Epidemiology: The Basic Science of Public Health

Paul E. Leaverton*

The author examines definitions of public health and epidemiology and briefly traces the history of epidemiology through the 20th century and suggests prospects for the next. He supports the traditional view that the main purpose of epidemiology is to search for and quantify causes of disease with a goal of pointing towards methods of prevention. Some of the classical epidemiological contributions to successful disease prevention programs are described. Also, the author outlines how epidemiological methods changed during the last century as the pattern of diseases dramatically shifted dominance from infectious to chronic disease.

Key words: epidemiology, public health, disease causation

INTRODUCTION

The definitions and relationships of the broad field of public health and the discipline of epidemiology recently have been the subject of public debate and discussion1,2. Although it might be viewed as a semantic argument, on the premise that definitions are important in science, we will elaborate on some of the relevant issues here.

Persons involved with medical training and research throughout the world, for years, have mostly agreed upon the basic sciences for clinical medicine. These are anatomy, physiology, pharmacology, microbiology and biochemistry. But, for public health training throughout the world, the core disciplines are different. In all schools of public health, these include epidemiology and biostatistics. Additional courses in environmental health, management, and behavioral science are almost always added.

The difference between clinical research and epidemiological research should be made at the outset; a distinction previously described by this author in this journal3. The clinician or clinical researcher generally is concerned with making improvements when presented with a sick person. The focus is on curing or improving the lot of a diseased or injured patient. The epidemiologist primarily is concerned with discerning the cause (not the cure) of the condition, whether illness or injury. The ultimate goal is prevention of disease in the future. Admittedly, this is a simplified model but it captures the essential difference. This attention to causes of disease versus cures differentiates public health from (most of) clinical medicine.

The author is with the Department of Epidemiology and Biostatistics, College of Public Health, University of South Florida, Tampa, Florida, USA. He is the current (2001-02) president of the American Epidemiological Society.

PUBLIC HEALTH AND EPIDEMIOLOGY

More than one definition of public health appears in the literature. One that is often cited is attributed to C.E.A. Winslow4, in 1920:

Public Health is the science and art of (1) preventing disease, (2) prolonging life, and (3) organized community efforts for (a) the sanitation of the environment, (b) the control of communicable infections, (c) the education of the individual in personal hygiene, (d) the organization of medicine and nursing services for the early diagnosis and preventive treatment of disease, and (e) the development of the social machinery to ensure everyone a standard of living adequate for the maintenance of health, so organizing these benefits as to enable every citizen to realize his birthright of health and longevity.

Winslow’s definition, though written over 80 years ago, has aged rather well. It is highlighted in the classic 1988 volume, The Future of Public Health5, published by the U.S. National Academy of Sciences’ Institute of Medicine (IOM). The distinguished committee involved in this project considered several definitions of public health and settled on the following statement (page 41):

The committee defines the substance of public health as: organized community efforts aimed at the prevention of disease and promotion of health. It links many disciplines and rests upon the scientific core of epidemiology.

Why has epidemiology been singled out as the core (basic science) of public health? It follows from the fact that public health’s first mission is to prevent disease. To pre-
vent disease, one must have some knowledge about its cause(s) as seen in populations. The study of causes (determinants) of disease in populations is the major function of epidemiology. So, simply by the generally accepted definitions, one must come to the conclusion of the IOM committee.

Another, more recent (1998) perspective by Last, suggests a concise "necessary sequence for the control of any public health problem." He lists:

- Awareness that the problem exists
- Understanding what causes the problem
- Capability to deal with the problem
- A sense of values that the problem matters
- Political will to control the problem

The first two requisites above are obvious. Awareness that the (a) public health problem exists depends upon an effective surveillance system for the target population. Surveillance, such as a reportable disease system, is often considered in the domain of epidemiology—but not always. The second point, understanding cause, clearly is the basis of epidemiology, if the problem is disease related (as is usually the case in public health). The last three important requisites increasingly depend more upon politics and less on epidemiology, or any other science.

Is epidemiology a science?

