use as late as 1650.

Avicenna emphasized the need for bringing experimentation and quantification into the study of physiology and medicine. He inferred from his observations that some infectious diseases were spread by contagion, and he suggested the use of quarantine to limit their spread. The following excerpt from the introduction to Avicenna's *Canon of Medicine* shows the influence of Hippocrates and his embrace of scientific study. In it, one can easily see the foreshadowing of modern epidemiologic theory[4]:

> The knowledge of anything, since all things have causes, is not acquired or complete unless it is known by its causes. Therefore in medicine we ought to know the causes of sickness and health. And because health and sickness and their causes are sometimes manifest, and sometimes hidden and not to be comprehended except by the study of symptoms, we must also study the symptoms of health and disease. Now it is established in the sciences that no knowledge is acquired save through the study of its causes and beginnings, if it has had causes and beginnings; nor completed except by knowledge of its accidents and accompanying essentials. … These are the subjects of the doctrine of medicine; whence one inquires concerning the disease and curing of the human body. One ought to attain perfection in this research; namely, how health may be preserved and sickness cured. And the causes of this kind are rules in eating and drinking, and the choice of air, and the measure of exercise and rest; and doctoring with medicines and doctoring with the hands. All this with physicians is according to three species: the well, the sick, and the medium of whom we have spoken.

Fracastoro (1478–1553)

The Renaissance physician and poet Fracastoro, from Verona, extended the concept of contagion by suggesting a theory about how contagious disease spreads. In his master work, *De contagione et contagiosis morbis et curatione,* which was published in 1546, he described many diseases, such as plague, typhus, and syphilis, that were transmitted from person to person, and he suggested a theory that disease was spread through self-replicating particles. He postulated that these particles, which he called *seminaria,* or seeds, were too small to see and were specific for each disease. His theory was the forerunner of the germ theory, although he had no concept that the seminaria were alive. He suggested that seminaria could infest an environment and could spread disease by direct person-to-person contact, by indirect contact with articles he called *fomites* (a term still in use for contaminated articles), and by transmission at a distance (eg, through air or water).

He thought that atmospheric conditions could influence the ability of the seminaria to spread and cause epidemics, an idea that traces back to Hippocrates, although for Hippocrates the causal role of airs and waters was more definitive.

Despite the importance of *De contagione*, Fracastoro is best known for his earlier epic poem *Syphilis sive morbus Gallicus* ("Syphilis or the French Disease"). This Latin verse was one of the most celebrated poems of the Renaissance, but at the same time it was a clinical dissertation on this new disease that had spread rapidly through Europe in the early years of the 16th century. He discussed the known treatments for syphilis, which included mercury and a New World treatment called *guaiacum*. The poem includes a detailed discussion of Columbus' voyage of discovery, not because of Columbus' possible role in importing syphilis to Europe but because the source of guaiacum was the New World that Columbus had explored.

John Graunt (1620–1674)

A London haberdasher by trade, John Graunt is known as the world's first epidemiologist and demographer. His avocation as a scientist led him to focus on an available data resource, the weekly *Bills of Mortality*, which summarized data collected in the parishes of London and later throughout England, originally to monitor deaths from the plague. The list included baptisms and deaths by cause in each parish. The collection of these data had begun late in the 15th century (uninterruptedly from 1603), and continued until the 19th century, when they were superseded by a more formal registration system (see section on William Farr). Graunt realized that much could be learned from a compilation of the data, thus setting a tradition for epidemiologists who still seek to use already collected data for epidemiologic research.

He published only one work based on his research, but this small volume contained an impressive number of original findings. The book had the unwieldy title *Natural and Political Observations Mentioned in a Following Index, and Made Upon the Bills of Mortality*. It included the first observation that more boys are born than girls. It presented the first actuarial table. It included the first reports of time trends for various diseases, adjusting for population size. He noted that some diseases changed in frequency because of changes in disease classification, as opposed to natural phenomena. He gave the first estimate of the population of London, demonstrating that it was growing rapidly through immigration rather than by an increase in births. He offered evidence to refute the theory that plague epidemics accompany the crowning of a new king. And he noted that physicians see more female patients than male patients.

