1

The Approach and Evolution of Epidemiology

LEARNING OBJECTIVES

By the end of this chapter the reader will be able to:

- Define and discuss the goals of public health.
- Distinguish between basic, clinical, and public health research.
- Define epidemiology and explain its objectives.
- Discuss the key components of epidemiology (population and frequency, distribution, determinants, and control of disease).
- Discuss important figures in the history of epidemiology including John Graunt, James Lind, William Farr, and John Snow.
- Discuss important modern studies including the Streptomycin Tuberculosis Trial, Doll and Hill's studies on smoking and lung cancer, and the Framingham Study.
- Discuss the current activities and challenges of modern epidemiologists.

Introduction

Most people do not know what epidemiology is or how it contributes to the health of our society. This fact is somewhat paradoxical, given that epidemiology pervades our lives. Consider, for example, the following statements that have recently made headline news:

- Mammographic screening for breast cancer does not benefit women aged 40 to 49 years.

- Cellular telephone users who talk on the phone while driving are more likely than other drivers to become involved in a car accident.
- Saint John's wort, a popular alternative medicine, is an ineffective treatment for severe depression.
- Disinfection byproducts in drinking water increase a pregnant woman's risk of miscarriage.
- Premarital screening for human immunodeficiency virus (HIV) infection is ineffective in reducing viral transmission.

The breadth and importance of these topics indicate that epidemiology directly affects the daily lives of most people. It affects the way that individuals make personal decisions about their lives and the way that the government, public health agencies, and medical organizations make policy decisions that affect how we live. For example, the results of epidemiologic studies described by the headlines might prompt a person to request a traditional medication for her depression, to wait until age 50 years to begin regular mammograms, or to install a water treatment device on her kitchen faucet. It might prompt a gynecologist to tailor recommendations about mammographic screening to suit the patient's preferences, a water utility to change its methods for disinfecting drinking water, cell phone manufacturers to design safer phones, or state legislators to vote against laws mandating premarital HIV screening of heterosexual couples and to vote for laws banning cell phone use by drivers.

This chapter helps the reader understand what epidemiology is and how it contributes to important issues affecting the public's health. In particular, it describes the definition, approach, and goals of epidemiology, as well as key aspects of its historical development, current state, and future challenges.

Definition and Goals of Public Health

Public health is a multidisciplinary field whose goal is to promote the health of the population through organized community efforts.[1(pp3–15)] In contrast to medicine, which focuses mainly on treating illness in separate individuals, public health focuses on preventing illness in the community. Key public health activities include assessing the health status of the population, diagnosing its problems, searching for the causes of those problems, and designing solutions for them. The solutions usually involve community-level interventions that control or prevent the cause of the problem. For example, public health interventions include establishing educational programs to discourage teenagers from smoking, implementing screening programs for the early detection of cancer, and passing laws that require automobile drivers and passengers to wear seat belts.

Unfortunately, public health achievements are difficult to recognize because it is hard to identify people who have been spared illness.[1(pp6–7)] For this reason, the field of public health has received less attention and fewer resources than the field of medicine has received. Nevertheless, public health has had a greater impact than medicine on the health of populations. For example, since the turn of the 20th century, the average life expectancy of Americans has increased by almost 30 years from 47.3 to 76.7 years.[2] Of this increase, 25 years can be attributed to improvements in public health and only 5 years can be attributed to improvements in the medical care system.[3] Public health achievements that account for improvements in health and life expectancy include the routine use of vaccinations for infectious diseases, improvements in motor vehicle and workplace safety, control of infectious diseases through improved sanitation and clean water, modification of risk factors for coronary heart disease and stroke (such as smoking cessation and blood pressure control), safer foods from decreased microbial contamination, improved access to family planning and contraceptive services, and the acknowledgment of tobacco as a health hazard and the ensuing antismoking campaigns.[4]

The public health system's activities in research, education, and program implementation have made these accomplishments possible. In the United States this system includes federal agencies such as the Centers for Disease Control and Prevention, state and local government agencies, nongovernmental organizations such as Mothers Against Drunk Driving, and academic institutions such as schools of public health. This complex array of institutions has achieved success through political action and gains in scientific knowledge.[1(pp13–14)] Politics enters the public health process when agencies advocate for resources, develop policies and plans to improve a community's health, and work to ensure that services needed for the protection of public health are available to all. Political action is necessary because the government usually has the responsibility for developing the activities required to protect public health.

Sources of Scientific Knowledge in Public Health

The scientific basis of public health activities mainly comes from (1) the basic sciences such as pathology and toxicology, (2) the clinical or medical sciences such as internal medicine and pediatrics, and (3) the public health sciences such as epidemiology, environmental health science, and health education and behavioral science. Research in these three areas provides complementary pieces of a puzzle that, when properly assembled, provide the scientific foundation for public health action. Other fields such as engineering and economics also contribute to public health. The three main areas approach research questions from different yet complementary viewpoints, and each field has its own particular strengths and weaknesses.

Basic scientists such as toxicologists study disease in a laboratory setting by conducting experiments on cells, tissues, and animals. The focus of this research is often on the disease mechanism or process. Because basic scientists conduct their studies in a controlled laboratory environment, they can regulate all important aspects of the experimental conditions. For example, a laboratory experiment testing the toxicity of a chemical is conducted on genetically similar animals who live in the same physical environment, eat the same diet, and follow the same daily schedule.[5(pp157–237)] Animals are assigned (usually by chance) to either the test group or the control group. Using identical routes of administration, researchers give the chemical under investigation to the test group and an inert chemical to the control group. Thus, the only difference between the two groups is the dissimilar chemical deliberately introduced by the investigator. This type of research provides valuable information on the disease process that cannot be obtained in any other way. However, the results are often difficult to extrapolate to real-life situations involving humans because of differences in susceptibility between species, and differences in the exposure level between laboratory experiments and real-life settings. In general, humans are exposed to much lower doses than those used in laboratory experiments.

Clinical scientists focus their research questions mainly on disease diagnosis, treatment, and prognosis in individual patients. For example, they try to determine if a diagnostic method is accurate or if a treatment is effective. Although clinicians are also involved in disease prevention, this activity has historically taken a back seat to disease diagnosis and treatment. As a consequence, clinical research studies are usually based on people who come to a medical care facility such as a hospital or clinic. Unfortunately, these people are often unrepresentative of the full spectrum of disease in the population at large because many sick people never come to the attention of health care providers.

Clinical scientists contribute to scientific knowledge in several important ways. First, they are usually the first to identify new diseases, the adverse effects of new exposures, and new links between an exposure and a disease. This information is typically published in case reports. For example, the epidemic of acquired immunodeficiency syndrome (AIDS) officially began in the United States in 1981 when clinicians reported several cases of *Pneumocystis carinii* pneumonia and Kaposi's sarcoma (a rare cancer of the blood vessels) among previously healthy, young homosexual men living in New York and California.[6,7] These cases were notable because

Pneumocystis carinii pneumonia had previously occurred only among immunocompromised individuals and Kaposi's sarcoma had occurred mainly among elderly men. We now know that these case reports described symptoms of a new disease that would eventually be called AIDS. Despite their simplicity, case reports provide important clues regarding the causes, prevention, and cures for a disease. In addition, they are often used to justify conducting more sophisticated and expensive studies.

Clinical scientists also contribute to scientific knowledge by recording treatment and response information in their patients' medical records. This information often becomes an indispensable source of research data for clinical and epidemiologic studies. For example, it would have been impossible to determine the risk of breast cancer following fluoroscopic x-ray exposure without patient treatment records from the 1930s through 1950s.[8] Investigators used these records to identify the subjects for the study and gather detailed information about subjects' radiation doses.

Public health scientists study ways to prevent disease and promote health in the population at large. Public health research differs from clinical research in two important ways. First, it focuses mainly on disease prevention rather than disease treatment. Second, the units of concern are groups of people living in the community rather than separate individuals visiting a health care facility. For example, a public health research project called "Lead Free Kids" determined the impact of removing lead-contaminated soil on children's blood lead levels.[9] About 150 healthy, community-dwelling children who were at risk for lead poisoning were targeted for this lead-poisoning prevention project.