Is this a rhetorical question? No, sadly there still are scientists today who still believe that a discipline is not a science unless it is based upon experimental (often laboratory) evidence. A major medical journal, when referring to a National Institutes of Health (NIH) conference on prevention, stated that, "What NIH seemed to want was epidemiologic evidence, not scientific evidence." To rebut this argument requires delving into the definition of a science. Over the past centuries every philosopher worth his/her salt has expressed an opinion on this matter. Kant summarized the issue by differentiating between fine art and science when he cited the impossibility of utilizing definite rules in art. Hume's comment, "It seems to me that the only objects of abstract science or demonstration are quantity and number..." is widely accepted. That science is based on categories and numbers further emphasizes the quantification requisite. It is true that most sciences derive their status from experimentation. However, it is not the ability to experiment that defines a science but the ability to quantify the essential entities with which it deals. Astronomy and economics are two other sciences in which experimentation plays a small role. Probably only about 10% of epidemiology is based upon experimental investigations. This important part of epidemiology occurs during the conduct of prevention trials evaluating the causal aspect of risk factors noted in "observational studies". (The author prefers the term "non-experimental" to "observational," not wishing to imply that observations are not made in experiments.)

RECENT DISSENTIONS

There has been a published dialogue on the matter of "the" basic science of public health. Savitz, et al, argue that epidemiology is simply one of many basic public health sciences. They cite clinical medicine, sociology, toxicology, molecular biology, anthropology, nutrition, sanitary engineering, policy analysis, risk assessment, industrial hygiene, economics, and political science as deserving of inclusion. (Curiously, biostatistics, a universally required public health core course, is omitted from their list.) It is true that epidemiology and public health rely on these and other sciences as well. However, the goal of these disciplines includes much more than uncovering the cause of disease and disability, which is the sole goal of epidemiology. The exception in their list is industrial hygiene, which limits its domain to industrial settings. This author is not dissuaded from the traditional view as apparent from the definitional argument cited earlier. Koplan, et al, while acknowledging the evolution of epidemiology to a more complicated enterprise, also maintain the more traditional viewpoint.

This preceding debate could be viewed simply as a matter of semantics and thus rather inconsequential. However, it has become part of a larger, more profound, discussion concerning the proper domain of public health and issues of disease causation. As quoted by Savitz, et al, the Institute of Medicine committee has defined a mission of public health to be, "the fulfillment of society's interest in assuring the conditions in which people can be healthy." This extremely broad perspective puts much of the focus of public health activities into the political arena where poverty and other aspects of environmental conditions may impact public health. The emphasis is on political (community) actions. This view contrasts with epidemiologically determined risk factors that are amenable to individual actions in almost any environment.

Both individual and societal actions clearly can have an impact on an individual's health. Geoffrey Rose's classic 1985 paper, "Sick Individuals and Sick Populations," logically explicates the difference in research goals regarding population risk and individual risk. In a recent commentary, McCormick cautions that, despite Rose's differentiation, "Good evidence of causal relationships must come from the study of individuals." He is rightly concerned with the possibility of the ecologic fallacy; that of wrongly attributing associations observed at the aggregate (population) level being assumed to operate at the individual level.

How far epidemiology and public health should venture into the political arena is a matter of debate. Sate recently has argued vehemently that if public health places too much emphasis on societal issues, it will lose its tradi-
tional focus. Much would be lost if public health defines itself too broadly. She states, "There is no reason, furthermore, why careful research on the relationship between social variables and health should not proceed and quality results published for the enlightenment of professionals and the lay public alike. But the more passionately public health experts pursue social justice, the less effort, time and money they can devote to promoting health for Americans today. With more than half of all deaths resulting from diseases that are preventable or modifiable, it is indeed reckless to downplay the virtues of self-care." Further, "Worse, putting social justice at the core of the public health enterprise undermines individual accountability." This important argument (on appropriate emphases in public health) is likely to continue for the foreseeable future. But this author sees much merit in Satel's warning. Focusing on plainly political matters will dilute the strength of both epidemiology and public health. The study of risk factors operating at the individual level has been the cornerstone of epidemiology's magnificent successes during the past half century. Epidemiology has identified many areas in which individual "lifestyle" choices unequivocally and importantly can affect one's long-term health consequences.

Relatedly, McMichael has argued for epidemiologists to return to their emphasis on population-based studies from their emphasis on individual risk studies. His point is that the entire world needs help in addressing the potentially devastating health problems that may be insidiously emerging from global warming and other man-made environmental effects. And the only way to develop effective surveillance for such possibilities is for epidemiologists to help other scientists in this difficult new arena by increasing their focus on population-level health effects. Leaverton has discussed this persuasive plea in an earlier article in this journal. Of course, this is a scientific issue, unrelated to whether or not persons should become political activists.

We shall next examine how the emphasis shifted during the past century from a population-based approach to (infectious) disease control (as exemplified by successful efforts to vaccinate at-risk populations) to identifying factors related to individual choices to "control" (or reduce the risk of) many chronic diseases.

