Essays of an Information Scientist: Science Literacy, Policy, Evaluation, and other Essays, Vol:11, p.290,1988 Current Contents, #36, p.3-8, September 5,1988

Current Comments®

EUGENE GARFIELD
INSTITUTE FOR SCIENTIFIC INFORMATION®
3501 MARKET ST. PHILADELPHIA, PA 19104

Mapping the World of Epidemiology. Part 2. The Techniques of Tracking Down Disease

Number 36 September 5, 1988

Part 2 focuses on epidemiological methods and techniques, including statistical analysis. The essay describes some of the technological advances that have broadened the scope of this discipline. Lists of journals, institutions, and organizations involved in epidemiological research are provided. We also point out that epidemiology deserves better recognition.

In the past century or so, epidemiology, once only the study of the causes of infectious disease, has broadened its scope considerably. Modern epidemiology encompasses all diseases, whether infectious or noninfectious, and is also concerned with controlling and preventing illness. Part 1 of this essay outlined the history of the field and highlighted some of its leading researchers and major breakthroughs.[1] This part focuses on the methods of the epidemiologist and lists institutes and organizations around the world that are involved with epidemiology, as well as the journals that publish epidemiological research.

As noted in Part 1, epidemiologists are concerned with patterns of disease in the human population and the factors that influence those patterns.[1] According to Abraham M. Lilienfeld (1920-1984), Department of Epidemiology, Johns Hopkins University School of Hygiene and Public Health, Baltimore, and David E. Lilienfeld, Division of Environmental and Occupational Medicine, Mt. Sinai School of Medicine, New York, epidemiologists are primarily interested in the occurrence of disease by time, place, and persons; they try to determine whether a specific disease has increased or decreased over time, whether one geographical area is experiencing a higher frequency of a disease than another, and whether the characteristics of those afflicted with a particular disease set them apart from those free of it.[2]

Among the factors that must be considered by epidemiologists are demographic characteristics, such as age, sex, and race; biological characteristics, such as blood levels of antibodies, enzymes, and chemicals; social and economic factors, including educational background and occupation; personal habits, such as the use of tobacco or drugs; and such genetic characteristics as blood type. The general purposes of epidemiological studies are to describe the causes or origins of a disease (its etiology), to formulate hypotheses and test them either clinically (at the bedside) or experimentally (in the laboratory), and, finally, to develop and evaluate procedures that will control or prevent the illness (or to recommend public-health practices).[2]

The "Epidemiological Method"

In Part 1 we discussed the fact that the field of epidemiology has expanded from its traditional interest in infectious diseases to encompass noninfectious diseases as well.[1] There is still considerable debate, however, regarding the definition and limits of the field.

According to Milton Terris, formerly (1960-1964) head of the Chronic Disease Unit, Division of Epidemiology, Public Health Research Institute of the City of New York, and now editor, *Journal of Public Health Policy*, "some epidemiologists have

tended to broaden the definition of epidemiology to the point where it becomes almost meaningless."[3] Terris wrote in 1962 that to claim the existence of an "epidemiological method" applicable to a wide variety of problems outside the scope of epidemiology is tantamount to reducing epidemiology to "a method rather than a field of study."[3] He claims that a wide variety of methods have been found useful in the course of the epidemiologist's work; some were developed by epidemiologists, while others have been borrowed from other sciences, including statistics, microbiology, biochemistry, clinical medicine, geography, demography, and sociology, among others.[3]

The standard measure of sickness in population groups, for example, is statistics; attributes of populations, such as height and weight, can be expressed as means (or averages) and distribution curves. Mathematical models of population groups are constructed from observations and measurements of samples drawn from that population; the reliability of the sample is evaluated using further statistical procedures. And since it is often impossible to precisely measure the influence of one factor upon another on the basis of a single investigation, epidemiologists often use a battery of data sets and analytical techniques in combination to draw conclusions about the occurrence of disease.[4]

In response to Terris's claims, Edward H. Kass, Channing Laboratory, Harvard Medical School, and Brigham and Women's Hospital, Boston, Massachusetts, among others, contends that "it is impossible to separate the methods of a field from the intellectual content of the field.... There is no such thing as *the* epidemiological method, but there certainly is a general body of knowledge and a special mind-set that characterizes most people who call themselves epidemiologists."[5] Haroutune K. Armenian, Department of Epidemiology, Johns Hopkins School of Hygiene and Public Health, responds by saying that "epidemiology is a purposive discipline and applies a variety of methods to elucidate etiology and prevent disease. The dynamism of the discipline is in its constantly changing and improving methodological base. The methods that we will be using in epidemiology in a few decades will be different from what they are today."[6]

