1. THE EMERGENCE OF EPIDEMIOLOGY

Among the sciences, epidemiology is an embryo. Although some excellent epidemiologic studies were conducted before the twentieth century, a systematized body of epidemiologic principles by which to design and judge such studies has begun to form only in the last two decades. These principles have evolved in tandem with an explosion of epidemiologic activity covering a wide range of health problems.

After the Second World War, the United States initiated many large-scale epidemiologic studies. A number of these studies have had far-reaching influence on the health of today's citizens. For example, the community intervention trials of fluoride supplementation in water that were started during the 1940s have led to widespread primary prevention of dental caries [Ast, 1965]. The Framingham Heart Study, initiated in 1949 to study risk factors for cardiovascular disease, is the most notable of several long-term follow-up studies of cardiovascular disease that have contributed importantly to understanding the etiology of this staggering public-health problem [Dawber et al., 1957; Kannel et al., 1961; Kannel et al., 1970; McKee et al., 1971]. This remarkable study is continuing to produce valuable findings 35 years after it was begun [Kannel et al., 1984] and has undoubtedly contributed essential knowledge that has stemmed the United States epidemic of cardiovascular mortality, which peaked in the mid 1960s [Stallones, 1980]. The largest formal human experiment ever conducted was the Salk vaccine field trial in 1954, with nearly a million school children as subjects [Francis et al., 1957]. This study provided the practical basis for the prevention of paralytic poliomyelitis.

The same era saw the first epidemiologic studies linking smoking and health, eventually leading to the landmark report *Smoking and Health,* issued by the Surgeon General in 1964 [U.S. Dept. of Health, Education and Welfare, 1964]. Since that time epidemiologic research has steadily attracted public attention. Along with a rising tide of social concern about environmental issues and health in general, epidemiologic studies on many subjects have been vaulted to prominence by the news media. Some of these studies were controversial, although the media may have been partly responsible in many cases for fueling the controversy. A few of the biggest attention-getters were studies related to

the efficacy of oral antidiabetic medication
the effect of diethylstilbestrol (DES) on offspring
clustering and infectious transmission of Hodgkin's disease
reserpine and breast cancer
Legionnaire's disease
low-level ionizing radiation and leukemia
saccharin and bladder cancer
swine flu vaccination and Guillain-Barré syndrome
hormonal drugs in pregnancy and birth defects
tampons and toxic-shock syndrome

Bendectin and birth defects
hazardous waste disposal sites
replacement estrogens and endometrial cancer
coffee drinking and pancreatic cancer
passive smoking
Agent Orange
Acquired Immune Deficiency Syndrome (AIDS)

Despite the surge of epidemiologic activity in recent years, there is still abundant evidence that, as a science, epidemiology remains in an early stage of development. In established sciences, one does not find wide disagreement and confusion about the most basic concepts or measures. Whereas physicists agree on the definition of mass or energy, epidemiologists often disagree on definitions of incidence (not to mention the definition of epidemiology itself). In 1975, a paper appeared in the *American Journal of Epidemiology* entitled "Definition of rates: some remarks on their use and misuse" [Elandt-Johnson, 1975]. No revolutionary concepts or definitions were proposed, but the paper was useful because so many readers did not know the definitions of the basic measures used in epidemiology. It is notable that all but one of the introductory texts published in the decade since Elandt-Johnson's paper appeared continue to corrupt the definitions of the basic measures that she discussed. Clear concepts of causation and related ideas such as induction period, like the definitions of basic measures, are fundamental to an understanding of epidemiologic research. Nevertheless, even these underpinnings have not yet been integrated into the conceptual bedrock of the discipline. Disagreement about basic conceptual and methodologic points has led, in some instances, to profound differences in the interpretation of data. In 1978, a controversy erupted about whether exogenous estrogens are carcinogenic to the endometrium: Several case-control studies had reported an extremely strong association, with up to a 15-fold increase in risk, but one group argued that a selection bias accounted for nearly the entire effect [Smith et al., 1975; Ziel and Finkle, 1975; Mack et al., 1976; Horwitz and Feinstein, 1978; Hutchison and Rothman, 1978; Jick et al., 1979; Greenland and Neutra, 1981]. Disagreement and confusion about basic ideas in epidemiology does not necessarily attest to the thick-headedness of epidemiologists; a more charitable interpretation would be that the basic ideas fundamental to the new science have not yet displaced the complacent languor of traditional thinking.