The main differences between the three branches of scientific inquiry are summarized in Table 1–1. Although this is a useful way to classify the branches of scientific research, the distinctions between these areas have become blurred. For example, epidemiologic methods are currently being applied to clinical medicine in a field called "clinical epidemiology." In

TABLE 1–1 Main Differences Between Basic, Clinical, and Public Health Science Research

Characteristic	*Basic*	*Clinical*	*Public health*
What/who is studied	Cells, tissues, animals in laboratory settings	Sick patients who come to health care facilities	Populations or communities at large
Research goals	Understanding disease mechanisms and the effects of toxic substances	Improving diagnosis and treatment of disease	Prevention of disease, promotion of health
Examples	Toxicology, immunology	Internal medicine, pediatrics	Epidemiology, environmental health science

addition, newly developed areas of epidemiologic research such as molecular and genetic epidemiology include the basic sciences.

Definition and Objectives of Epidemiology

The term *epidemiology* is derived from the Greek words *epi,* which means "on or upon"; *demos,* which means "the common people"; and *logy,* which means "study."[10(pp484,599,1029)] Putting these pieces together yields the following definition of epidemiology: "the study of that which falls upon the common people." Epidemiology can also be defined as the "branch of medical science which treats epidemics."[11] The latter definition was developed by the London Epidemiological Society, which was formed in 1850 to determine the causes of cholera and other epidemic diseases and methods of preventing them.[12] Over the last century, many definitions of epidemiology have been set forth. Some early definitions reflect the field's initial focus on infectious diseases, and later ones reflect a broader scope encompassing all diseases.[12]

We define epidemiology as follows: *The study of the distribution and determinants of disease frequency in human populations and the application of this study to control health problems.*[13(p1),14(p55)] Our definition is a combination of a popular one coined by MacMahon and Pugh in 1970 and another described by Last in the third edition of the *Dictionary of Epidemiology.*[14(p55),15(p1)] Note that the term "disease" refers to a broad array of health-related states and events including diseases, injuries, disabilities, and death.

We prefer this hybrid definition because it describes both the scope and ultimate goal of epidemiology. In particular, the objectives of epidemiology are to: (1) study the natural course of disease from onset to resolution, (2) determine the extent of disease in a population, (3) identify patterns and trends in disease occurrence, (4) identify the causes of disease, and (5) evaluate the effectiveness of measures that prevent and treat disease. All of these activities contribute scientific knowledge for making sound policy decisions that protect public health.

Our definition of epidemiology has five key words or phrases: (1) population, (2) disease frequency, (3) disease distribution, (4) disease determinants, and (5) disease control. Each term is described in more detail in the following sections.

Population

Populations are at the heart of all epidemiologic activities because epidemiologists are concerned with disease occurrence in groups of people rather than in individuals. The term *population* refers to a group of people with a common characteristic such as place of residence, gender, age, or use of certain medical services. For example, people who reside in the city of Boston are members of a geographically defined population. Determining the size of the population in which disease occurs is as important as counting the cases of the disease, because it is only when the number of cases is related to the size of the population that we know the true frequency of disease. The size of the population is often determined by a census—that is, a complete count—of the population. Sources of these data range from the decennial census, in which the federal government attempts to count every person in the United States every 10 years, to computerized records from medical facilities that provide counts of patients who use the facilities.

Disease Frequency

This phrase refers to quantifying how often a disease arises in a population. Counting, which is a key activity of epidemiologists, includes three steps: (1) developing a definition of disease, (2) instituting a mechanism for counting cases of disease within a specified population, and (3) determining the size of that population.

Diseases must be clearly defined in order to determine accurately who should be counted. Usually disease definitions are based on a combination of physical and pathological examinations, diagnostic test results, and signs and symptoms. For example, a case definition of breast cancer might include findings of a palpable lump during a physical exam and mammographic and pathological evidence of malignant disease.

Currently available sources for identifying and counting cases of disease include hospital patient rosters, death certificates, special reporting systems such as registries of cancer and birth defects, and special surveys. For example, the National Health Interview Survey is a federally funded study that has collected data on the health status of the U.S. population since the 1950s. Its stated purpose is “to secure accurate and current statistical information on the amount, distribution, and effects of illness and disability in the United States.”[16] Data are collected on many diseases including cancer; diabetes; and cardiovascular, respiratory, gastrointestinal, and renal disease.

Disease Distribution

Disease distribution refers to the analysis of disease patterns according to the characteristics of person, place, and time. In other words, who is getting the disease, where is it occurring, and how is it changing over time? Variations in disease frequency by these three characteristics provide useful information that helps epidemiologists understand the health status of a population; formulate hypotheses about the determinants of a disease;

and plan, implement, and evaluate public health programs to control and prevent adverse health events.

Disease Determinants

Disease determinants are factors that bring about a change in a person's health—that is, factors that either cause a healthy individual to become sick or cause a sick person to recover.[14(p46)] Thus, determinants include both causal and preventive factors. Determinants also include individual, environmental, and societal characteristics. Individual determinants include a person's genetic makeup, gender, age, immunity level, diet, behaviors, and existing diseases. For example, the risk of breast cancer is increased among women who carry genetic alterations such as BRCA1 and BRCA2, are elderly, give birth at a late age, have a history of benign breast disease, or have a history of radiation exposure to the chest.[17]

Environmental and societal determinants are external to the individual, and thereby encompass a wide range of natural, social, and economic events and conditions. For example, the presence of infectious agents, reservoirs in which the organism multiplies, vectors that transport the agent, poor and crowded housing conditions, and political instability are environmental and social factors that cause many communicable diseases around the world.

Epidemiologic research involves generating and testing specific hypotheses about disease determinants. A hypothesis is defined as "a tentative explanation for an observation, phenomenon, or scientific problem that can be tested by further investigation."[10(p866)] Generating hypotheses is a process that involves creativity and imagination and usually includes observations on the frequency and distribution of disease in a population. Epidemiologists test hypotheses by making comparisons, usually within the context of a formal epidemiologic study. The goal of a study is to harvest valid and precise information about the determinants of disease in a particular population. As described in Chapter 6, epidemiologic research encompasses several types of study designs; each type of study merely represents a different way of harvesting the information.

Disease Control

Epidemiologists accomplish disease control through epidemiologic research, as described above, and through surveillance. The purpose of surveillance is to monitor aspects of disease occurrence that are pertinent to effective control.[18(p507)] For example, since 1985 the Centers for Disease Control and Prevention has collected information on the occurrence of AIDS across the United States.[19] In addition, after the development of serum HIV-antibody tests, states began reporting confirmed cases of HIV infection. For every case of AIDS and HIV infection, the surveillance system gathers data on the individual's demographic characteristics, exposure category (such as injecting drug users or men who have sex with men), and diagnosis date. These surveillance data are essential for formulating and evaluating programs to reduce the spread of AIDS.

Historical Development of Epidemiology

The historical development of epidemiology spans almost 400 years and is best described as slow and unsteady. Only since World War II has the field experienced a rapid expansion. The following sections, which are not meant to be a comprehensive history, highlight several historic figures and studies that made significant contributions to the evolution of epidemiologic thinking. These include John Graunt, who summarized the pattern of mortality in 17th-century London; James Lind, who used an experimental study to discover the cause and prevention of scurvy; William Farr, who pioneered a wide range of activities during the mid-19th century that are still used by modern epidemiologists; John Snow, who showed that cholera was transmitted by fecal contamination of drinking water; members of the Streptomycin in Tuberculosis Trials Committee, who conducted one of the first modern controlled clinical trials; Richard Doll and A. Bradford Hill, who conducted early research on smoking and lung cancer; and Thomas Dawber and William Kannel, who began the Framingham Study—one of the most influential and longest-running studies of heart disease in the world. It is clear that epidemiology has played an important role in the achievements of public health throughout its history.

John Graunt

The logical underpinnings for modern epidemiologic thinking evolved from the scientific revolution of the 17th century.[20(p23)] During this period, scientists believed that the behavior of the physical universe was orderly and could therefore be expressed in terms of mathematical relationships called "laws." These laws are generalized statements based on observations of the physical universe such as the time of day that the sun rises and sets. Some scientists believed that this line of thinking could be extended to the biological universe and reasoned that there must be "laws of mortality" that describe the patterns of disease and death. These scientists believed that the "laws of mortality" could be induced by observing the patterns of disease and death among humans.