Biostatistics and Computer Modeling

As the scope of epidemiology broadened, new methods had to be developed to deal with new or different types of disease. As noted by Samuel W. Greenhouse, Department of Statistics, George Washington University, Washington, DC, among others, when epidemiology expanded from the study of infectious diseases to include noninfectious diseases as well, certain difficulties, both theoretical and methodological, were encountered.[7]

However, thanks to advances in computer science and biostatistics, epidemiologists may now investigate causal and risk factors in chronic disease by participating in large-scale intervention trials to decrease risk factors, surveying entire populations, and conducting longitudinal studies of thousands of people over a period of several years. Case-control techniques also have been crucial in the investigation of such factors. Indeed, Roger I. Glass, Center for Infectious Diseases, Centers for Disease Control (CDC), Atlanta, feels that modern technology "is making epidemiology more interesting and versatile, by providing tools to refine population-based studies of risk factors, disease associations, and causality. This new methodology is already having an impact on the prevention and control of disease."[8]

In the realm of biostatistics, two major contributions were made by Jerome Cornfield, National Cancer Institute (NCI), NIH. In 1951 Cornfield demonstrated a relatively simple method for accurately estimating the relative risk of contracting a disease;[9] about 10 years later Cornfield introduced a method of analyzing tabular material that was more sensitive and accurate than those in use at the time.[10] According to the Lilienfelds, Cornfield's work "ushered in the modern era of case-control studies."[11] However, Greenhouse claims that "the estimate of relative risk is now so commonplace that most application papers no longer reference the Cornfield source."[7] And in

fact, data from the *Science Citation Index®* show that Cornfield's classic 1951 paper has been cited "only" 185 times since 1955; I have discussed the Mertonian "obliteration phenomenon" so often that references to it seem superfluous—the ultimate compliment.

Another major statistical advance came in 1959, when Nathan Mantel and William Haenszel, Biometry Branch, NCI, published their landmark paper, "Statistical aspects of the analysis of data from retrospective studies of disease."[12] In the article the authors discuss the relationship of the retrospective study to the prospective study in the investigation and occurrence of disease. They provide a chi-square test for the statistical significance of an observed association between a disease and a specific factor under study. This classic work, published in the *Journal of the National Cancer Institute*, has received nearly 2,100 citations. (The NIH is one of a number of organizations around the world that are involved in epidemiological research; a select list of these institutes appears in Table 1.) Interestingly, in his *Citation Classic®* commentary on the work, Mantel writes, "In a way, our work was an extension of still earlier work by Jerome Cornfield, who had suggested the effective utilization of retrospective studies."[13]

The publication of some of the first results of what is now commonly referred to as "The Framingham Study" in 1971 represented another advance in the statistical methodology of epidemiology.[14] Published by William B. Kannel and colleagues, Heart Disease Epidemiology Study, Framingham, Massachusetts, and the National Heart and Lung Institute, NIH, it has been cited in almost 850 publications. The study, started in 1949 and still ongoing, is an investigation of the effects of a large number of variables on the risk of developing coronary heart disease, the number-one killer in the US. In his *Citation Classic* commentary, Kannel states that the 1971 study "was one of the largest bodies of data showing the impact of cholesterols and lipoproteins on risk using prospective data."[15] The techniques of multifactorial analysis developed for this study have revolutionized the analysis of epidemiological data. We identified this study in an analysis of highly cited papers from the *Annals of Internal Medicine*.[16]

Table 1: Selected list of organizations and research institutes providing information on and conducting research in the field of epidemiology.

American Cancer Society
Department of Epidemiology and Statistics
4 West 35th Street
New York, NY 10016

American Epidemiological Society
Emory University
School of Medicine
735 Gatewood Road, NE
Atlanta, GA 30322

American Health Foundation
Mahoney Institute for Health Maintenance
Division of Epidemiology
320 East 43rd Street
New York, NY 10017

Centers for Disease Control
1600 Clifton Road, NE
Atlanta, GA 30333

Chaim Sheba Medical Center
Tel Aviv University
Sackler School of Medicine
Department of Clinical Epidemiology
Tel Hashomer
Israel

International Epidemiological Association
University of Pennsylvania
School of Medicine
Room 229 L NEB/S2
Philadelphia, PA 19104

