COMMENTARY

Invited Commentary: Circular Epidemiology

Lewis H. Kuller

Circular epidemiology can be defined as the continuation of specific types of epidemiologic studies beyond the point of reasonable doubt of the true existence of an important association or the absence of such an association. Circular epidemiology is an extreme example of studies of the consistency of associations. A basic problem for epidemiology is the lack of a systematic approach to acquiring new knowledge to reach a goal of improving public health and preventive medicine. For epidemiologists, research support unfortunately is biased toward the continued study of already proven hypotheses. Circular epidemiology, however, freezes at one point in the evolution of epidemiologic studies, failing to move from descriptive to analytical case-control and longitudinal studies, for example, to experimental, clinical trials. Good epidemiology journals are filled with very well-conducted epidemiologic studies that primarily repeat the obvious or are variations on the theme. Am J Epidemiol 1999;150:897-903.

Epidemiology is the basic science of public health and preventive medicine. Good epidemiologic studies progress from descriptive to analytical to experimental epidemiology and then to studies of effectiveness leading to prevention programs. Unfortunately, there is a tendency in epidemiology to perseverate at one level of evidence, for example, on one type of study design, without moving forward or, worst yet, to regress to or rediscover the evidence documented from many past studies. Circular epidemiology can be defined as the continuation of specific types of epidemiologic studies beyond the point of reasonable doubt of the true existence of an important association or the absence of such an association. Circular epidemiology is an extreme example of studies of the consistency of associations (1).

There are at least two major reasons for circular epidemiology. First, epidemiology is subcategorized into specific topics. Methodology in one specialized area is often rediscovered in a newly described category of epidemiology. Second, subcategories of epidemiology mature. The tendency is to further rediscover the methodology, that is, the study designs, and the results of past studies.

A basic problem for epidemiology is the lack of a systematic approach to acquiring new knowledge to reach a goal of improving public health and preventive medicine. Our research support unfortunately is biased toward the continued study of already proven hypotheses. A new hypothesis for which there is a lack of substantial prior data is unlikely to be successful in terms of peer review.

Replication by investigators, both believers and non-believers of a hypothesis, of studies among different populations is obviously the pillar of good epidemiologic research. However, circular epidemiology freezes at one point in the evolution of epidemiologic studies, failing to move from descriptive to analytical case-control and longitudinal studies, for example, to experimental, clinical trials. Instead, investigators often identify a population or a methodology and then go round and round in circles. Specific errors or difficult, insoluble methodological issues are suddenly rediscovered. The new study repeats the errors of pre-
Prior to World War II, descriptive epidemiologic studies and clinical epidemiologic studies of dietary intake of cholesterol and saturated fat with risk of coronary artery disease began as a result of animal laboratory studies and clinical epidemiologic studies of genetic familial hypercholesterolemia conducted prior to World War II. Descriptive epidemiologic studies by Keys demonstrated the striking geographic variations in coronary heart disease mortality in relation to dietary saturated fat and total cholesterol levels in the 1950s (5). The International Atherosclerosis Project further documented the geographic variation in atherosclerosis among countries (6) and its association with cardiovascular risk factors. Animal and short-term feeding experiments in the 1960s showed the relationship between elevated blood cholesterol levels with dietary intake of specific-chain-length saturated fats, cholesterol, and polyunsaturated fatty acids (7). These studies resulted in equations such as those by Keys et al. and Hegsted of the relation of the nutrients with blood cholesterol levels (7, 8). Numerous subsequent feeding studies have verified the Keys et al. and Hegsted equations.

Longitudinal studies such as the Framingham Study documented the association between high blood cholesterol levels and blood pressure levels, smoking, and risk of coronary heart disease as early as the 1950s, 40 years ago (9). The National Diet-Heart Study clearly demonstrated the association of diet with cholesterol levels (10). This study also showed that changes in blood cholesterol levels in relation to diet were a function of adherence to the experimental diet. The reduction in blood cholesterol levels was greater in closed than in open populations (10). International migrant studies, such as those conducted in Japan, Hawaii, and California, also documented that an increase in saturated fat and cholesterol in the diet resulted in high blood cholesterol levels and risk of coronary heart disease (11).