Although Graunt was an amateur scientist, he had a natural appreciation for the fine points of epidemiology. For example, he was meticulous in describing the method of data collection for the *Bills of Mortality*[5]:

> These *Bills* were Printed and published, not onely every week on *Thursdays*, but also a general Accompt of the whole Year given in, upon the *Thursday* before *Christmas Day*: which said general Accompts have been presented in the several manners following, *viz.* from the Year 1603, to the Year 1624, *inclusive*…
>
> We have hitherto described the several steps, whereby the *Bills of Mortality* are come up to their present state; we come next to shew how they are made, and composed, which is in this manner, *viz.*

> When any one dies, then, either by tolling, or ringing of a Bell, or by bespeaking of a Grave of the *Sexton*, the same is known to the *Searchers*, corresponding with the said *Sexton*.
>
> The *Searchers* hereupon (who are antient Matrons, sworn to their Office) repair to the place, where the dead Corps lies, and by view of the same, and by other enquiries, they examine by what *Disease*, or *Casualty* the Corps died. Hereupon they make their Report to the *Parish-Clerk*, and he, every *Tuesday* night, carries in an Accompt of all the *Burials*, and *Christnings*, hapning that Week, to the *Clerk* of the *Hall*. On *Wednesday* the general Accompt is made up, and Printed, and on *Thursdays* published and dispersed to the several Families, who will pay four shillings *per Annum* for them.

The classification of deaths in the *Bills of Mortality* would seem strange to the modern eye. For example, categories of death included "Suddenly," "Killed by Several Accidents," and "Found Dead in the Streets." Nevertheless, Graunt's concern about possible misclassification of the events that were tabulated in the *Bills of Mortality* was very modern. He suspected, for example, that plague deaths were underascertained, because in years that had more plague deaths, there were also more deaths from other causes. He inferred that about 20% of plague deaths were mistakenly recorded as deaths attributable to other causes and on that basis produced a refined estimate of mortality from the plague. This was the first recorded example of correction for misclassification error. He also considered the reasons for misclassification of deaths. The matrons who served as searchers could recognize a corpse as well as anyone and could easily establish whether a death should be attributed to hanging, leprosy, drowning, or other obvious causes. But some causes of death were more difficult to ascertain. With respect to consumption (ie, tuberculosis), he noted[5] that

> … all dying thereof so emaciated and lean (their *Ulcers* disappearing upon Death) that the Old-women *Searchers* after the mist of a cup of *Ale*, and the bribe of a two-groat fee, instead of one, given them, cannot tell whether this emaciation, or leanness were from a *Phthisis*, or from an *Hectic Fever*, *Atrophy*, &c or from an Infection of the *Spermatick* parts.

Graunt's work provides several lessons that still apply to good epidemiologic work.[6] Among these are the following:

1. He was succinct. His short work contained many original findings and yet had an ample description of methods and results.
2. He provided clear explanations for his reasoning, for example with the calculations showing the underestimation of plague deaths.
3. He subjected his theories and novel findings to repeated tests. For example, he estimated the population of London to be 384,000, but he derived this figure using five different approaches.
4. He invited readers to criticize his work, to "correct my *Positions*, and raise others of their own: For herein I have, like a silly Scholeboy, coming to say my Lesson to the World (that Peevish, and Tetchie Master) brought a bundle of Rods wherewith to be whipt, for every mistake I have committed."[5] Such humility is less common nowadays.
5. He described how he had revised his opinions in the light of new data.
6. He relied on estimation, a quantitative approach to his subject matter rather than a qualitative approach, such as the presence or absence of statistical significance, which infests much of today's research (see Chapter 8).

Bernardino Ramazzini (1633–1714)

Ramazzini was born in Carpi, Italy, in the year when Galileo was interrogated by the Inquisition. He was a polymath who became a pioneer in epidemiology and especially in the field of occupational medicine. As a physician, he was noted as an early proponent for the use of cinchona bark (a source of quinine) to treat malaria. His crowning achievement, however, was the text *De morbis artificum diatriba* ("On Artificially Caused Disease"), the first comprehensive work on occupational diseases and industrial hygiene. Published in 1700, it described risks related to dozens of occupational hazards. He counseled physicians to inquire about the work activities and workplace exposures of their patients[7]:

> "When you come to a patient's house, you should ask him what sort of pains he has, what caused them, how many days he has been ill, whether the bowels are working and what sort of food he eats." So says Hippocrates in his work *Affections*. I may venture to add one more question: what occupation does he follow?