National Cancer Center
Research Institute
Division of Epidemiology
1-1, Tsukiji 5-chome,
Chuo-ku, Tokyo 104
Japan

National Center for Health Statistics
Office of Analysis and Epidemiology Program
3700 East-West Highway
Hyattsville, MD 20782

National Health & Medical Research Council
Unit of Epidemiology & Preventive Medicine
University of Western Australia
Nedlands V1C2 185
WA 6009
Australia

Society for Epidemiologic Research
c/o American Journal of Epidemiology
624 North Broadway, Suite 225
Baltimore, MD 21205

WHO (World Health Organization)
Epidemiology and Statistics Department
Avenue Appia
CH-1211 Geneva 27
Switzerland

Among the first to recognize the need for the mechanized management of data were Sir Richard Doll, then of the Statistical Re-

search Unit, Medical Research Council, London, and A. Bradford Hill, London School of Hygiene and Tropical Medicine, UK, who in the early 1950s began an observational study of the relationship between tobacco use and lung cancer in England.[17] They surveyed 34,440 British male physicians, aged 35 and over, on their smoking habits and followed them over the course of decades to determine the cause of death when it occurred, paying particular attention to deaths from lung cancer. Initially they reported that smokers had 10 times the mortality rate from lung cancer than nonsmokers.[18] In 1976 Doll and Richard Peto, Radcliffe Infirmary, University of Oxford, UK, reported the results of a 20-year follow-up of these doctors.[19] They found that between one-half and one-third of the deaths of cigarette smokers were linked to smoking-related diseases, such as heart disease, lung cancer, and various other cardiac and pulmonary disorders. Without the aid of machines, analysis of the vast amount of data by age, smoking history, and cause of death would have been impossible. We are still awaiting a commentary on this *Citation Classic*.

Breakthroughs in Molecular and Genetic Biology

New techniques that are revolutionizing molecular biology are also being put to use in epidemiology, especially in the study of infectious disease. Kaye Wachsmuth, Division of Bacterial Disease, CDC, points out that genetic probes, DNA hybridization, and DNA-sequence analyses have been used for the rapid identification and characterization of microorganisms causing infectious disease, as well as for genes involved in genetically caused illness and cancer. Genetic-sequence studies have also been used to identify the organisms involved in *Escherichia coli* infections, cholera, and polio.[20]

Such advances have helped bring about the emergence of the discipline "genetic epidemiology." Among the many researchers whose work helped marry genetics with epidemiology was Abraham Lilienfeld, who made both substantive and methodological contributions to the field. Recognizing that family studies could provide the best means of studying the interaction of genetics with environment, he tried to incorporate fundamental epidemiological principles into the methodology of genetics.[21]

According to Muin J. Khoury and colleagues, Department of Epidemiology, Johns Hopkins School of Hygiene and Public Health, the central theme of genetic epidemiology is the study of genetic factors in disease and their interaction with environmental factors. In this discipline, genetics also brings its own tools for counting and analysis, and epidemiology brings methodological principles from widespread medical research. The result is a field that can address medical and public-health issues from a broad perspective.[22] In recognition of this growing field, the journal *Genetic Epidemiology* was launched in 1984. It appears on our selected list of key epidemiology journals, shown in Table 2.

Where Epidemiological Research Is Published and Where It Is Used

Table 2 is by no means exhaustive, but it does represent a fair slice of the journals publishing research related to epidemiology. It is perhaps not surprising that the oldest journal on the list, dating back to 1904, is the *Journal of Infectious Diseases*, since it harks back to epidemiology's roots. The newest journal on the list is *Epidemiology and Infection*, founded in 1987.

Table 2 not only allows us to see where epidemiological research is being published, but also gives us a clue as to who is using it. In an attempt to quantify the research specialties using epidemiological results, Andrew L. Dannenberg, Department of Epidemiology, Johns Hopkins School of Hygiene and Public Health, analyzed citations to and from the *American Journal of Epidemiology* from 1974 through 1982.[23] Dannenberg believes that epidemiological methods have been applied unevenly among medical specialties and that identifying current uses and areas of potential research would help clarify and define the field. In 1985 he reported that internal-medicine and public-health/epidemiology journals accounted for most of the citations. On the other hand, allergy, anesthesiology, and dermatology

journals are among the publications that accounted for few of the citations either to or from the *American Journal of Epidemiology*. Dannenberg concluded that the interchange between clinicians and epidemiologists is adequate and that citation analysis is a useful tool in examining interactions and trends within a field.[23]

Is Epidemiology an Unsung Hero?