Criticism of the diet-heart hypothesis in the 1960s was based on data from the Framingham Study that failed to show an association between diet, as measured by various instruments in observational studies, and either blood cholesterol levels or coronary heart disease mortality (12). In a seminal paper, Jacobs et al. demonstrated that difficulties in measuring dietary intake, and the limited variability of blood cholesterol levels in homogeneous human populations, resulted in a null relation in spite of the substantial, proven experimental association of dietary variables, blood cholesterol levels, and atherosclerosis (13). The Oslo Study Diet and Antismoking Trial documented that a reduction in saturated fat and cholesterol in the diets of persons with high blood cholesterol levels resulted in decreased blood cholesterol levels and risk of coronary heart disease (14). Other dietary trials focusing on both reducing saturated fat and increasing polyunsaturated fat demonstrated reductions in the risk of coronary artery disease. In the United States, there was a substantial decrease in consumption of cholesterol primarily from eggs and saturated fat, specifically a decrease...
in butter and whole milk products and an increase in margarine consumption, and a substantial decrease in blood cholesterol levels and coronary heart disease mortality (15). Subsequent clinical trials including, most recently, the very powerful Statin Drug Trials have clearly documented the great benefits of lowering LDL cholesterol levels to reduce the risk of new and recurrent heart attacks and the extent of atherosclerosis, and at LDL cholesterol levels that include almost 50 percent of the US population (16).

In a circular epidemiology fashion, a new crop of studies in the 1980s and 1990s rediscovered the lack of an observational association between diet and coronary heart disease morbidity and mortality. The studies had all of the pitfalls previously identified by Jacobs et al. (13) and others. Not surprisingly, these studies over time ultimately concluded that, yes, saturated fat, cholesterol, and polyunsaturated fat are important determinants of heart disease. The problems with these studies had been noted for generations. The inability to measure individual exposure to a common source, in this case diet, within a homogeneous population had been recognized from the studies by Goldberger on pellagra (17), in which the primary units of comparison were comprised of groups of persons. Investigation of a common source was the method used in prior studies of air pollution and disease. The unit of measurement was often a city or geographic area rather than a specific individual exposure. In occupational epidemiologic studies, which also examine common-source exposure, the unit of measurement is often a class of workers, as these studies have a limited ability to measure individual exposure within specific groups of workers. The studies by Keys et al. (7) were based on the methods used to study common-source epidemics. They compared nutrient intake and blood cholesterol levels in populations and the subsequent risk of heart attack. The investigators needed a well-defined marker of exposure with less variance within persons as compared to between persons, that is, body burden.

The use of ecologic correlations not based on sound biologic plausibility can result in fallacious associations, often referred to as the ecologic fallacy. Keys, on the other hand, tested the hypothesis of a possible common-source epidemic related to specific dietary nutrients that was based on sound animal experimental and human high-risk, genetic models; he and others then tested their hypothesis in longitudinal studies among these populations that varied in their exposure to the common-source nutrients and then in nutrition-experimental studies that firmly established the diet-heart hypothesis (5).

Many investigations of common-source epidemics therefore depended on the next step to prove the hypothesis, experimental epidemiology. The experiment could be a true randomized trial or often was a natural experiment, such as determining changes in air pollution levels in a community (18), conducting migrant studies, or studying changes in manufacturing processes. Unfortunately, the model of investigating common-source exposures within persons moved to cancer. In spite of numerous studies over many years, there still is no consistent evidence of an association between any major nutrients and cancer risk. This certainly does not mean that the nutrients are not related to cancer but rather that the case-control or longitudinal studies of the association of nutrients with chronic disease and cancer have been conducted inappropriately in populations in which the nutrient intake is homogeneous, thus resulting in erroneous interpretation of the data. Recently, epidemiologists have rediscovered the approaches to studying common-source epidemics and have called for more “macro epidemiology” (19).

Previous studies also stressed that a low blood cholesterol level and decline in blood cholesterol over time is often a manifestation of chronic disease, especially liver disease, infections, and cancer. Suddenly, however, investigators rediscovered low cholesterol levels (20). The adverse effects of low cholesterol became a new research topic. Again, little new information has been generated over time, and the likelihood that lowering the blood cholesterol level causes a person to be killed in an accident seems rather remote (21). The basic problem with these studies is the lack of recognition of the comorbidity, namely, the associated chronic diseases and other health behaviors, as noted in the Commission on Chronic Illness studies of the 1950s (22).