THE EVOLUTION OF EPIDEMIOLOGY DURING THE 20TH CENTURY

At the outset of the 20th century, most of the health world was focused on infectious diseases. They were the leading causes of death in every country. Continuing the emerging trend from the late 19th century, microbiological success stories were an almost everyday occurrence and there was a strong feeling that this laboratory science would continue its dramatic progress through nearly every important disease. The development of vaccines and the success of vaccine programs were impressive. Indeed, the contribution of vaccines to the conquest of many important infectious diseases constitutes perhaps the most important public health achievement of the past century. The eradication of smallpox from the face of the earth was the pinnacle of that effort. Polio seems likely to be next.

Inexorably, the pattern of disease (at least as measured by mortality) shifted gradually, but dramatically over the century in the United States (and other developed nations) from domination by infectious diseases to domination by chronic diseases by the year 2000. Heart disease, cancers, and stroke became the top three killers, by far. Epidemiologists, led by Abraham Lilienfeld (Johns Hopkins) and Brian MacMahon (Harvard), demonstrated that new methodologic tools could be used to successfully examine factors related to the occurrence of chronic diseases even though the time period between exposures and disease was years, not weeks. The case-control study and the cohort study became the fundamental approaches taught and employed by epidemiologists to examine suspected causative variables (later termed "risk factors") in the development of several chronic diseases in individuals. Long-term cohort studies became the most popular method for studying risk factors in prevalent diseases such as coronary heart disease. Case-control studies became the method of choice for studies of rare diseases such as site-specific cancers. These new designs, along with methods for addressing bias and confounding, became the essence of the new epidemiology. Because of the changing disease patterns, interest in infectious disease epidemiology rapidly waned and chronic disease epidemiology became king.

The notion that a single (microbiological) agent was responsible for a disease was dominant in the infectious disease era. Dubos termed this the "doctrine of specific etiology". It was demonstrably true with infectious disease that if an effective vaccine could be developed and deployed against this single agent, that disease control problem was essentially solved. The reduction in the incidence of such diseases by this mechanism was astounding. However, it gradually became clear that this "doctrine" did not apply to chronic diseases. Heart disease, cancer, neurological and mental disorders were due to a complex mix of both genetic and environmental factors. There was no single causal agent. Multiple causes could have multiple adverse health effects; the several diseases affected by cigarette smoking is the classic example. The new epidemiology was proving not to be an easy enterprise.

The identification of several chronic disease risk factors, and the establishment of their causal role through preventive clinical trials (when possible), has had a tremendous impact on public health practice and the health of populations. Attention to cigarette smoking, diet, and physical activity has escalated dramatically in the U.S. and other nations. The substantial, positive effect on the world's health is undeniable.
Do we really prevent chronic disease?

The public health and epidemiology communities still talk about prevention of disease, a concept that is carried over from the infectious disease era where true prevention could occur through effective vaccine programs. Such "prevention" programs are touted even though reduction in risk for chronic disease usually cannot prevent the disease from occurring, but rather delays the onset. This is not to demean the effort; delaying a first heart attack from age 50 to age 60 is an important contribution. But, this author believes in precise definitions and prefers the more accurate goal of delaying the onset of a chronic disease rather than prevention. Next, we will take a brief look at a few of the remarkable public health success stories resulting from the "new epidemiology" which have delayed substantially the onset of many chronic diseases.

SOME EPIDEMIOLOGIC CONTRIBUTIONS TO PUBLIC HEALTH DURING THE LAST HALF OF THE 20TH CENTURY

The Framingham Heart Study

The prototypical cohort study demonstrating chronic disease epidemiology is the Framingham Heart Study. This long-term investigation, begun in 1948 in Framingham, Massachusetts, continues to this day. A representative (though not fully random) sample of this small community, who were free of coronary heart disease and stroke at the outset, was selected. Every two years this cohort was examined for evidence of disease utilizing a standardized approach to data collection. Although it involved only 5200 men and women, this study changed the face of public health. An excellent, comprehensive website, describing all facets of Framingham, is supported by the National Institutes of Health (NIH)®. A few of the many highlights, by year of publication, are:

- 1960 Cigarette smoking found to increase the risk of heart disease
- 1961 Cholesterol level, blood pressure, and ECG abnormalities found to increase the risk of heart disease
- 1967 Physical activity found to reduce the risk of heart disease and obesity to increase the risk
- 1970 High blood pressure found to increase the risk of stroke
- 1988 High levels of HDL cholesterol found to reduce the risk of death