Why has epidemiology been so slow to blossom? The answer lies partially in the difficulty of conducting epidemiologic research. Measures of disease incidence are the basic building blocks of epidemiologic inferences. These measures involve the observation of disease occurrence in relation to the people and time spans in which they occur. This is not a simple process. Typically, disease occurs rarely in the person-time experience (see Chap. 3), so that considerable time and effort are needed to make the basic measurements. Epidemiologists also face the problem of obtaining cooperation from other people to make their observations. The investigator has no control over the "experimental" setting and usually must reckon with limitations imposed by budget and concerns for the privacy of subjects. The end product of such an excruciating and often frustrating exercise is just the first step in accumulating epidemiologic knowledge.

Such difficulties have long discouraged epidemiologic research and will continue to do so. Economies of scale resulting from these observational problems have favored epidemiologic research in settings where medical records and vital statistics are carefully collected and available for use, or where the wealth of society can support the expensive efforts needed to gather the necessary information. The logistic problems encountered in measuring disease incidence have also led to the ascendance of the case-control study as a central tool of modern epidemiology. Case-control research is in many ways emblematic of the modern synthesis of epidemiologic concepts. The methodology of case-control studies has a sound theoretical basis, and as a means of increasing measurement efficiency in epidemiology, it is an attractive option. Unfortunately, the case-control approach has often been misunderstood to be a second-rate substitute for follow-up studies. Only through a firm conceptual grounding in epidemiologic principles can the student of epidemiology see that there is no basis for this derogation of case-control research. Since this type of understanding, covering a wide range, is critical to the successful conduct and interpretation of epidemiologic research of all types, a focus on epidemiologic concepts and methods is crucial to anyone who aspires to understand modern epidemiology.

The past two decades have also seen rapid growth in the understanding and synthesis of epidemiologic concepts. The main stimulus for the growth of theory seems to have been practice: The explosion of epidemiologic activity accentuated the need to improve understanding of the theoretical underpinnings. For example, the signal studies on smoking and lung cancer in the early 1950s were scientifically noteworthy not only for their substantive findings but also because they demonstrated the efficacy and great efficiency of the case-control study [Wynder and Graham, 1950; Doll and Hill, 1952]. Likewise, analysis of data from the Framingham Heart Study stimulated the development of the most popular multivariate methodology used today, multiple logistic regression analysis [Cornfield, 1962; Truett et al., 1967].

The fundamental concepts of epidemiology depend little on other scientific disciplines, nor do they depend on empirical results. Thus, the capacity to formulate a theory of epidemiologic concepts has been possible for centuries; that it is a twentieth century phenomenon is independent of any recent scientific and technical breakthroughs. Rather, the economic development of the prosperous nations in the twentieth century afforded the luxury of conducting epidemiologic research, which in turn motivated the conceptual development that is the scientific "emergence" of epidemiology.

Until recently, virtually all epidemiologists were physicians. Their interest in epidemiology was typically focused on the occurrence patterns of a particular disease. Perhaps because these researchers subordinated an interest in epidemiologic principles to their substantive goals in understanding disease etiology, there was no movement to pursue the development of a theory of epidemiologic investigation. Many epidemiologic investigations, now best forgotten, were designed and conducted poorly for want of such a theory.

Historically, physicians have collaborated fruitfully with statisticians, who contributed expertise in making observations on large populations as well as in data analysis. Much of the theoretical development of modern epidemiology was contributed by statisticians—Cornfield, Mantel, Cox, Breslow, and Prentice are a few of the outstanding contributors. The influence of statistical thinking in epidemiology has not been wholly positive, however. It was natural for statisticians, bringing their skills to bear on epidemiologic problems, to borrow methods with which they were familiar in other areas of application. These methods often became incorporated into epidemiologic practice, not always with a sound basis in theory.