Some epidemiologists decided that studying the diet-heart hypothesis was rather dull and that they needed to find new risk factors. Not surprisingly, they found that people who swallowed all kinds of pills, from beta-carotene to vitamin E to garlic, had a reduced risk of heart attack. It is likely that at least some of these pills have some effect on coronary artery disease, although the study design used to measure specific pill consumption might have been severely limited by selection bias. Previous studies clearly documented that adherence to almost any therapy is associated with a decreased risk of disease. Canner and others, in the Coronary Drug Project conducted in the 1970s, showed that good adherence to consuming placebos was associated with a substantial reduction in the risk of coronary heart disease death (23). Only well-conducted clinical trials will substantiate whether an association exists between risk and benefit regarding many of these agents, often considered the “pill of the week.” It is probably still too early to drink herbal...
A risk factor for older persons (24). These investigators
noted that blood cholesterol levels are not an important
factor in the development of atherosclerosis, a marker of subclinical
disease, and the association of measures of subclinical
disease with cardiovascular risk factors. If, during the
incubation period, development of atherosclerosis is a
function of the level and duration of exposure to cer-
tain blood cholesterol levels, then, in older persons,
lower blood cholesterol levels and a longer duration of
exposure to these levels will result in atherosclerosis,
for example, higher blood cholesterol levels, and a
shorter duration of exposure will lead to atherosclero-
sis at a somewhat younger age.

Infectious disease epidemiologists clearly recognize
the importance of the incubation period. Few epidemi-
ologic studies of chronic diseases have defined the
incubation period of disease, its relation to the risk fac-
tors, and, in particular, measurements of subclinical
disease (25). If a risk factor is measured before onset of
clinical disease, the study is prospective. The patho-
physiologic changes, those caused by the subclinical
disease, are not recognized and are assumed to be a
cause of the clinical disease. We have an outpouring of
papers in which markers of inflammation, endothelial
function, hormonal changes, and psychological
changes are identified as causes of disease rather than
as consequences of "subclinical disease." Neither cancer
nor atherosclerosis has a 1- or 2-year incubation
period.

The classic occupational epidemiologic studies rec-
ognize the importance of the incubation period in rela-
tion to the defined exposure (26). Measurements of
recent exposures are not likely to predict disease. Does it
make sense for epidemiologic studies to demonstrate
that blood cholesterol level is not related to risk of
heart disease in older persons, even though we know
that blood cholesterol level is an important determi-
nant of atherosclerosis and that atherosclerosis is a
major pathophysiologic process leading to clinical
coronary artery disease? The results of clinical trials
(i.e., experimental epidemiology) are consistent with
the fact that reduction of LDL cholesterol levels in
older persons results in a decreased risk of clinical car-
diovascular disease (16, 24).

Epidemiologic studies have suddenly begun to
switch their enthusiasm to genetics. They estimate that
60—70 percent of a specific lipoprotein or disease is
due to "genetic factors" (27). Unfortunately forgotten
are the tremendous international and national varia-
tions in diseases and risk factors and the results of the
migrant studies and of clinical trials and natural exper-
iments. In infectious disease epidemiology, specific
measures of host susceptibility have been inferred but
not identified clearly, and epidemiologists have recog-
nized that genetic variations in both the host and the

Am J Epidemiol Vol. 150, No. 9, 1999
agent can contribute to "cause of disease." The specific reason that only one in a thousand persons exposed to the poliomyelitis virus developed clinical disease is still not understood but, at least in part, must have some genetic basis. Nevertheless, the development of the polio vaccine was remarkably successful. Host susceptibility was considered a continuum from subclinical disease (infection and immunity) to clinical disease (paralysis) to bulbar polio and finally death (28).

The need for good measures of host, agent, and environment has come full circle in the generation of new genetic epidemiologic studies of chronic diseases. In this "new epidemiology," better definitions of host susceptibility, for example, genotype, can be determined. The interrelations between environmental exposures and specific agents have often been lost. Thus, we have gone around in a circle, from studies in which we could measure the environment and the agents fairly well by using relatively crude determinations of host susceptibility to studies that attempt to define in great detail the genotype and host susceptibility but not measures of environment and specific agents.