John Graunt, a London tradesman and founding member of the Royal Society of London, was a pioneer in this regard. He became the first epidemiologist, statistician, and demographer when he summarized the Bills of Mortality for his 1662 publication *Natural and Political Observations*

Mentioned in a Following Index, and Made Upon the Bills of Mortality.[21] The Bills of Mortality were a weekly count of people who died that had been conducted by the parish clerks of London since 1592 because of concern about the plague. According to Graunt, the Bills were collected in the following manner:

> When any one dies, then, either by tolling, or ringing 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 enqueries, they examine by what Disease, or Casualty the Corps did die. 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.[21(pp25–26)]

This method of reporting deaths is not very different from the system used today in the United States. Like the "searchers" of John Graunt's time, modern physicians and medical examiners inspect the body and other evidence such as medical records to determine the official cause of death, which is recorded on the death certificate. The physician typically submits the certificate to the funeral director, who files it with the local office of vital records. From there, the certificate is transferred to the city, county, state, and federal agencies that compile death statistics. While 17th-century London families had to pay 4 shillings for the Bills of Mortality, these U.S. statistics are available free of charge.

Graunt drew many inferences about the patterns of fertility, morbidity, and mortality by tabulating the Bills of Mortality.[21] For example, he noted new diseases such as rickets, and he made the following observations:

- Some diseases affected a similar number of people from year to year, while others varied considerably over time.
- Common causes of death included old age, consumption, smallpox, plague, and diseases of teeth and worms.
- Many greatly feared causes of death were actually uncommon, including leprosy, suicide, and starvation.
- Four separate periods of increased mortality caused by the plague occurred from 1592 to 1660.
- The mortality rate for men was higher than for women.
- Fall was the "most unhealthful season."

Graunt was the first to estimate the number of inhabitants, age structure of the population, and rate of population growth in London and the first to construct a life table that summarized patterns of mortality and survival from birth until death (see Table 1–2). He found that the mortality rate for children was quite high; only 25 individuals out of 100 survived to age 26 years. Furthermore, even though mortality rates for adults were much lower, very few people reached old age (only 3 of 100 London residents survived to age 66 years).

Graunt did not accept the statistics at face value but carefully considered their errors and ambiguities. For example, he noted that it was often difficult for the "antient matron" searchers to determine the exact cause of death. In fact, by cleverly comparing the number of plague deaths and nonplague deaths, Graunt estimated that London officials had overlooked about 20% of deaths resulting from plague.[22]

Although Graunt modestly stated that he merely "reduced several great confused *Volumes* into a few perspicuous *Tables* and abridged such *Observations* as naturally flowed from them," historians consider his work much more significant. Statistician Walter Willcox summarized Graunt's importance:

> Graunt is memorable mainly because he discovered the numerical regularity of deaths and births, of ratios of the sexes at death and birth, and of the proportion of deaths from certain causes to all causes in successive years and in different areas; or in general terms, the uniformity and predictability of many important biological phenomena taken in the mass. In doing so, he opened the way both for the later discovery of uniformities in many social and volitional phenomena like marriage, suicide and crime, and for a study of these uniformities, their nature and their limits.[21(p xiii)]

TABLE 1–2 Life Table of the London Population Constructed by John Graunt in 1662

Age (years)	*Number dying*	*Number surviving*
Birth	0	100
6	36	64
16	24	40
26	15	25
36	9	16
46	6	10
56	4	6
66	3	3
76	2	1
86	1	0

Source: Data from Graunt J. *Natural and Political Observations Made upon the Bills of Mortality,* p. 69. Baltimore, MD: The Johns Hopkins Press; 1932.

James Lind

Only a few important developments occurred in the field of epidemiology during the 200-year period following the publication of John Graunt's *Bills of Mortality*. One notable development was the realization that experimental studies could be used to test hypotheses about the laws of mortality. As described in Chapter 7, these studies involve designed experiments that investigate the role of some factor or agent in the causation, improvement, postponement, or prevention of disease.[23] Their hallmarks are: (1) the comparison of at least two groups of individuals (an experimental group and a "control" group) and (2) the active manipulation of the factor or agent under study by the investigator (that is, the investigator assigns individuals either to receive or not to receive a preventive or therapeutic measure).

In the mid-1700s, James Lind conducted one of earliest experimental studies on the treatment of scurvy, a common disease and cause of death at the time.[24(pp145–148)] Although scurvy affected people living on land, sailors often became sick and died from this disease while at sea. As a ship's surgeon, Lind had many opportunities to observe the "epidemiology" of this disease. His astute observations led him to dismiss the popular ideas that scurvy was a hereditary or infectious disease and to propose that "the principal and main predisposing cause" was moist air and that its "occasional cause" was diet.[24(pp64–67,85,91)] He evaluated his hypothesis about diet with the following experimental study.

> On the 20th of May, 1747 I took twelve patients in the scurvy, on board the *Salisbury* at sea. Their cases were as similar as I could have them. They all in general had putrid gums, the spots and lassitude, with weakness of their knees. They lay together in one place, being a proper apartment for the sick in the fore-hold; and had one diet common to all, *viz,* water-gruel sweetened with sugar in the morning; fresh mutton-broth often times for dinner; at other times puddings, boiled biscuit with sugar, etc.; and for supper barley and raisins, rice and currents, sago and wine, or the like. Two of these were ordered each a quart of cyder a day. Two others took twenty-five gutts of *elixir vitriol* three times a-day. . . . Two other took two spoonfuls of vinegar three times a-day. . . . Two of the worst patients, with the tendons in the ham rigid (a symptom none of the rest had), were put under a course of sea-water. . . . Two others had each two oranges and one lemon given them every day. . . . They continued but six days under this course having consumed the quantity that could be spared. . . . The two remaining patients took bigness of a nutmeg three times a day, of an electuary recommended by an hospital-surgeon made of garlic, mustard seed.[24(pp145–148)]

After 4 weeks, Lind reported the following: "The consequence was, that the most sudden and visible good effects were perceived from the use of the oranges and lemons; one of those who had taken them being at the end of six days fit for duty. . . . He became quite healthy before we came into Plymouth which was on the 16th of June. . . . The other was the best recovered of any in his condition; and being now deem pretty well, was appointed nurse to the rest of the sick."24(p146) Lind concluded, "I shall here only observe, that the result of all my experiments was, that oranges and lemons were the most effectual remedies for this distemper at sea. I am apt to think oranges preferable to lemons though perhaps both given together will be found most serviceable."24(p148)

Although the sample size of Lind's experiment was quite small by today's standards (12 men divided into six groups of two), Lind followed one of the most important principles of experimental research—ensuring that important aspects of the experimental conditions remained similar for all study subjects. Lind selected sailors whose disease was similarly severe, who lived in common quarters, and who had a similar diet. Thus, the main difference between the six groups of men was the dietary addition purposefully introduced by Lind. He also exhibited good scientific practice by confirming "the efficacy of these fruits by the experience of others."24(p148) In other words, Lind did not base his final conclusions about the curative powers of citrus fruits on a single experiment, but rather he gathered additional data from other ships and voyages.

Lind used the results of this experiment to suggest a method for preventing scurvy at sea. Because fresh fruits were likely to spoil and were difficult to obtain in certain ports and seasons, he proposed that lemon and orange juice extract be carried on board.24(pp155–156) The British Navy took 40 years to adopt Lind's recommendation; within several years of doing so, it had eradicated scurvy from its ranks.24(pp377–380)

William Farr

William Farr made many important advances in the field of epidemiology in the mid-1800s. Now considered one of the founders of modern epidemiology, Farr was the compiler of Statistical Abstracts for the General Registry Office in Great Britain from 1839 through 1880. In this capacity, Farr was in charge of the annual count of births, marriages, and deaths. A trained physician and self-taught mathematician, "Farr pioneered a whole range of activities encompassed by modern epidemiology. He described the state of health of the population, he sought to establish the determinants of public health, and he applied the knowledge gained to the prevention and control of disease."25(p xi–xii)

One of Farr's most important contributions involved calculations that combined registration data on births, marriages, and deaths (as the numerator) with census data on the population size (as the denominator). As he stated, "The simple process of comparing the deaths in a given time out of a given number" was "a modern discovery."25(p170) His first annual report in 1839 demonstrated the "superior precision of numerical expressions" over literary expressions.[25(p214)] For example, he quantified and arranged mortality data in a manner strikingly similar to modern practice (see Table 1–3). Note that the annual percentage of deaths increased with age for men and women, but for most age groups the percentage was higher for men than for women.

Farr drew numerous inferences about the English population by tabulating vital statistics. For example, he reported the following findings:

- The average age of the English population remained relatively constant over time at 26.4 years.
- Widowers had a higher marriage rate than bachelors.
- The rate of illegitimate births declined over time.
- People who lived at lower elevations had higher death rates resulting from cholera than did those who lived at higher elevations.
- People who lived in densely populated areas had higher mortality rates than did people who lived in less populated areas.
- Decreases in mortality rates followed improvements in sanitation.