He gave detailed descriptions of specific occupational diseases, including a review of existing knowledge and suggestions for prevention and treatment. He was also concerned with the physical demands of specific occupations and with repetitive physical tasks. His interest in ergonomics can be seen in this excerpt[8]:

> Nowadays women sit to weave, but in such a posture that they somehow look as though they were standing. This kind of work is certainly very fatiguing, for the whole body is tasked, both hands, arms, feet, and back, so that every part of the body at once shares in the work. . . . Now an occupation so fatiguing naturally has its drawbacks, especially for women, for if pregnant they easily miscarry and expel the fetus prematurely and in consequence incur many ailments later on. It follows that women weavers, I mean those who are engaged wholly in this occupation, ought to be particularly healthy and robust, otherwise they break down from overwork and as they get on in years are compelled to abandon this trade. . . . Therefore in work so taxing moderation would be the best safeguard against these maladies, for men and women alike; for the common maxim "Nothing to excess" is one that I excessively approve.

In 1982, an international group of public-health scientists founded the Collegium Ramazzini, headquartered in Carpi. The Collegium was founded to advance the study of environmental and occupational disease and, according to its bylaws, to be "a bridge between the world of scientific discovery and those social and political centers which must act on these discoveries to conserve life and to prevent disease."[9]

William Farr (1807–1883)

Farr was the first Compiler of Abstracts at the General Register Office in England. The General Register Office was created by an act of Parliament in 1836 to record the civil registration of births, deaths, and marriages in England and Wales. Around that time, information on cause of death was collected along with age and occupation for all deaths. Farr's appointment began in 1839, and he remained in the General Register Office for 40 years, where he continued the tradition that began with John Graunt of using routinely collected vital statistics to learn about disease occurrence.

Farr occupied himself with many of the same issues that Graunt addressed. He studied demographic issues of population size, sex ratio at birth, fecundity, and time trends. He tested the theories that had been proposed by Malthus on population growth. He constructed actuarial tables, examined infant mortality, and, following in the footsteps of Ramazzini, studied the relation between specific occupations and mortality. He was famously engaged in a series of analyses of the cholera epidemics that struck London in the mid–19th century, in which he determined that local and comparatively small changes in altitude were strongly related to mortality from cholera. At first, he attributed the effect of altitude to atmospheric influences, consistent with the then-popular theory that foul air was the medium by which cholera was spread. In this, he was in conflict with John Snow, who attributed the spread of cholera to contaminated water (see next section). But eventually his own data persuaded Farr to change his mind, and he became a champion of Snow's theory that sewage contaminated with excreta from cholera victims perpetuated the epidemics.

Farr's work defined the field of vital statistics and had a lasting influence on epidemiology and public health. He was a dispassionate scientist, but he also saw the harsh social burden of industrialization and was an ardent reformer as well as a member of the sanitary movement. On the fundamental role of epidemiology in public health, Farr remarked, "Diseases are more easily prevented than cured and the first step to their prevention is the discovery of their exciting causes."[10]

John Snow (1813–1858)

Esteemed for his meticulous research on cholera as well as his pioneering work in the use of anesthetic gases, John Snow is considered the founding father of both epidemiology and anesthesiology. His renown as an anesthetist stems from his study of anesthetic gases, the design of inhalers to control the flow of these gases, and the sensational administration of chloroform to Queen Victoria during delivery. His renown as an epidemiologist stems from his investigation of the London cholera epidemics in the mid–19th century and his famous persuasion of the Board of Guardians of Saint James' Parish to remove the pump handle from the Broad Street pump in Golden Square, so as to contain the epidemic of cholera that raged there in the summer of 1854.