We have seen in both parts of this essay the key role that epidemiology has played in the eradication of a number of diseases, in identifying new agents in epidemics before their causes were known, and in the control and prevention of others. Yet the field has been neglected by the recognition system of science. There is no Nobel Prize specifically for epidemiology, nor anything even comparable. The John Scott Award, given to makers of useful inventions,[24] has, on occasion, recognized contributions that have had epidemiological applications. In 1981 the award was given to Benjamin A. Rubin, the inventor of the bifurcated needle that was used by WHO in its worldwide campaign against smallpox; while working on ways to administer the smallpox vaccine, Rubin designed a needle that made the vaccination process quicker and easier.[25] However, this is not a specific recognition of the field of epidemiology. In Part 1 we mentioned that more recently the award was presented to David W. Fraser, now president, Swarthmore College, Pennsylvania, for his part in solving the Legionnaires' disease riddle.[1]

Terris claims that epidemiology has been ignored largely because of an "almost completely therapeutic orientation of the medical and allied health professions, whose attitudes range, with relatively rare exceptions, from sheer indifference to outright hostility toward epidemiology, preventive medicine, and public health.... The public's heroes... are the surgeons and other clinicians who work miraculous cures in hospital settings."[26] But at least part of the problem of the lack of acceptance with which epidemiology has been met lies in the nature of its evidence. Many epidemiological findings are based on statistical analysis, which is hard for many clinicians and laboratory investigators to accept.[26] Many biological scientists are uncomfortable with essentially mathematical concepts, which accounts for much of this reluctance.[5]

Table 2: Selected list of journals reporting on epidemiology. The first year of publication is included in parentheses.

American Journal of Epidemiology (1921)
M. Szklo, ed.
Johns Hopkins University School of Hygiene and Public Health
Baltimore, MD

Bulletin of the World Health Organization (1947)
WHO (World Health Organization)
Geneva, Switzerland

Epidemiologic Reviews (1979)
H.K. Armenian, ed.
Johns Hopkins University School of Hygiene and Public Health
Baltimore, MD

Epidemiology and Infection (1987)
J.R. Pattison, ed.
Cambridge University Press
New York, NY

Genetic Epidemiology (1984)
D.C. Rao, ed.
Alan R. Liss
New York, NY

International Journal of Epidemiology (1972)
C. du Ve Florey, ed.
Oxford University Press
Oxford, United Kingdom

Journal of Clinical Epidemiology (1955)
W.O. Spitzer & A.R. Feinstein, eds.
Pergamon Press
Oxford, United Kingdom

Journal of Epidemiology and Community Health (1947)
J.R.T. Colley, ed.
British Medical Association
London, United Kingdom

Journal of Infectious Diseases (1904)
M.D. Yow, ed.
University of Chicago Press
Chicago, IL

Journal of Public Health Policy (1980)
M. Terris, ed.
Journal of Public Health Policy
South Burlington, VT

MMWR—Morbidity and Mortality Weekly Report (1950)
M.B. Gregg, ed.
Centers for Disease Control
Atlanta, GA

Revue d'Epidemiologie et de Sante Publique (1953)
D. Schwartz & R. Sohier, eds.
Masson
Paris, France

However, with the recent contributions that epidemiologists have made in uncovering the etiologies of many infectious and noninfectious diseases, it is hard to believe that epidemiology will not soon earn more explicit recognition of its central role in modern medical and environmental research. Some indicators would include the memberships of leading epidemiologists in groups such as the National Academy of Sciences and the US Institute of Medicine (IOM). This year, Doll, who, incidentally, received an honorary PhD from Harvard this past June, became a foreign member of the IOM. I've also been told that Alexander D. Langmuir, Department of Preventive and Social Medicine, Harvard Medical School, mentioned in Part 1, and Kass are also senior physicians in the IOM.[5]

That epidemiology is receiving an increasing level of recognition over the past few decades is evidenced by the changes introduced in the titles of some of the journals in the field. The *American Journal of Hygiene* became the *American Journal of Epidemiology* in 1965; the *British Journal of Preventive and Social Medicine* became the *Journal of Epidemiology and Community Health* in 1977; and the *Journal of Chronic Diseases* became the *Journal of Clinical Epidemiology* in 1988.[6]

Etymologists will note that such changes were long overdue, and merely confirm, *a posteriori*, what had happened long ago. One wonders where this exciting field will go in the years to come.

* * * * *

My thanks to Stephen A. Bonaduce and Marianne Zajdel for their help in the preparation of this essay.