Farr used these data to form hypotheses about the causes and preventions of disease. For example, he used data on smallpox deaths to derive a general law of epidemics that accurately predicted the decline of the rinderpest epidemic in the 1860s.[25(p x)] He used the data on the association between cholera deaths and altitude to support the hypothesis that

TABLE 1–3 Annual Mortality per Hundred Males and Females in England and Wales, 1838–1871

Ages (years)	*Males*	*Females*
0–4	7.26	6.27
5–9	0.87	0.85
10–14	0.49	0.50
15–24	0.78	0.80
25–34	0.99	1.01
35–44	1.30	1.23
45–54	1.85	1.56
55–64	3.20	2.80
65–74	6.71	5.89
75–84	14.71	13.43
85–94	30.55	27.95
95+	44.11	43.04

Source: Data from Farr W. *Vital Statistics: a Memorial Volume of Selections from the Reports and Writings of William Farr,* p. 183. New York, NY: Academy of Medicine; 1975.

an unhealthful climate was the disease's cause—a theory that was subsequently disproved.

Farr made several practical and methodological contributions to the field of epidemiology. First, he constantly strove to ensure that the collected data were accurate and complete. Second, he devised a categorization system for the causes of death so that these data could be reduced to a usable form. The system that he devised is the antecedent of the modern *International Classification of Diseases*, which categorizes diseases and causes of death. Third, Farr made a number of important contributions to the analysis of data, including the invention of the "standardized mortality rate," an adjustment method for making fair comparisons between groups with different age structures.

John Snow

Another important figure in the development of epidemiologic methods during the mid-1800s was John Snow (see Figure 1–1). A respected physician who was a successful anesthetist and researcher on anesthetic gases, Snow was also interested in the cause and spread of cholera.[26(p xxxiv)] Although Farr mistakenly thought that an unhealthful climate accounted for the variation in cholera mortality by altitude, Snow used these data to support an innovative hypothesis that cholera was an infectious disease spread by fecal contamination of drinking water.

Snow argued, "Cholera must be a poison acting on the alimentary canal by being brought into direct contact with the alimentary mucous surface . . . the symptoms are primarily seated in the alimentary canal and all the after-symptoms of a general kind are the results of the flux from the canal. His inference from this was, that the poison of cholera is taken direct into the canal by mouth. This view led him to consider the media through which the poison is conveyed and the nature of the poison itself. Several circumstances lent their aid in referring him to water as the chief, though not the only, medium, and to the excreted matters from the patient already stricken with cholera, as the poison."[26(p xxxiv–xxxv)]

In 1849 Snow published his views on the causes and transmission of cholera in a short pamphlet titled *The Mode of Communication of Cholera*. During the next few years he continued groundbreaking research testing the hypothesis that cholera was a waterborne infectious disease. The second edition of his pamphlet, published in 1855, describes in greater detail "the whole of his inquiries in regard to cholera."[26(p xxxvi)] The cholera investigations for which Snow is best known are described in the following paragraphs.

One such investigation focused on the Broad Street epidemic. During August and September of 1854, one of the worst outbreaks of cholera occurred in the London neighborhood surrounding Broad Street. Almost

FIGURE 1–1. John Snow investigated the cause and spread of cholera in 19th-century London.

Source: Reprinted with permission from the Wellcome Trust Library, London, England.

500 fatalities from cholera occurred within a 10-day period within 250 yards of the junction between Broad and Cambridge Streets (see Figure 1–2). According to Snow, "The mortality in this limited area probably equals any that was ever caused in this country, even by the plague; and it was much more sudden, as the greater number of cases terminated in a few hours."[26(p38)] Snow continued,

> As soon as I became acquainted with the situation and extent of this irruption of cholera, I suspected some contamination of the water of the much-frequented street-pump in Broad Street, near the end of Cambridge Street; but on examining the water, on the evening of the 3rd of September, I found so little impurity in it of an organic nature that I hesitated to come to a conclusion. Further inquiry, however, showed me that there was no other circumstance or agent common to the cholera occurred, and not extending beyond it, except the water of the above mentioned pump.[26(p39)]

His subsequent investigations included a detailed study of the drinking habits of 83 individuals who died between August 31 and September 2, 1854.[26(pp39–40)] He found that 73 of the 83 deaths occurred among indi-

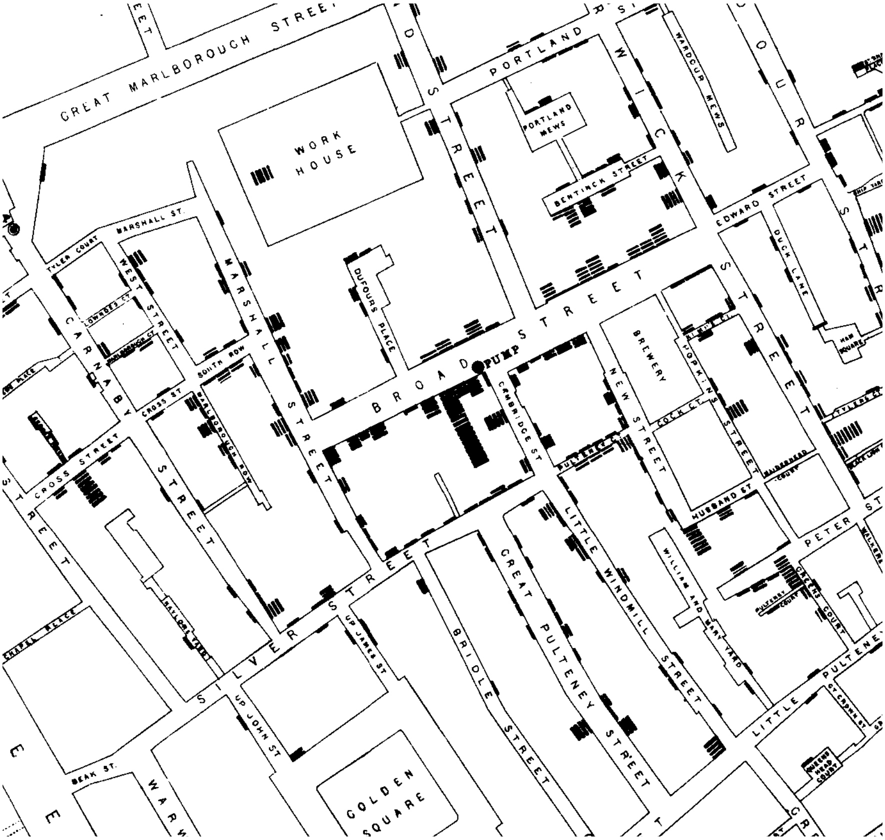

FIGURE 1–2. Distribution of Deaths from Cholera in the Broad Street Neighborhood from August 19 to September 30, 1854. "A black mark or bar for each death is placed in the situation of the house in which the fatal attack took place. The situation of the Broad Street Pump is also indicated, as well as that of all the surrounding Pumps to which the public had access."

Source: Reprinted from Snow J. *Snow on Cholera,* Map 1. New York, NY: Hafner Publishers; 1965. With permission from the Commonwealth Fund, New York, NY.

viduals living within a short distance of the Broad Street pump and that 10 deaths occurred among individuals who lived in houses that were near other pumps. According to the surviving relatives, 61 of the 73 individuals who lived near the pump drank the pump water and only 6 individuals did not. (No data could be collected for the remaining 6 because everyone connected with these individuals had either died or departed the city). The drinking habits of the 10 individuals who lived "decidedly nearer to another street pump" also implicated the Broad Street pump. Surviving relatives reported that 5 of the 10 drank water from the Broad Street pump because they preferred it, and 2 drank its water because they attended a nearby school.

Snow also investigated pockets of the Broad Street population that had fewer cholera deaths. For example, he found that only 5 cholera deaths occurred among 535 inmates of a workhouse located in the Broad Street neighborhood.[26(p42)] The workhouse had a pump-well on its premises, and "the inmates never sent to Broad Street for water." Furthermore, no cholera deaths occurred among 70 workers at the Broad Street brewery who never obtained pump water but instead drank a daily ration of malt liquor.