The popularization of the dramatic vignette of the pump handle has Snow's intervention stopping the epidemic cold. As I describe in Chapter 4, however, the epidemic was already waning when the pump handle was removed, and it is not clear that any cholera cases were prevented by its removal. On the other hand, it is possible that removal of the handle may have prevented a second outbreak. The likely index case of the outbreak was an infant who lived near the pump and whose mother soaked the soiled diapers of the infant in pails and dumped them into a cesspool near the pump. Later investigation showed that the brickwork around the pump was defective and water from the cesspool could enter the pump. The infant's father developed cholera the day that the pump handle was removed and therefore might have ignited a second epidemic had the pump still been in operation.[11,12]

The real story was not so much the removal of the pump handle but Snow's meticulous investigations that demonstrated the connection between consuming fouled water and risk of cholera. By 1854, Snow had already been pursuing for several years the theory that cholera was transmitted by ingestion of an agent that was spread through fecal contamination of drinking water. He brought many lines of inquiry to bear on his research. At the time, the germ theory was not accepted—most experts still believed that miasmas were the means for spreading communicable diseases. Farr, for example, believed in miasmas, despite his own analyses that showed that altitude in the London area was a strong determinant of cholera risk. Altitude could affect the flow of bad air, so Farr and others thought that the altitude data supported the miasma theory. But altitude is also related to the flow of water, with individuals living at lower altitudes much more likely to be drinking water containing the sewage of those at higher elevations. Snow noted that inhalation of bad air would make more sense if the cholera patient developed a respiratory disease, but he saw that clinically the disease begins as a gastrointestinal affliction. Snow marshaled a wide array of facts to make his case. For example, he noted that miners experienced a greater risk of cholera than those in any other occupation. Miners also spent long hours underground, with no access to privies, and carried food along that they ate with bare hands that they were unable to wash adequately.

Although the incident with the Broad Street pump is iconic, Snow's most important contribution to epidemiology was his groundbreaking work on the relation between water supply and cholera conducted in certain neighborhoods of London. The key study was one in which water pipes from two competing companies, one carrying clean water and the other water that was contaminated with London sewage, supplied the same neighborhood. Most residents did not even know which company supplied their apartment, because landlords typically contracted to provide water. This situation, elegantly described by Snow as being better than any experiment that could have been devised, became known as a *natural experiment*. It is described in greater detail in Chapters 4 and 5. Snow is celebrated today by epidemiologists even more for being a pioneer of epidemiologic study design than for his seminal work on the mode of cholera transmission.

Ignasz Semmelweis (1818–1865)

Today recognized as a pioneering epidemiologist and revered as a hero in his native Hungary, in his time Semmelweis was an outcast who was widely spurned by the medical profession. He was born in Budapest and studied medicine and practiced obstetrics in Vienna at a time when puerperal fever raged throughout the maternity hospitals of Europe with no satisfactory explanation. The Vienna Maternity Hospital was the largest hospital of its kind in the world. Patients were admitted on alternate days to each of its two obstetrics clinics. Medical students were taught in the first clinic and midwives in the second. During the period 1840–1846, the risk of maternal death during delivery was about 10% for women admitted to the first clinic but less than 4% in the second clinic, where the midwives trained. Almost all maternal deaths were from puerperal fever. The mortality

rate was so high in the first clinic that women who could afford to do so gave birth at home rather than risk being assigned to the first clinic.

Semmelweis investigated a theory that doctors themselves were the source of disease for their patients. In the first clinic, when a woman died with puerperal fever, students and their teachers would conduct an autopsy, and from there they would proceed back to the clinic to examine other patients. Semmelweis noted that when he left Vienna for a sabbatical, puerperal fever deaths in the first clinic declined, but they rose again after his return. He also noted that a colleague who nicked himself with a scalpel during an autopsy died of a disease that resembled puerperal fever. From these and other observations, Semmelweis theorized that the source of disease, the "morbid matter," was being spread by doctors and students from cadavers in the morgue to patients on the ward. Semmelweis instituted a policy of washing hands in a chlorinated lime disinfectant before returning to the wards, and the risk of puerperal fever in the first clinic was reduced to the same level as in the second clinic.

Unfortunately, the germ theory of disease was not yet accepted. Most physicians believed that miasmas, or bad air, were the source of communicable disease. Semmelweis was also handicapped by his own personality. He was slow to publish, loath to accept criticism, quick to anger, and dogmatic.[13] Consequently, he was not effective in persuading others to take his theory seriously, despite the remarkable evidence that he assembled. After a series of professional setbacks, he died young and ignominiously while hospitalized in a mental institution, and his work was forgotten. Eventually, the English surgeon Joseph Lister, influenced by Pasteur, introduced antiseptic technique to surgery practice, and from there it was carried over to maternity clinics. All this happened without any influence from Semmelweis, whose reputation rose only after his work was rediscovered.