REFERENCES

1. **Garfield E.** Mapping the world of epidemiology. Part 1. The disease detectives. *Current Contents* (35):3-11, 29 August 1988.
2. **Lilienfeld A M & Lilienfeld D E.** Laying the foundations: the epidemiologic approach to disease. *Foundations of epidemiology.* New York: Oxford University Press, 1980. p. 3-22.
3. **Terris M.** The scope and methods of epidemiology. *Amer. J. Public Health* 52:1371-6, 1962.
4. **Knox E G, Acheson R M, Anderson D O, Bice T W & White K L,** eds. The methods of epidemiology. *Epidemiology in health care planning.* New York: Oxford University Press, 1979. p. 89-109.
5. **Kass E H.** Personal communication. 26 July 1988.
6. **Armenian H K.** Personal communication. 26 July 1988.
7. **Greenhouse S W.** Jerome Cornfield's contributions to epidemiology. *Biometrics* 38(Supp.):33-45, 1982.
8. **Glass R I.** New prospects for epidemiologic investigations. *Science* 234:951-5, 1986.
9. **Cornfield J.** A method of estimating comparative rates from clinical data. Applications to cancer of the lung, breast, and cervix. *J. Nat. Cancer Inst.* 11:1269-75, 1951.
10. ----------. Joint dependence of risk of coronary heart disease on serum cholesterol and systolic blood pressure: a discriminant function analysis. *Fed. Proc.* 21:58-61, 1962.
11. **Lilienfeld A M & Lilienfeld D E.** A century of case-control studies: progress? *J. Chronic Dis.* 32:5-13, 1979.
12. **Mantel N & Haenszel W.** Statistical aspects of the analysis of data from retrospective studies of disease. *J. Nat. Cancer Inst.* 22:719-48, 1959.
13. **Mantel N.** Citation Classic. Commentary on *J. Nat. Cancer Inst.* 22:719-48, 1959. (Barrett J T, ed.) *Contemporary classics in the life sciences. Volume 2: the molecules of life.* Philadelphia: ISI Press, 1986. p. 260.
14. **Kannel W B, Castelli W P, Gordon T & McNamara P M.** Serum cholesterol, lipoproteins, and the risk of coronary heart disease. *Ann. Intern. Med.* 74:1-12, 1971.
15. **Kannel W B.** Citation Classic. Commentary on *Ann. Intern. Med.* 74:1-12, 1971. (Barrett J T, ed.) *Contemporary classics in the life sciences. Volume 2: the molecules of life.* Philadelphia: ISI Press, 1986. p. 39.
16. **Garfield E.** 101 *Citation Classics* from *Annals of Internal Medicine. Essays of an information scientist: the awards of science and other essays.* Philadelphia: ISI Press, 1985. Vol. 7. p. 374-84.
17. **Doll R & Hill A B.** Smoking and carcinoma of the lung: preliminary report. *Brit. Med. J.* 2:739-48, 1950.
18. ----------. Mortality in relation to smoking: ten years' observations of British doctors. *Brit. Med. J.* 1:1399-410, 1964.
19. **Doll R & Peto R.** Mortality in relation to smoking: 20 years' observations on male British doctors. *Brit. Med. J.* 2:1525-36, 1976.
20. **Wachsmuth K.** Molecular epidemiology of bacterial infections: examples of methodology and of investigations of outbreaks. *Rev. Infec. Dis.* 8:682-92, 1986.
21. **Lilienfeld A M.** Problems and areas on genetic-epidemiology field studies. *NY Acad. Sci.* 91:797-805, 1961.
22. **Khoury M J, Beaty T H & Cohen B H.** The interface of genetics and epidemiology. *J. Chronic Dis.* 39:963-78, 1986.
23. **Dannenberg A L.** Use of epidemiology in medical specialties: an examination by citation analysis. *Amer. J. Epidemiol.* 121:140-51, 1985.
24. **Garfield E.** Since 1816 the John Scott and other Philadelphia awards have recognized "useful" scientific discoveries—James Black and Benjamin Rubin head a list of recent distinguished recipients. *Essays of an information scientist.* Philadelphia: ISI Press, 1983. Vol. 5. p. 686-94.
25. **Rubin B A.** A note on the development of the bifurcated needle for smallpox vaccination. *WHO Chron.* 34:180-1, 1980.
26. **Terris M.** Epidemiology as a guide to health policy. *Annu. Rev. Public Health* 1:323-44, 1980.