One of the early Framingham papers (1961) by Kannel and colleagues often is cited as the first scientific publication to use the term "risk factors" in the manner that has come to be accepted in public health. Labarthe, in his superb cardiovascular disease epidemiology overview comments (page 458), "Referring to serum cholesterol concentration, blood pressure, and electrocardiographic evidence of left ventricular hypertrophy, Kannel and colleagues wrote: Combinations of the three risk factors under consideration appear to augment further the risk of subsequent development of coronary heart disease. It has been demon-strated... that the incidence of coronary heart disease rises progressively as these factors are combined." Further, "As additional longitudinal observations are made, it is hoped that additional risk factors will be determined. This will allow further identification of susceptible individuals and hopefully suggest methods of control."

Other population-based studies of individuals and their risk factors emerged, mostly corroborating Framingham's findings. One of the most famous, The Seven Countries Study, recently has been summarized by one of the active U.S. investigators, Henry Blackburn and colleagues. Other American studies were summarized and known as the pooling project®. These landmark epidemiologic investigations changed way persons viewed their susceptibility to cardiovascular disease (CVD). In the middle of the 20th century, the prevailing view was that CVD (by then the leading cause of death) was an inexcusable result of "modernization", a natural occurrence based upon society's rather rapid transition from agrarian to industrialized. Keys (organizer of the Seven Countries Study), and later Blackburn and Rose, became champions of the notion that, because of the great differences among developed nations in CVD incidence rates, there must be something modifiable in the environment. Of course, this has turned out to be the case. Although dependent upon laboratory and clinical evidence for biologic plausibility, the bulk of society's knowledge about the causes (not treatment) of CVD has come from these and other many other epidemiologic studies.

Risk factors, epidemiologically identified as consistently associated with the development of CVD had to be evaluated, when possible, by the use of (preventive) randomized clinical trials. These were needed to clarify whether the risk factors were part of the causal pathway or were merely "ride-along" markers of disease propensity. If this latter situation were the case, then modification of them would make no difference in disease risk. Such trials were (and are) expensive to conduct but many have been convincing enough to persuade the health community about an individual's ability to substantially reduce his/her risk. Many trials of blood pressure reduction and cholesterol lowering have conclusively demonstrated reduced risks of heart disease. There are modifiable risk factors for CVD.
Smoking and Cancer

McGinnis and Foege in their classic 1993 article23, "Actual Causes of Death in the United States", utilized best estimates of risk factor contributions to several causes of death. They estimated that tobacco use (primarily cigarettes) was the number one cause, accounting for 19 per cent of all deaths. This was followed by diet and activity patterns (14%) and alcohol abuse (5%). As illustrated in this paper, cigarettes often have been cited as the number one preventable cause of death in the U.S. Although cigarettes are estimated to take a greater toll in heart disease than with lung cancer (due to the large difference in incidence), the linkage of smoking with adverse health came about primarily because of case-control studies of lung cancer patients during the 1950s and 60s.

Several epidemiologists are credited with the earliest published studies incriminating cigarette smoking as the likely major cause of lung cancer. Wynder46, as well as Doll and Hill23, were among the first. The relative risk (smokers vs. nonsmokers) in men was on the order of 10 to 1, providing much credence to its being a causal risk factor even though a randomized trial could not be conducted, for both ethical and logistic reasons. The surgeon general's landmark 1964 report54 on this issue (Is taking public health action warranted in the absence of experimental proof of causality?) carefully weighed the evidence and concluded that the epidemiological, and other, evidence was strong enough to begin taking steps to warn the public. The evidence against tobacco use (now including heart disease, lung disease and some other cancers) continues to mount every year. The success of this important linkage can be traced to elegant and groundbreaking studies conducted several decades ago. It is one of epidemiology's most important stories.

Other Major Contributions

Many other, less publicized, epidemiological investigations which add to our understanding of disease causation occur with regularity. The contributions are legion and too numerous to detail in this brief overview. These include: folic acid deficiency and birth defects, aspirin and Rey's syndrome, Hepatitis B and cancer of the liver, chemical exposure and various cancers, and a sedentary life-style as a risk factor for the development of CVD and diabetes.

The discovery77, by CDC epidemiologists in the early 1980s, that HIV/AIDS was a sexually and blood transmitted disease nearly a year before the virus was identified by laboratory scientists, is an outstanding example of "shoe-leather" epidemiology. This term is attributed to Alexander Langmuir, the visionary physician/epidemiologist who founded the Epidemic Intelligence Service (EIS) at the Centers for Disease Control (then the Communicable Disease Center) in 1951. It refers to the wearing out of shoe leather while investigating disease outbreaks, which often entail going door-to-door to locate cases and obtain other relevant data.