Florence Nightingale (1820–1910)

Nightingale is best known for her contribution to nursing, but she also was an accomplished epidemiologist and statistician who made major contributions to the field of public health.[14,15] Born to a wealthy family, she entered the nursing profession as a calling to work for the public good, facing parental opposition and societal roadblocks. Her early professional efforts were aimed at caring for people in poverty, helping to improve medical care available to the poor, and working to reform the Poor Law of the United Kingdom. In 1854, she was asked by the British Secretary of War to go to the British front where the Crimean war was being waged. She went with a staff of trained volunteer nurses and discovered that poor nutrition and hygiene, lack of medicines, and indifference were responsible for more deaths among soldiers than battlefield injuries. She instigated improvements such as better sewage and ventilation. Death rates soon dropped in the military field hospital.

Nightingale's efforts in the Crimea brought her fame and with it a mandate to continue her efforts in public health. She wrote a treatise on health conditions in the Crimea, for which she invented a diagram now known as a *polar-area diagram* or *coxcomb* (Fig. 2–1). It plotted monthly deaths attributable to preventable and unpreventable causes (deaths were proportional to area, not to radius); the

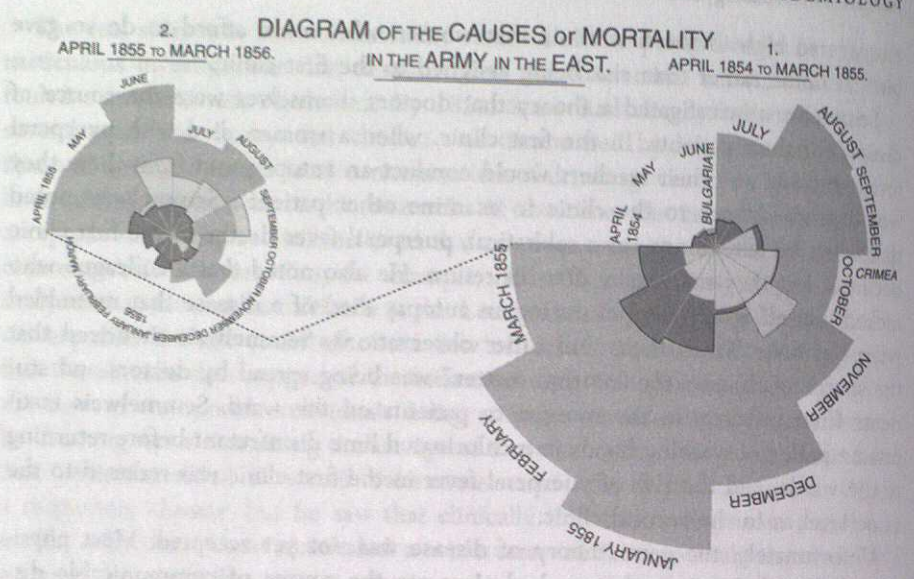

Figure 2–1. Diagram of mortality by cause during the Crimean war, by Florence Nightingale, 1858. Each wedge represents one month. The area of the outer wedges is proportional to death from "preventable or mitigable zymotic diseases," the area of the white wedges represents deaths from wounds, and the area from the dark shaded wedges represents deaths from all other causes. The graph shows that sickness took a much greater toll than injuries, and deaths were much greater during the first year than the second.

data revealed that most deaths were preventable and showed clear time trends. Back in England, she noted that the poor nutrition and hygiene she had found at the front also afflicted military men stationed at home. She wrote a report that sparked an overhaul of medical care of soldiers and led to the establishment of a medical school for the army. Later, she turned her attention to sanitation in rural India and became a voice for improvement of public health throughout India. Her work was characteristically meticulous, grounded in carefully collected data, and presented in compelling graphics.

Janet Lane-Claypon (1877–1967)

Another female British pioneer in epidemiology, who has been largely overlooked despite her seminal contribution, is the inventor of the modern case-control study. Lane-Claypon, like Nightingale, was a public-health leader who became an innovator in epidemiology in pursuit of her public-health objectives. In 1912, she published the results of a study that examined breast feeding versus bottle feeding in relation to weight gain during infancy and secondarily compared the effect of boiled versus raw cow's milk among the bottle-fed infants. This study has been described as the first retrospective cohort study,[16] although one could argue that the earlier studies of Snow on cholera or of Semmelweiss on puerperal fever were retrospective cohort studies. Regardless of priority, Lane-Claypon's work was ahead of its time, not only for the basic study design but also for

her attention to both systematic and random error. She excluded sick infants to prevent confounding, and she wrote about possible confounding by social class, a factor that was not controlled and could affect the data she reported.