Although Snow never found direct evidence of sewage contamination of the Broad Street pump-well, he did note that the well was near a major sewer and several cesspools. He concluded, "There had been no particular outbreak or increase of cholera, in this part of London, except among the persons who were in the habit of drinking the water of the above-mentioned pump-well."[26(p40)] He presented his findings to the Board of Guardians of St. James's Parish on September 7, and "the handle of the pump was removed on the following day."[26]

Snow's investigation of the Broad Street epidemic is noteworthy for several reasons. First, Snow was able to form a hypothesis implicating the Broad Street pump after he mapped the geographic distribution of the cholera deaths and studied that distribution in relation to the surrounding public water pumps (see Figure 1–2). Second, he collected data on the drinking water habits of unaffected as well as affected individuals, which allowed him to make a comparison that would support or refute his hypothesis. Third, the results of his investigation were so convincing that they led to immediate action to curb the disease—namely, the pump handle was removed. Public health action to prevent disease seldom occurs so quickly.

Another series of Snow's groundbreaking investigations on cholera focused on specific water supply companies. In particular, he found that districts supplied by the Southwark and Vauxhall Company and the Lambeth Company had higher cholera mortality rates than all of the other water companies.[26(pp63–64)] A few years later, a fortuitous change occurred in the water source of several of the south districts of London. As Snow stated, "The Lambeth Company removed their water works, in 1852, from opposite Hungerforth Market to Thames Ditton; thus obtaining a supply of water quite free from the sewage of London."[26(p68)]

Following this change, Snow obtained data from William Farr to show that "districts partially supplied with the improved water suffered much less than the others"[26(p69)] (see Table 1–4). Districts with a mixture of the clean and polluted drinking water (Southwark and Vauxhall and Lambeth Companies combined) had 35% fewer cholera deaths (61 versus 94 deaths per 10,000) than districts with only polluted drinking water (Southwark and Vauxhall Company alone).

TABLE 1–4 Mortality from Cholera in Relation to the Water Supply Companies in the Districts of London, November 1853

Water supply company	*Number of cholera deaths*	*Size of population*	*Death rate resulting from cholera*
Southwark and Vauxhall	111	118,267	94 / 100,000
Southwark and Vauxhall, Lambeth	211	346,363	61 / 100,000

Source: Data from Snow J. *Snow on Cholera,* p. 69. New York, NY: Hafner Publishers; 1965. With permission from the Commonwealth Fund, New York, NY.

TABLE 1–5 Mortality from Cholera in Relation to the Water Supply Companies in the Subdistricts of London, 1853

Water supply company	*Number of cholera deaths*	*Size of population*	*Death rate resulting from cholera*
Southwark and Vauxhall alone	192	167,654	114 / 100,000
Southwark and Vauxhall, and Lambeth combined	182	301,149	60 / 100,000
Lambeth alone	0	14,632	0 / 100,000

Source: Data from Snow J. *Snow on Cholera,* p. 73. New York, NY: Hafner Publishers; 1965. With permission from the Commonwealth Fund, New York, NY.

Snow next analyzed the cholera mortality data in smaller geographic units—London subdistricts—in order to make an even clearer distinction between the polluted and clean water supplies. In particular, he examined the death rates in the London subdistricts supplied by: (1) the Southwark and Vauxhall Company alone (heavily polluted water), (2) the Lambeth Company alone (nonpolluted water), and (3) both companies combined (a mixture of polluted and nonpolluted water). The cholera death rates were highest in subdistricts supplied by the heavily polluted water of the Southwark and Vauxhall Company and were intermediate in subdistricts supplied by the mixed water from the Southwark and Vauxhall Company and Lambeth Company combined. No cholera deaths were observed in subdistricts supplied with the nonpolluted water of the Lambeth Company (see Table 1–5).

Although Snow thought that these data provided "very strong evidence of the powerful influence which the drinking of water containing the sewage of a town exerts over the spread of cholera, when that disease is present," he thought that further study of the people living in the subdistricts supplied by both companies would "yield the most incontrovertible proof on one side or another." Snow understood that the differences in cholera deaths rates between the two companies might not have been caused by the water supply itself but rather by differences between the groups such as differences in gender, age, and socioeconomic status. Fortunately for Snow, further study revealed that the two groups were strikingly similar.

Snow made the following observation:

> In the subdistricts enumerated in the above table as being supplied by both companies, the mixing of the supply is of the most intimate kind. The pipes of each Company go down all the street, and into nearly all the courts and alleys. A few houses are supplied by one Company and a few by the other, according to the decision of the owner or occupier at that time when the water companies were in active competition. In many cases a single house has a supply different from that on either side. Each company supplied both rich and poor, both large houses and small; there is no difference either in the condition or occupation of the persons receiving the water of different companies. . . . As there is no difference whatever, either in the houses or the people receiving the supply of the two Water Companies, or in any of the physical conditions with which they are surrounded, it is obvious that no experiment could have been devised which would more thoroughly test the effect of water supply on the progress of cholera than this, which circumstances placed ready made before the observer. The experiment, too, was on the grandest scale. No fewer than three hundred thousand people of both sexes, of every age and occupation, and of every rank and station, from gentlefolks down to the very poor, were divided into two groups without their choice, and, in most cases, without their knowledge; one group being supplied with water containing the sewage of London, and amongst it, whatever might have come from the cholera patients, the other group having water quite free from such impurity.[26(pp74–75)]

Snow's next step was to obtain a listing from the General Register of the addresses of persons dying of cholera in the subdistricts that used water from both suppliers. Then he had the difficult task of going door to door to inquire about the drinking water supplier. According to Snow, "The inquiry was necessarily attended with a good deal of trouble. There were very few instances in which I could get the information I required. Even when the water-rates are paid by the residents, they can seldom remember the name of the Water Company till they have looked for the receipt."[26(p76)] However, Snow found an ingenious solution to this problem:

> It would, indeed, have been almost impossible for me to complete the inquiry, if I had not found that I could distinguish the water of the two companies with perfect certainty by a chemical test. The test I employed was founded on the great difference in the quantity of chloride of sodium contained in the two kinds of water. On adding solution of nitrate of silver to a gallon of water of the Lambeth Company . . . only 2.28 grains of chloride of silver were obtained. . . . On treating the water of Southwark

TABLE 1–6 Mortality from Cholera in Relation to the Water Supply Companies in the Subdistricts of London, July–August 1854

Water supply company	*Number of cholera deaths*	*Number of houses*	*Death rate resulting from cholera*
Southwark and Vauxhall Company	1,263	40,046	315/10,000
Lambeth Company	98	26,107	37/10,000
Rest of London	1,422	256,423	59/10,000

Source: Adapted from Snow J. *Snow on Cholera,* p. 86. New York, NY: Hafner Publishers; 1965. With permission from the Commonwealth Fund, New York, NY.

> and Vauxhall Company in the same manner, 91 grains of chloride of silver were obtained.[26(pp77–78)]

Thus, Snow identified the drinking water source of each household and was able to link the death rate from cholera to the water supply companies (see Table 1–6). He concluded, "The mortality in the houses supplied by the Southwark and Vauxhall Company was therefore between eight and nine times as great as in the houses supplied by the Lambeth Company."[26(p86)]

Based on his findings, Snow made a series of recommendations for the prevention of cholera. For example, he recommended, "Care should be taken that the water employed for drinking and preparing food . . . is not contaminated with the contents of cesspools, house-drains, or sewers; or in the event that water free from suspicion cannot be obtained, is [sic] should be well boiled, and if possible, also filtered."[26(pp133–134)] Even though his results and recommendations were reported at once to William Farr and others, it took several years for Snow's theories to be accepted.[27]

Fortunately, over time we have come to recognize the importance of John Snow's contributions to our understanding of infectious diseases, in general, and cholera, in particular. For several reasons, Snow's investigations are considered "a nearly perfect model" for epidemiologic research.[26(pix)] First, Snow organized his observations logically so that meaningful inferences could be derived from them.[20(p29)] Second, he recognized that "a natural experiment" had occurred in the subdistricts of London that would enable him to gather unquestionable proof either for or against his hypothesis. Third, he conducted a quantitative analysis of the data contrasting the occurrence of cholera deaths in relation to the drinking water company.