EPIDEMIOLOGY IN THE 21ST CENTURY

The Impact of the New Genetics

No one can predict the future of any science very far into the future. Nevertheless, there are some real possibilities that seem likely to develop in the field of epidemiology, based upon recent events. One of these events is the cataloging of the human genome. As understanding of this truly monumental scientific achievement grows, it is likely that genetic therapy finally will attain some of its promise as an effective treatment for some diseases that clearly are of genetic origin. However, as we have noted, most of the important chronic diseases are caused by a complex mix of genetic and environmental factors. The gene-environment interaction is a popular phrase for this concept. Cardiovascular disease, most cancers, neurological diseases, and mental disorders may all be so categorized. What this portends, in this author's view, is not less work for epidemiology, but more.

One wonders if many of the risk factor studies will need to be repeated for newly identified genetic types that have starkly different predispositions to certain diseases. New epidemiological studies of the major risk factors in major genetic sub-types will be argued for and some will be conducted. However, this possibility of new studies and trials will prove to be too costly a strategy for every situation. Consequently, statistical modeling, utilizing the best genetic and epidemiological data, will likely attempt to handle the problem, at least until flaws in this approach for some groups are identified. This debate between conducting new population-based studies and the less-expensive modeling approach is likely to heat up among epidemiologists during the next few years.

The Blurring Line

The division between infectious and chronic diseases, that this author and many others have used, seems less clear as the new century begins. Many, if not most, of the major chronic diseases now appear to have some microbiological agent in at least part of their complex causal pathway.

The discovery that most peptic ulcers are strongly related to prior infection with H. pylori bacteria, created a dramatic shift in thinking about this disease; both in prevention and treatment. Several cancers are believed to be related in a similarly causal manner to specific viruses. H. pylori is the only bacteria in this domain, being strongly implicated in the development of stomach cancer. Human papilloma virus (cervical cancer), Hepatitis B and C (liver cancer), and Epstein-Barr virus (Burkitt's lymphoma) are
the most prominent viruses that have been so incriminated to date.

Even coronary heart disease appears to have an infectious disease component in some instances. Evidence is accumulating that chlamydial (and possibly other) infections can cause early insults to arterial walls thus initiating the atherosclerotic process. Of course, this is the pathology that underlies the gradual development of most cardiovascular diseases.

That HIV/AIDS now may be properly categorized both as an infectious and a chronic disease, perhaps best exemplifies the change. These relatively recent infectious disease linkages to major chronic disease conditions create an even more blurred line between infectious and chronic diseases. It has been a useful, though somewhat artificial, distinction. Now, epidemiologists will have to create even more complicated models for disease causation as the new century unfolds and as more such linkages are discovered.

Implementing Epidemiology's Findings

Prevention strategies to ensure that societies can benefit maximally from epidemiologic knowledge have been well described by Rose. He differentiates between a high-risk approach and a population-based approach (these are not mutually exclusive). Disease control programs that target high-risk individuals require (sometimes costly) screening to identify asymptomatic persons at high risk. Population-based programs "treat" the whole target population. This latter strategy attempts to persuade (perhaps augmented by legislation) entire communities to adopt healthier lifestyles. Several extremely well-planned community intervention trials, with risk factor reduction as the goal, have been conducted. Sadly, the results generally have been disappointing. Feinleib, has succinctly summarized the likely reasons. It has proven far more difficult than anticipated to create and sustain improved health behavior either through individual emphasis or community efforts.

This enigma, of not being able to fully utilize knowledge obtained from epidemiological studies on a population basis, lies at the intersection of "applied epidemiology" and health promotion. It will surely occupy many epidemiologists (at least in program design and evaluation) and other public health professionals for the foreseeable future. Society should have the capacity to reduce greatly the burden of CVD and some other diseases if only we knew how to efficiently translate that knowledge into improved health behavior among the majority of citizens. Effective public health programs, based only on today's knowledge, could substantially reduce death and disability in the developed nations and forestall (or even eliminate) the CVD epidemic expected to occur in most of the developing world. There is legitimate concern that many countries will see chronic disease burdens accelerate as the west's unhealthy lifestyles are adopted. To be able to effectively utilize disease risk reduction information, obtained through epidemiological studies during the last century (as well as new information certain to emerge), presents one of the major challenges for epidemiology and all of public health during the next.

REFERENCES