Her seminal work came in 1926 with the publication of results from the first modern case-control study, which was aimed at evaluating risk factors for breast cancer.[17] She selected 500 cases and an equal number of controls from hospitals in London and Glasgow. Exposure information was obtained by interview using a questionnaire designed for the study. Her analyses laid the groundwork for breast cancer epidemiology, establishing the relation of age at first pregnancy, age at menopause, number of children, and surgically induced menopause to breast cancer risk. Although it was not recognized at the time, this study introduced a novel study design that has since become a hallmark of modern epidemiology. Lane-Claypon seemed well aware, however, of the strengths and weaknesses of the method. She discussed the possibility of recall bias, long considered an important source of error in case-control studies that rely on interviews to gather data. She also acknowledged possible problems stemming from the facts that cases were survivors and that cases and controls were drawn from a set of hospitals. Many decades before the principles of epidemiologic study design were codified and discussed, Lane-Claypon had the insight to describe many of the important issues in both cohort and case-control studies.[18,19]

Wade Hampton Frost (1880–1938)

Wade Hampton Frost was the first professor of epidemiology in the first university department of epidemiology, at the Johns Hopkins School of Hygiene and Public Health (now known as the Johns Hopkins Bloomberg School of Public Health).[20] After graduating with a degree in medicine from the University of Virginia in 1903, Frost chose to pursue a career in public health. (One speculative view suggests that Frost chose a career in public health because he contracted tuberculosis while a student and was consequently advised to avoid working in general practice.) At the start of the 20th century, the germ theory of disease was still new, and the control of disease spread by infectious agents was in its infancy. Soon after graduation, Frost was called to New Orleans to investigate an outbreak of yellow fever. The role of the mosquito vector, *Aedes aegypti*, in the transmission of yellow fever had only recently been established. Frost and a team of colleagues spent weeks eliminating breeding spots for the mosquito and eventually stopped the outbreak after 459 people had died. It marked the first time that a yellow fever outbreak in the United States was halted before the arrival of winter, and it was also the last epidemic of yellow fever in the United States.

Frost's subsequent work and teaching were influenced by the classic investigations by William Budd on typhoid fever and by Snow on cholera. His later work focused on typhoid, poliomyelitis, meningococcal meningitis, tuberculosis, and influenza. His studies on polio led to the understanding that paralytic cases constitute only a small proportion of those infected by the virus and that childhood infection with the virus confers lasting immunity. His influenza research documented the spread of the global pandemic of 1918–1919, showed that the

effects of influenza epidemics could be tracked using death rates for pneumonia, and revealed that the case-fatality rate for influenza during the pandemic had a bimodal age distribution, hitting hard among young adults and the very old. His work on tuberculosis was an early and elegant demonstration of the power of analyzing mortality rates by birth cohort, a method first used a decade earlier by the Norwegian physician Kristian Andvord. Frost is also known for his collaboration with Lowell Reed in developing the elegant Reed-Frost mathematical model for infectious disease transmission and the development of herd immunity (see Chapter 6).

ALSO NOTEWORTHY

The contributions of these pioneers highlight only a few of the important milestones in the history of epidemiology and public health. Many others could easily have been mentioned—William Budd, Edward Goldberger, Major Greenwood, Joseph Jenner, James Lind, Pierre Louis, Peter Panum, Geoffrey Rose, and Edgar Sydenstricker, to suggest a few. By the mid–20th century, epidemiology research and teaching had expanded rapidly. It had also extended its purview to include an ever-widening range of disease foci, such as psychiatric illness, violent behavior, and obesity. Leaders such as Austin Bradford-Hill, Richard Doll, Brian MacMahon, Abraham Lilienfeld, John Cassel, and many others nurtured the growth of epidemiology during this period, influencing others through their teaching but even more so by the quality of their work. The surge in epidemiologic activity has continued through the 20th century and into the present.