Modern Experimental Studies

The development and application of epidemiologic methods advanced slowly during the late 1800s and early 1900s. Only during the 1930s and

1940s did physicians begin to realize that it was necessary to refine the methods used to evaluate the effectiveness of disease treatments.[28] Although some physicians still thought that they could assess the usefulness of a treatment merely by observing the patient's response and comparing it to what they expected on the basis of their education and experience, many realized that "modern" experimental studies with comparable treatment and control groups of patients, and comparable methods for assessing the disease changes were needed to yield correct conclusions.[29]

Streptomycin Tuberculosis Trial

In the late 1940s the Streptomycin in Tuberculosis Trials Committee of the British Medical Research Council conducted one of the first modern experimental studies on the use of streptomycin to treat pulmonary tuberculosis.[30] According to the investigators:

> The natural course of pulmonary tuberculosis is in fact so variable and unpredictable that evidence of improvement or cure following the use of a new drug in a few cases cannot be accepted as proof of the effect of that drug. The history of chemotherapeutic trials in tuberculosis is filled with errors due to empirical evaluation of drugs. . . . It had become obvious that . . . conclusions regarding the clinical effect of a new chemotherapeutic agent in tuberculosis could be considered valid only if based on adequately controlled clinical trials.[30(p4582)]

This controlled clinical trial of streptomycin included 107 patients with acute progressive bilateral pulmonary tuberculosis.[30] The investigators decided to include only cases of tuberculosis that were unsuitable for other forms of treatment in order "to avoid having to make allowances for the effect of forms of therapy other than bed-rest." In addition, they excluded cases in which spontaneous regression was likely and cases in which there was little hope of improvement.

One group of 55 patients was treated with bed rest and streptomycin, and a second group of 52 patients was treated with bed rest alone.[30] Patients were assigned to these groups by an innovative method known as randomization, which is defined as "an act of assigning or ordering that is the result of a random process."[23(p220)] As described in Chapter 7, random assignment methods include flipping a coin or using a sequence of random numbers. The exact process used in the Streptomycin Tuberculosis Trial was as follows: "Determination of whether a patient would be treated by streptomycin and bed rest (S case) or by bed rest alone (C case) was made by reference to a statistical series based on random sampling numbers drawn up for each sex at each centre by Professor Bradford Hill; the details of the series were unknown to any of the investigators or to the co-ordinator and were contained in a set of sealed envelopes."[30(p770)]

Patients in the streptomycin group received the drug by injection four times a day.[30] Although investigators observed toxic effects in many patients, these effects were not so severe as to require the termination of treatment. During the 6-month follow-up period, 7% of the streptomycin patients died and 27% of the control patients died. Investigators observed x-ray evidence of considerable pulmonary improvement in 51% of the streptomycin patients and only 8% of the control patients. Clinical improvement was also more common in the streptomycin group. The investigators reached the following conclusion: "The course of bilateral acute progressive disease can be halted by streptomycin therapy. . . . That streptomycin was the agent responsible for this result is attested by the presence in this trial of the control group of patients, among whom considerable improvement was noted in only four (8%)."[30(p780)]

According to Richard Doll, "Few innovations have made such an impact on medicine as the controlled clinical trial that was designed by Sir Austin Bradford Hill for the Medical Research Council's Streptomycin in Tuberculosis Trials Committee in 1946."[29(p343)] Four features of the trial were particularly innovative. First and foremost was its use of randomization to assign patients to the streptomycin and control groups. Although randomization had been used in agriculture and laboratory research, this trial was one of the first instances in which it was used in medical research. As described in Chapter 7, the main advantage of randomization is that the order of future assignments cannot be predicted from that of past ones. Lack of predictability is the key to minimizing bias, which is defined as a systematic error in the study that causes a false conclusion.

The second innovation was the placement of restrictions on the type of patient eligible for the trial.[29] Patients with the type of tuberculosis that was unsuitable for therapies other than bed rest were excluded so that the results would not be obscured by the effects of other treatments. Patients who were likely to get better without any treatment or who were so ill that the streptomycin was unlikely to help were also excluded.

Third, the data collection methods helped ensure that the results would be free of bias.[29] These methods included using a precise and objective endpoint such as death, and masking the investigators who were assessing the radiological improvements. Masking means that the investigators who reviewed the x-rays were unaware of the person's treatment assignment and so the chances of their making a biased judgment were reduced.

Fourth, the investigators considered the ethical issues involved in conducting the trial, including whether it was ethical to withhold the streptomycin treatment from the control group.[29] Before the trial was conducted, researchers had already shown that streptomycin inhibited the tubercle bacillus in vitro and reduced experimental infections of guinea pigs. Preliminary results of clinical studies had also been encouraging. However, only a small amount of the drug was available in Britain, and it was impossible to treat all patients with tuberculosis. Thus, the committee reasoned, "It would . . . have been unethical *not* to have seized the opportunity to design a strictly controlled trial, which could speedily and effectively reveal the value of the treatment."[29(p339)]

Doll and Hill's Studies on Smoking and Lung Cancer

Most epidemiologists consider Richard Doll and A. Bradford Hill's 1950 study on smoking and lung cancer to be one of the major milestones of epidemiology.[31] Doll and Hill undertook the study because of the striking increase in lung cancer death rates in England and Wales following World War I.[32] Some scientists argued that the increase was the result of improvements in lung cancer diagnosis. However, Doll and Hill believed that improved diagnosis could not be entirely responsible, because the number of lung cancer deaths had increased in areas with and without modern diagnostic facilities. Thus, Doll and Hill thought it was "right and proper" to justify searching for an environmental cause. Their work is emblematic of an important shift in epidemiology following World War II that redirected the focus of epidemiologic research from infectious to chronic diseases.[31] The shift was fueled by the idea that chronic diseases were not merely degenerative disorders of old age but rather were potentially preventable diseases with environmental origins.

Doll and Hill's first study was a "case–control study,"[32] an epidemiologic design that is described in more detail in Chapter 9. The study included 709 subjects who had lung cancer (the cases) and 709 subjects who had diseases other than cancer (the controls). Control patients were purposely selected to be of the same sex, within the same 5-year age group, and in the same hospital at approximately the same time as the lung cancer patients.

Patients from each group were interviewed about their smoking habits while they were in the hospital for treatment. In particular, they were asked: "(a) if they had smoked at any period of their lives; (b) the ages at which they had started and stopped; (c) the amount they were in the habit of smoking before the onset of the illness which had brought them to the hospital; (d) the main changes in their smoking history and the maximum they had ever been in the habit of smoking; (e) the varying proportions smoked in pipes and cigarettes; and (f) whether or not they inhaled."[32(p741)]

Doll and Hill found that proportionately more lung cancer patients than noncancer patients were smokers.[32] In particular, 99.7% of male lung cancer patients and 95.8% of male noncancer patients smoked; 68.3% of female lung cancer patients and only 46.7% of female noncancer patients were smokers. Furthermore, a higher proportion of patients with lung cancer described themselves as heavy smokers. For example, 26.0% of the male lung cancer patients and 13.5% of the male noncancer patients reported that they had smoked 25 or more cigarettes per day before their illness began. Although the authors acknowledged that they did not know what carcinogens in tobacco smoke might be responsible, they concluded that "smoking is an important factor in the cause of lung cancer."

Three other case–control studies published in 1950 also showed an association between smoking and lung cancer. However, modern epidemiologists consider the Doll and Hill study to be "a classic exemplar for the investigation of a given outcome and an array of exposures. . . . No previous research paper lays out the essentials of the case–control method with such understanding and meticulous care."[31(p163)] Far in advance of their peers, Doll and Hill considered a wide range of problems in the design and analysis of their study, including errors that may have occurred when they recruited and interviewed their subjects. These topics are described in Chapters 10 and 11.

In the years following the 1950 Doll and Hill study, several more studies were conducted using the case–control approach of comparing the smoking histories of patients with and without lung cancer (such as Wynder and Cornfeld's 1953 study).[33] These studies all found that the proportion of smokers, particularly heavy smokers, was higher among lung cancer patients than among noncancer patients. However, Doll and Hill believed that additional "retrospective" studies were "unlikely to advance our knowledge materially or to throw any new light upon the nature of the association." (Retrospective studies investigate diseases that have already occurred.) They asserted that if there were "any undetected flaw in the evidence that such studies have produced, it would be exposed only by some entirely new approach." The new approach that they proposed was a "prospective" study, a study that follows participants into the future in order to observe the occurrence of disease (see Chapter 8).

Doll and Hill initiated a prospective study in 1951 by inviting 59,600 male and female members of the British Medical Association to complete a short questionnaire about their smoking habits.[34] The investigators then divided the respondents into four groups on the basis of their answers: nonsmokers and light, moderate, and heavy smokers. The investigators obtained information on the causes of death among those who answered the questionnaire from the Registrars-General Office in the United Kingdom.