One example of the negative influence of statistical thinking in epidemiologic practice is the dominance of statistical hypothesis testing in epidemiologic data analysis. The motivation for the development of statistical hypothesis testing was to provide a basis for decision making in agricultural and quality-control experiments. These experiments were designed to answer questions that called for specific actions, so that the results had to be classified, if possible, into qualitatively discrete categories. Thus arose the practice of declaring associations in data as "statistically significant" or "nonsignificant," using arbitrary criteria that became conventional. The notion of statistical significance has come to pervade epidemiologic thinking as well as that of other disciplines. Unfortunately, statistical hypothesis testing is a mode of analysis that offers less insight into epidemiologic data than alternative methods that emphasize estimation of interpretable measures.

Another example of the misapplication of statistics in epidemiology has been in the area of multivariate analysis. Statistical methodology in multivariate modeling has often been transferred wholesale to epidemiology without giving sufficient thought to the underlying epidemiologic concepts. Many practices common in multivariate analysis are often inappropriate in an epidemiologic context: The use of continuous independent variates, product terms to evaluate interactions, stepwise algorithms to determine the model, and variance reduction to evaluate the model are all potentially problematic. Multivariate analysis is an important analytic tool

for the epidemiologist, but it cannot be used appropriately without first considering the epidemiologic principles that govern its use. Today, notwithstanding the important contributions to the field by many who consider themselves first as statisticians or physicians, epidemiologists have achieved a separate identity. Being either a physician or a statistician or even both simultaneously is not sufficient qualification for being an epidemiologist. What is sufficient is a theoretical understanding of the principles of epidemiologic research and the experience to apply them.

Epidemiology has established a toehold as a scientific discipline. Whereas epidemiologic results were once greeted mainly with skepticism, they are now generally accorded some degree of respect. At midcentury, epidemiologists had trouble persuading the scientific community of a relation between smoking and lung cancer. By 1984, the situation had changed so much that a weak epidemiologic association observed between beta-carotene and cancer occurrence was the stimulus for a biochemical hypothesis on anti-oxidants, which was published in *Science*. The paper begins with the observation that

> [E]pidemiological studies indicate that the incidence of cancer may be slightly lower among individuals with an above-average intake of beta-carotene and other carotenoids [Burton and Ingold, 1984].

The respectability evinced by this integration of epidemiology into the fold of the biologic sciences stems in large part from the emergence of a clearer understanding of the epidemiologic concepts that have become the basis of modern epidemiology.

REFERENCES

Ast, D. B. Dental public health. In P. E. Sartwell (ed.), *Preventive Medicine and Public Health* (9th ed.). New York: Meredith, 1965.

Burton, G. W., and Ingold, K. U. Beta-carotene: An unusual type of lipid antioxidant. *Science* 1984; 224:569–573.

Cornfield, J. Joint dependence of the risk of coronary heart disease on serum cholesterol and systolic blood pressure: A discriminant function analysis. *Fed. Proc.* 1962; 21:58–61.

Dawber, T. R., Moore, F. E., and Mann, G. V. II. Coronary heart disease in the Framingham study. *Am. J. Public Health* 1957; 47:4–24.

Doll, R., and Hill, A. B. A study of the aetiology of carcinoma of the lung. *Br. Med. J.* 1952; 2:1271–1286.

Elandt-Johnson, R. C. Definition of rates: some remarks on their use and misuse. *Am. J. Epidemiol.* 1975; 102:267–271.

Francis, T., Napier, J. A., Voight, B. S., et al. Evaluation of the 1954 Field Trial of Poliomyelitis Vaccine. Final Report. Ann Arbor: Poliomyelitis Vaccine Evaluation Center, University of Michigan, 1957.