QUESTIONS

1. Although epidemiologic research is greatly facilitated by computers, its conduct is based on enumerating events in populations and is not especially dependent on technology. The ideas needed to conduct epidemiologic research could have been developed and applied long ago. Speculate as to why epidemiologic research did not become commonplace until the middle of the 20th century.

2. Both Snow and Semmelweiss conducted research that showed how one might prevent infectious disease before the germ theory was accepted. Yet the existence of microorganisms had been known for almost 200 years, since van Leeuwenhoek first described what he saw under a microscope. Speculate on the nature of social and scientific thinking of the time that hindered the scientists of the 19th century from accepting the role of microorganisms in disease.

3. Graunt and Farr advanced the field of epidemiology by studying public records and vital statistics, setting examples for generations of epidemiologists. Others, such as Snow and Semmelweiss, devised studies to collect their own data, which required considerable time and effort but was possible as an independent effort. Today, collection of ad hoc epidemiologic data is expensive and often requires collaborative efforts and substantial funding that must be approved in a peer-review process. It appears that peer review, had it been necessary, might have been an obstacle to Snow and Semmelweiss. Give arguments listing the strengths and weaknesses of a system of peer review for funding of scientific research.

ADDITIONAL READING

Buck C, Llopis A, Nájera E, Terris M. *The Challenge of Epidemiology: Issues and Selected Readings.* Pan American Health Organization Scientific Publication No. 505. Geneva: World Health Organization, 1988.

REFERENCES

1. Rosen G. *History of Public Health.* Expanded edition. Baltimore: Johns Hopkins University Press; 1993.
2. Svoboda T, Henry B, Shulman L, et al. Public health measures to control the spread of the severe acute respiratory syndrome during the outbreak in Toronto. *N Engl J Med.* 2004;350:2352–2361.
3. Hippocrates. *On Airs, Waters, and Places.* Adams F, trans. http://classics.mit.edu/Hippocrates/airwatpl.html. Accessed September 30, 2011.
4. Horne CF, ed. *The Sacred Books and Early Literature of the East.* Vol VI: Medieval Arabia. New York: Parke, Austin, & Lipscomb; 1917:90–91.
5. Graunt J. *Natural and Political Observations Mentioned in a Following Index, and Made Upon the Bills of Mortality.* Facsimile edition. New York: Arno Press; 1975.
6. Rothman KJ. Lessons from John Graunt. *Lancet.* 1996;347:37–39.
7. Ramazzini B. Diseases of workers. *The Classics of Medicine Library.* Special edition. Chicago: University of Chicago Press; 1983.
8. Franco G, Franco F. Bernardino Ramazzini: The father of occupational medicine. *Am J Publ Health.* 2001;91:1380–1382.
9. Bylaws of the Collegium Ramazzini. http://www.collegiumramazzini.org/download/Bylaws.pdf. Accessed September 30, 2011.
10. Farr W. Letter to the Registrar General. In: *First Annual Report of the Registrar General.* London: Her Majesty's Stationery Office; 1839.
11. Vinten-Johansen P, Brody H, Paneth N, et al. *Cholera, Chloroform, and the Science of Medicine: A Life of John Snow.* New York: Oxford University Press; 2003.
12. Rothman KJ. My interview with John Snow. *Epidemiology.* 2004;15:640–644.
13. Nuland SB. *The Doctors' Plague: Germs, Childbed Fever, and the Strange Story of Ignac Semmelweis.* New York: WW Norton; 2003.
14. Winkelstein W. Florence Nightingale: founder of modern nursing and hospital epidemiology. *Epidemiology.* 2009;20:311.
15. Cook E. *The Life of Florence Nightingale.* Vol 2. London: Macmillan; 1913.
16. Winkelstein W Jr. Vignettes of the history of epidemiology: three firsts by Janet Elizabeth Lane-Claypon. *Am J Epidemiol.* 2004;160:97–101.
17. Lane-Claypon JE. *A Further Report on Cancer of the Breast, With Special Reference to Its Associated Antecedent Conditions.* Reports on Public Health and Medical

Subjects No. 32, Ministry of Health. London: His Majesty's Stationery Office; 1926.
18. Winkelstein W Jr. Janet Lane-Claypon, a forgotten epidemiologic pioneer. *Epidemiology.* 2006;17:705.
19. Paneth N, Susser E, Susser M. Origins and early development of the case-control study. Part 2: The case-control study from Lane-Claypon to 1950. *Social Prev Med.* 2002;47:359–365.
20. Daniel TM. *Wade Hampton Frost, Pioneer Epidemiologist 1880–1938.* Rochester, NY: University of Rochester Press; 2004.