During the 29-month period following the administration of the questionnaire, 789 deaths were reported among the 24,389 male doctors aged 35 years and older. Of these deaths, 36 were reported to have died of lung cancer as either the direct or contributory cause. After accounting for age differences between the smoking groups, the investigators found that death rates caused by lung cancer increased from 0.0 per 1000 among nonsmokers to 0.48 per 1000 among light smokers, 0.67 per 1000 among moderate smokers, and 1.14 per 1000 among heavy smokers.[34]

The investigators continued to follow the doctors for the next 20 years.[35] During this period they updated the smoking and mortality data. Of the 34,400 men studied, 10,072 were known to have died from November 1951 through October 1971. Among men under the age of 70 years, the death rate was about twice as high among cigarette smokers as among lifelong nonsmokers. The causes of death related to smoking included not only lung cancer but also heart disease, chronic obstructive lung disease, and a variety of vascular diseases.

Because the proportion of doctors who smoked cigarettes declined over the 20-year period, the investigators were also able to examine the death rate among former smokers who had stopped smoking for various lengths of time. They found that, as compared to lifelong nonsmokers, the risk of lung cancer death among ex-smokers steadily declined in relation to the number of years since they had stopped smoking. However, even 15 or more years after quitting, the risk of lung cancer death was still twice as high among former smokers than among lifelong nonsmokers.

Like their first case–control study, Doll and Hill's prospective study broke new ground. First, the study included tens of thousands of subjects, and so it had adequate "power" to examine numerous health effects of several levels of smoking. Second, the investigators followed the subjects for a long period of time. A long follow-up period is particularly important in the study of diseases such as cancer that take decades to develop. Third, Doll and Hill incorporated changes in smoking habits over time, and so were able to examine the health benefits of smoking cessation.

The Framingham Study

Like the work of Doll and Hill, the Framingham Study is notable for bringing about a shift in focus from infectious to noninfectious diseases following World War II. Considered "the epitome of successful epidemiologic research," this study "has become the prototype and model of the cohort study."[31(p157)] The cohort study is one of the three main study designs used in epidemiologic research, and is described in Chapters 6 and 8.

According to Susser, the Framingham Study is "undisputably the foundation stone for current ideas about risk factors in general and the prevention of ischemic heart disease in particular." In addition, it has provided the impetus for solving difficult design and analysis issues in epidemiologic research, including the development of appropriate methods for measuring the major risk factors for coronary heart disease (such as high blood pressure, elevated serum cholesterol levels, physical activity, and life stress) and for solving problems associated with measurements that vary over time.[31(pp157–161)] The study has also served as a stimulus for developing other cohort studies of cardiovascular disease and other topics.

When the Framingham Study was started in 1947, its goal was to develop ways of identifying latent cardiovascular disease among healthy volunteers.[31] Within a few years, investigators expanded the study's purpose to include determining the causes of cardiovascular disease. The study now investigates a wide variety of diseases, including stroke, cerebrovascular disease, diabetes, Alzheimer's disease, and cancer, and it includes the offspring and grandchildren of the original participants.

Initially, the investigators enrolled about 5000 healthy adult residents in Framingham, Massachusetts, a town located about 18 miles west of Boston.[36(pp14–29)] In the late 1940s, Framingham was a self-contained community of about 28,000 residents who obtained their medical care from local physicians and two hospitals near the center of town. Framingham residents were considered an excellent population for a community-based prospective study because (1) the town's population was stable, (2) the investigators could identify a sufficient number of people with and without risk factors for heart disease, and (3) local medical doctors were eager to help recruit study subjects.

For the past 50 years Framingham Study participants have undergone interviews, physical exams, laboratory tests, and other tests every 2 years.[37] The interviews have gathered information on each subject's medical history and history of cigarette smoking, alcohol use, physical activity, dietary intake, and emotional stress. The physical exams and laboratory tests have measured characteristics such as height and weight, blood pressure, vital signs and symptoms, cholesterol levels, glucose levels, and bone mineral density. These data-gathering efforts have left an immeasurable legacy of research findings on numerous topics.

Modern Epidemiology

In recent years the field of epidemiology has expanded tremendously in size, scope, and influence. The number of epidemiologists has grown rapidly along with the number of epidemiology training programs in schools of public health and medicine. Many subspecialties have been established that are defined either by (1) disease, (2) exposure, or (3) population being studied. Disease-specific subspecialties include reproductive, cancer, cardiovascular, infectious disease, and psychiatric epidemiology. Exposure-specific subspecialties include environmental, behavioral, and nutritional epidemiology and pharmaco-epidemiology. Population-specific subspecialties include pediatric and geriatric epidemiology.

In addition, the scope of epidemiologic research has expanded in two divergent directions. First, some epidemiologists examine health determinants at the molecular and genetic level and so combine the basic and public health sciences. For example, genetic epidemiology investigates whether certain diseases cluster in families, whether the clustering is caused by inherited factors or a shared environment, and how genes influence the risk of disease.[38] Molecular epidemiology involves the use of "molecular markers" to improve the measurement of exposure and diseases.[39(pp413–415)] For example, molecular markers such as serum micronutrient levels can determine a person's fruit and vegetable intake more accurately than can personal interviews.

The second new direction of epidemiologic research involves the study of determinants at the biological and societal level.[40] Social epidemiology is the study of exposures and disease susceptibility and resistance at diverse levels including the individual, household, neighborhood, and region. For example, social epidemiologists investigate how neighborhoods, racial discrimination, and poverty influence a person's health.

In addition, the theories and methods of epidemiologic research have evolved during recent years. For example, epidemiologists have developed new views on disease causation such as the "sufficient-component cause model." The theoretical framework underlying epidemiologic study designs has matured, particularly the conceptualization and design of the case–control study. Finally, the availability of high-powered computer hardware and software has facilitated the analysis of large data sets with many risk factors, enabling epidemiologists to explore new public health questions and to assess the effects of multiple risk factors simultaneously.

Not surprisingly, epidemiology is currently being used to investigate a wide range of important public health topics. Noteworthy topics that have been examined recently include the risk of brain cancer among cell phone users,[41] the epidemiology of emerging and re-emerging infectious diseases such as HIV and tuberculosis,[42] unexplained chronic illnesses among Persian Gulf War veterans,[43] molecular epidemiology of Alzheimer's disease,[44] the effectiveness of mammography for women under the age of 50 years,[45] the safety of hormone replacement therapy during menopause,[46] and the effectiveness of alternative medicines in treating disease.[47]

The 21st century poses even more challenging problems for epidemiologists such as "air, water and soil pollution; global warming; population growth; poverty and social inequality; and civil unrest and violence."[48(p5)] A recent editorial on epidemiology in the 21st century noted that, like public health achievements of the past, solutions to these problems will occur through "the complementary contributions of different facets of epidemiology: calculating disease trends and probabilities, communicating findings to the public and policy makers, and designing and implementing interventions based on data."[49(p1154)] The editorialists went on to observe: "Epidemiology's full value is achieved only when its contributions are placed in the context of public health action, resulting in a healthier populace. . . . Like others in epidemiology's rich history, we should keep our eyes on the prizes of preventing disease and promoting health."[49(p1155)] The prospect of preventing disease and death through "analytic prowess" has attracted many great minds to epidemiology throughout its history, and it will undoubtedly continue to attract them in the coming century.

Summary

Disease prevention and health promotion are the main goals of public health, a multidisciplinary field that focuses on populations and communities rather than separate individuals. Epidemiology, one of the basic sciences of public health, is defined as "the study of the distribution and determinants of disease frequency in human populations and the application of this study to control health problems." Epidemiology has played an important role in public health achievements of the last 400 years. Key historic figures and studies have included John Graunt, who summarized the patterns of mortality in 17th-century London; James Lind, who discovered the cause and prevention of scurvy using an experimental study design in the 18th century; William Farr, who originated many modern epidemiologic methods in the 19th century including the combination of numerator and denominator data; John Snow, who demonstrated that contaminated drinking water was the mode of cholera transmission in the 19th century; members of the Streptomycin in Tuberculosis Trials Committee, who conducted one of the first modern controlled clinical trials in the 1940s; Doll and Hill, who conducted case–control studies on smoking and lung cancer in the 1950s; and investigators who have worked on the Framingham Study, a study that was started in 1947 and has become one of the most influential studies of heart disease in the world. In recent years the field of epidemiology has greatly expanded in size, scope, and influence, and epidemiologists currently investigate a wide range of important public health problems. The 21st century will pose even more challenging problems for epidemiologists. Like past public health achievements, the solutions to these problems will be found by placing the contributions of epidemiology in the context of public health action.