Greenland, S., and Neutra, R. An analysis of detection bias and proposed correc-

tions in the study of estrogens and endometrial cancer. *J. Chron. Dis.* 1981; 34:433–438.
Horwitz, R. I., and Feinstein, A. R. Alternative analytic methods for case-control studies of estrogens and endometrial cancer. *N. Engl. J. Med.* 1978; 299:1089–1094.
Hutchison, G. B., and Rothman, K. J. Correcting a bias? *N. Engl. J. Med.* 1978; 299:1129–1130.
Jick, H., Watkins, R. N., Hunter, J. R., et al. Replacement estrogens and endometrial cancer. *N. Engl. J. Med.* 1979; 300:218–222.
Kannel, W. B., Dawber, T. R., Kagan, A., et al. Factors of risk in the development of coronary heart disease—six year follow-up experience: The Framingham study. *Ann. Intern. Med.* 1961; 55:33–50.
Kannel, W. B., Wolf, P. A., Verter, J., et al. Epidemiologic assessment of the role of blood pressure in stroke: The Framingham study. *J.A.M.A.* 1970; 214:301–310.
Kannel, W. B., and Abbott, R. D. Incidence and prognosis of unrecognized myocardial infarction. An update on the Framingham study. *N. Engl. J. Med.* 1984; 311:1144–1147.
Mack, T. M., Pike, M. C., Henderson, B. E., et al. Estrogens and endometrial cancer in a retirement community. *N. Engl. J. Med.* 1976; 294:1262–1267.
McKee, P. A., Castelli, W. P., McNamara, P. M., et al. The natural history of congestive heart failure: The Framingham study. *N. Engl. J. Med.* 1971; 285:1441–1446.
Smith, D. C., Prentice, R., Thompson, D. J., et al. Association of exogenous estrogen and endometrial carcinoma. *N. Engl. J. Med.* 1975; 293:1164–1167.
Stallones, R. A. The rise and fall of ischemic heart disease. *Sci. American* 1980; 243:(5)53–59.
Truett, J., Cornfield, J., and Kannel, W. A multivariate analysis of the risk of coronary heart disease in Framingham. *J. Chron. Dis.* 1967; 20:511–524.
U.S. Department of Health, Education and Welfare. Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service. Public Health Service Publication No. 1103. Washington, D.C. Government Printing Office, 1964.
Wynder, E. L., and Graham, E. A. Tobacco smoking as a possible etiologic factor in bronchogenic carcinoma. A study of six hundred and eighty-four proved cases. *J.A.M.A.* 1950; 143:329–336.
Ziel, H. K., and Finkle, W. D. Increased risk of endometrial carcinoma among users of conjugated estrogens. *N. Engl. J. Med.* 1975; 293:1167–1170.

2. CAUSAL INFERENCE IN EPIDEMIOLOGY

In *The Magic Years,* Selma Fraiberg [1959] characterizes every toddler as a scientist, busily fulfilling an earnest mission to develop a logical structure for the strange objects and events that make up the world that he or she inhabits. None of us is born with any concept of causal connections. As a youngster, each person develops an inventory of causal explanations that brings meaning to the events that are perceived and ultimately leads to increasing power to control those events. Parents can attest to the delight that children take in forming causal hypotheses and then meticulously testing them, often through exasperating repetitions that are motivated mainly by the joy of scientific confirmation. At a certain age, a child will, upon entering a new room, search for a wall switch to operate the electric lighting, and upon finding one that does, repeatedly switch it on and off merely to confirm the discovery beyond any reasonable doubt. Experiments such as those designed to test the effect of gravity on free-falling liquids are usually conducted with careful attention, varying the initial conditions in subtle ways and reducing extraneous influences whenever possible by conducting the experiments safely removed from parental interference. The fruit of these scientific labors is the essential system of causal beliefs that enables each of us to navigate our complex world.

Although the method of proposing and testing causal theories is mastered intuitively by every youngster, the inferential process involved has been the subject of philosophic debate throughout the history of scientific philosophy. It is worthwhile to consider briefly the history of ideas describing the inductive process that characterizes causal inference, to understand better the modern view and its implications for epidemiology.

PHILOSOPHY OF SCIENTIFIC INFERENCE

The dominant scientific philosophy from the birth of historic scientific inquiry until the beginning of the scientific revolution was the doctrine of *rationalism.* According to this doctrine, scientific knowledge accumulated through reason and intuition rather than by empirical observation. In ancient Greece, the only prominent empirical science was astronomy. Nevertheless, even the observation of the heavens was belittled by Plato, who considered celestial observations an unreliable source of knowledge compared with reason [Reichenbach, 1951]. The highest form of knowledge was considered to be mathematics, a system of knowledge built upon a framework of axioms by deductive logic. The geometry of Euclid exemplifies the rationalist ideal.

Skeptics of rationalism who believed that perceptions of natural phenomena are the source and ultimate judge of knowledge developed a competing doctrine known as *empiricism.* The great pioneers of modern empiricism were Francis Bacon, John Locke, and David Hume. Bacon saw that earlier empiricists, though they exalted empirical science, overemphasized observation to the extent that logic played little role in the accumu-