3

What Is Causation?

The acquired wisdom that certain conditions or events bring about other conditions or events is an important survival trait. Consider an infant whose first experiences are a jumble of sensations that include hunger, thirst, color, light, heat, cold, and many other stimuli. Gradually, the infant begins to perceive patterns in the jumble and to anticipate connections between actions such as crying and effects such as being fed. Eventually, the infant assembles an inventory of associated perceptions. Along with this growing appreciation for specific causal relations comes the general idea that some events or conditions can be considered causes of other events or conditions.

Thus, our first appreciation of the concept of causation is based on our own observations. These observations typically involve causes with effects that are immediately apparent. For example, changing the position of a light switch on the wall has the instant effect of causing the light to go on or off. There is, however, more to the causal mechanism for getting the light to shine than turning the light switch to the on position. If the electric lines to the building are down because of a storm, turning on the switch will have no effect. If the bulb is burned out, manipulating the switch also will have no effect. One cause of the light going on is having the switch in the proper place, but along with it we must include a supply of power to the circuit, a working bulb, and intact wiring. When all other factors are in place, turning the switch will cause the light to go on, but if one or more of the other factors is not playing its causal role, the light will not go on when the switch is turned. There is a tendency to consider the switch as the unique cause of turning on the light, but we can define a more intricate causal mechanism in which the switch is one component of several. The tendency to identify the switch as the unique cause stems from its usual role as the final factor that acts in the causal mechanism. The wiring can be considered part of the causal mechanism, but after it is installed, it seldom warrants further attention. The switch is typically the only part of the mechanism that needs to be activated to turn on the light. The effect usually occurs immediately after turning the switch, and as a result, we tend to identify the switch as a unique cause. The

Subjects No. 32, Ministry of Health. London: His Majesty's Stationery Office; 1926.
18. Winkelstein W Jr. Janet Lane-Claypon, a forgotten epidemiologic pioneer. *Epidemiology.* 2006;17:705.
19. Paneth N, Susser E, Susser M. Origins and early development of the case-control study. Part 2: The case-control study from Lane-Claypon to 1950. *Social Prev Med.* 2002;47:359–365.
20. Daniel TM. *Wade Hampton Frost, Pioneer Epidemiologist 1880–1938.* Rochester, NY: University of Rochester Press; 2004.

3

What Is Causation?

The acquired wisdom that certain conditions or events bring about other conditions or events is an important survival trait. Consider an infant whose first experiences are a jumble of sensations that include hunger, thirst, color, light, heat, cold, and many other stimuli. Gradually, the infant begins to perceive patterns in the jumble and to anticipate connections between actions such as crying and effects such as being fed. Eventually, the infant assembles an inventory of associated perceptions. Along with this growing appreciation for specific causal relations comes the general idea that some events or conditions can be considered causes of other events or conditions.

Thus, our first appreciation of the concept of causation is based on our own observations. These observations typically involve causes with effects that are immediately apparent. For example, changing the position of a light switch on the wall has the instant effect of causing the light to go on or off. There is, however, more to the causal mechanism for getting the light to shine than turning the light switch to the on position. If the electric lines to the building are down because of a storm, turning on the switch will have no effect. If the bulb is burned out, manipulating the switch also will have no effect. One cause of the light going on is having the switch in the proper place, but along with it we must include a supply of power to the circuit, a working bulb, and intact wiring. When all other factors are in place, turning the switch will cause the light to go on, but if one or more of the other factors is not playing its causal role, the light will not go on when the switch is turned. There is a tendency to consider the switch as the unique cause of turning on the light, but we can define a more intricate causal mechanism in which the switch is one component of several. The tendency to identify the switch as the unique cause stems from its usual role as the final factor that acts in the causal mechanism. The wiring can be considered part of the causal mechanism, but after it is installed, it seldom warrants further attention. The switch is typically the only part of the mechanism that needs to be activated to turn on the light. The effect usually occurs immediately after turning the switch, and as a result, we tend to identify the switch as a unique cause. The