Are we now at the threshold of being able to provide a genetic host susceptibility scorecard for each person that will define the exact dose of a specific agent that the person can consume per day (e.g., 80 vs. 100 mEq of sodium/day) to prevent elevated blood pressure? Too bad that we cannot measure the agent (sodium) or the outcome (individual "blood pressure") with the precision required for this careful definition of a host susceptibility scorecard. Failure to recognize the broad gradient of susceptibility to most agents and of environmental exposure to disease is generating a new round of studies that may delay implementation of effective preventive medicine programs.

Should Salk have waited for the specific host-susceptibility genetic markers to be identified before developing polio vaccines? Similarly, is it necessary to identify the specific genes associated with elevated LDL cholesterol before intervening successfully at the individual and population level to reduce LDL cholesterol levels and the risk of atherosclerosis? We must avoid going around in a circle from prior good studies of agents and environment and poor measures of host susceptibility to good measures of host susceptibility and poor measures of agent and environment. The public health approach still has the greatest impact on reducing morbidity and mortality.

To survive as an important discipline, epidemiology must be the basic science of preventive medicine and public health (29). The emphasis in epidemiologic research must be to link the causal chains to provide a scientific foundation for preventive programs (30). Epidemiology cannot succeed by describing new methods of analyzing the same types of data sets or by rediscovering the important social ills of society (31).

Epidemiologic studies should cover the breadth from phase I exploratory studies, often called descriptive epidemiology, to phase IV clinical trials and evaluation of the application to policy and programs. Public health and preventive medicine not based on solid science, especially epidemiology, is very likely to be unsuccessful and even hazardous to the populations involved. Taking the fruit-, vegetable-, vitamin-, or toxic-chemical-of-the-month-club approach does not benefit public health.

On the other hand, epidemiology that does not focus on how it applies to public health and preventive medicine will not be supported in the future. The need to define the implications of an epidemiologic study in publications should be encouraged rather than discouraged. Epidemiologic papers should clearly describe hypotheses and implications of results in terms of the direction of future studies and the potential importance of results to the practice of public health, prevention, and clinical medicine. Papers that include only caveats about more limitations of the implications of their studies should receive a very low priority. The role of a current study in the evolving process of epidemiology, from descriptive to experimental epidemiology, should be clarified. In addition, the paper should present the future direction of epidemiologic research, not just replicate the current study in different population, race, and sex groups.

Alternative hypotheses and criticism of existing dogma clearly should be encouraged. Epidemiologic studies should be based on sound scientific methodology and not just replication of previous errors. Previous study design problems that have already been solved should be documented carefully in the new studies. Replication of the same methodological problems and errors is circular epidemiology and will not move the field of epidemiology forward.

There is a very substantial risk that erroneous results of epidemiologic studies based on repetition of poor methodology—circular epidemiology—will be used to deter effective public health and preventive medicine programs. For example, industry has been very effective in using observational epidemiologic studies that have failed to demonstrate a relation between the amount of salt in the diet, especially from processed foods, and blood pressure levels to prevent a reduction in the amount of salt in processed foods (32).

Epidemiology is in part dependent on new technology and concepts of pathophysiology provided by other disciplines. The use of new technology may overcome some of the limits of epidemiologic studies,
and these studies should not hesitate to include information on new advances in scientific methodology. Examples include molecular genetics, noninvasive methods of measuring atherosclerosis and cancer, and new techniques of measuring hormone levels, growth factors, immune function, clotting, and thrombosis. Epidemiologic studies can also point other scientific disciplines in new directions (33).

We should continually monitor the successes and failures of epidemiologic studies in improving the "public health." The heroes of epidemiology identify a problem (descriptive epidemiology), develop methods to test specific hypotheses (observational epidemiology), experiment (clinical trials) to prove or disprove a hypothesis, and then apply good public health and preventive medicine strategies to enable the information they have acquired to be used to reduce morbidity and mortality. They do not stop at one point in their search for information and replicate the same study over and over again. Their solution is not to develop more extensive regression analysis that adds yet another variable to a long list of variables or to define a new method that might reduce the variance and change the relative risk from 2.3 to 2.4 but rather to succeed in documenting the benefits of their research in terms of improving the public's health.

The future of epidemiology is very bright if we continue to stress that it is an important basic science of preventive medicine and public health (34-40) and that the study designs that link host, agent, and environment variables to a long list of variables or to define a new method that might reduce the variance and change the relative risk from 2.3 to 2.4 but rather to succeed in documenting the benefits of their research in terms of improving the public's health.

REFERENCES


Am J Epidemiol Vol. 150, No. 9, 1999