References

1. Schneider MJ. *Introduction to Public Health.* Gaithersburg, MD: Aspen Publishers, Inc; 2000.
2. National Center for Health Statistics. *Health, United States, 2000 with Adolescent Health Chart Book.* Hyattsville, MD: Public Health Service; 2000.
3. Bunker JP, Frazier HS, Mosteller F. Improving health: measuring effects of medical care. *Milbank Q.* 1994;72:225–258.
4. Centers for Disease Control and Prevention. Ten great public health achievements—United States, 1900–1999. *MMWR.* 1999;48:241–243.

5. Loomis TA. *Essentials of Toxicology.* 3rd ed. Philadelphia, PA: Lea & Febiger; 1978.
6. Gottlieb MS, Schanker HM, Fan PT, Saxon A, Weisman JD, Pozalski I. *Pneumocystis* pneumonia—Los Angeles. *MMWR.* 1981;30:250–252.
7. Friedman-Kien A, Laubenstein L, Marmor M, et al. Kaposi's sarcoma and *Pneumocystis* pneumonia among homosexual men—New York City and California. *MMWR.* 1981;30:305–308.
8. Boice JD, Monson RR. Breast cancer in women after repeated fluoroscopic examination of the chest. *J Natl Cancer Inst.* 1977;59:823–832.
9. Weitzman M, Aschengrau A, Bellinger D, Jones R, Hamlin JS, Beiser A. Lead-contaminated soil abatement and urban children's blood lead levels. *JAMA.* 1993;269:1647–1654.
10. Pickett JP, ed. *The American Heritage Dictionary of the English Language.* 4th ed. Boston, MA: Houghton Mifflin; 2000.
11. Simpson JA, Weiner ESC, eds. *Oxford English Dictionary.* 2nd ed. Oxford, England: Clarendon Press; 1989. OED Online. Oxford University Press; Accessed 20 July 2001.
12. Lilienfeld D. Definitions of epidemiology. *Am J Epidemiol.* 1978;107:87–90.
13. MacMahon B, Trichopoulos D. *Epidemiology: Principles and Methods.* 2nd ed. Boston, MA: Little, Brown and Co; 1996.
14. Last JM. *A Dictionary of Epidemiology,* 3rd ed. New York, NY: Oxford University Press; 1995.
15. MacMahon B, Pugh TF. *Epidemiology: Principles and Methods.* Boston, MA: Little, Brown and Co; 1970.
16. Centers for Disease Control and Prevention. National Health Interview Survey. Available at: http://www.cdc.gov/nchs/about/major/nhis. Accessed January 2001.
17. Kelsey JL. Breast cancer epidemiology. *Epidemiol Rev.* 1993;15:256–263.
18. Benenson AS, ed. *Control of Communicable Diseases in Man.* 15th ed. Washington DC: American Public Health Association; 1990.
19. Centers for Disease Control and Prevention. *HIV/AIDS Surveillance Rep.* 1999;11:1–44.
20. Lilienfeld DE, Stolley. *Foundations of Epidemiology.* 3rd ed. New York, NY: Oxford University Press; 1994.
21. Graunt J. *Natural and Political Observations Made upon the Bills of Mortality.* Willcox WF, ed. Baltimore, MD: The John Hopkins Press; 1939.
22. Rothman KJ. Lessons from John Graunt. *Lancet.* 1996;347:37–39.
23. Meinert CL. *Clinical Trials Dictionary, Terminology and Usage Recommendations.* Baltimore, MD: The Johns Hopkins Center for Clinical Trials; 1996.
24. Stewart CP, Guthrie D, eds. *Lind's Treatise on Scurvy. A Bicentenary Volume Containing a Reprint of the First Edition of a Treatise of the Scurvy by James Lind, MD, with Additional Notes.* Edinburgh, Scotland: Edinburgh University Press; 1953.
25. Farr, W. *Vital Statistics: A Memorial Volume of Selections from the Reports and Writings of William Farr.* New York, NY: New York Academy of Medicine; 1975.
26. Snow J. *On the Mode of Communication of Cholera.* 2nd ed. London, England: Churchill, 1855. Reprinted as Snow J. *Snow on Cholera.* New York, NY: Hafner Publishing Co; 1965.

27. Winkelstein W. A new perspective on John Snow's communicable disease theory. *Am J Epidemiol.* 1995;142:S3–S9.
28. Vandenbroucke JP. Second thoughts: a short note on the history of the randomized controlled trial. *J Chron Dis.* 1987;40:985–987.
29. Doll R. Clinical trials: retrospect and prospect. *Stat Med.* 1982;1:337–344.
30. Medical Research Council. Streptomycin in Tuberculosis Trials Committee. Streptomycin treatment of pulmonary tuberculosis. *Br Med J.* 1948;2:769–782.
31. Susser M. Epidemiology in the United States after World War II: the evolution of technique. *Epidemiol Rev.* 1985;7:147–177.
32. Doll R, Hill AB. Smoking and carcinoma of the lung. *Br Med J.* 1950;2:739–748.
33. Wynder EL, Cornfeld J. Cancer of the lung in physicians. *N Engl J Med.* 1953;248:441–444.
34. Doll R, Hill AB. The mortality of doctors in relation to their smoking habits. *Br Med J.* 1954;1:1451–1455.
35. Doll R, Peto R. Mortality in relation to smoking: 20 years' observations on male British doctors. *Br Med J.* 1976;2:1525–1536.
36. Dawber TR. *The Framingham Study: The Epidemiology of Atherosclerotic Disease.* Cambridge, MA: Harvard University Press; 1980.
37. Kannel WB. The Framingham Study: its 50-year legacy and future promise. *J Atheroscler Thromb.* 2000;6:60–66.
38. Beaty TH, Khoury MJ. Interface of genetics and epidemiology. *Epidemiol Rev.* 2000;22:120–125.
39. Friis RH, Sellers TA. *Epidemiology for Public Health Practice.* 2nd ed. Gaithersburg, MD: Aspen Publishers, Inc; 1999.
40. Kreiger N. Epidemiology and social sciences: towards a critical reengagement in the 21st century. *Epidemiol Rev.* 2000;22:155–163.
41. Hardell L, Nasman A, Pahlson A, Hallquist A, Hanssion Mild K. Use of cellular telephones and the risk for brain tumours: a case–control study. *Int J Oncol.* 1999;15:113–116.
42. Horsburgh CR Jr, Pozniak A. Epidemiology of tuberculosis in the era of HIV. *AIDS.* 1993;7(suppl 1):S109–S114.
43. Centers for Disease Control and Prevention. Leads from the *Morbidity and Mortality Weekly Report,* Atlanta, GA: Unexplained illness among Persian Gulf War veterans in an Air National Guard Unit: preliminary report, August 1990–March 1995. *JAMA.* 1995;274:16–17.
44. Hyman BT, Tanzi R. Molecular epidemiology of Alzheimer's disease. *N Engl J Med.* 1995;333:1283–1284.
45. Smith RA. Breast cancer screening among women younger than age 50: a current assessment of the issues. *CA Cancer J Clin.* 2000;50:312–336.
46. Finkel ML, Cohen M, Mahoney H. Treatment options for the menopausal woman. *Nurse Pract.* 2001;26:5–17.
47. Moyad MA, Hathaway S, Ni HS. Traditional Chinese medicine, acupuncture, and other alternative medicines for prostate cancer: an introduction and the need for more research. *Semin Urol Oncol.* 1999;17:103–110.
48. Winkelstein W. Interface of epidemiology and history: a commentary on past, present, and future. *Epidemiol Rev.* 2000;22:2–6.
49. Koplan JP, Thacker SB, Lezin NA. Epidemiology in the 21st century: calculation, communication, and intervention. *Am J Public Health.* 1999;89:1153–1155.

EXERCISES

1. Define each of the following terms:
   A. Public health
   B. Epidemiology
   C. Population
2. What is the primary difference between public health and medicine?
3. What are the main objectives of epidemiology?
4. How do epidemiologists quantify the disease frequency in a population?
5. State the contribution that was made by each of the following historical figures:
   A. John Graunt
   B. John Snow
   C. Richard Doll and Austin Bradford Hill