Workshop on Guidelines to the Epidemiology of Weak Associations

Introduction

ERNST L. WYNDER, M.D.

American Health Foundation, 320 E. 43rd Street, New York, New York 10017

Epidemiology, from the days of scurvy, smallpox, and childbed fever to those of lung cancer and coronary artery disease, has made the definitive contributions to identifying causative factors in disease—an identification that contributed to appropriate public health measures and disease reduction. Epidemiology, therefore, is not only a science concerned with the discovery of risk factors that contribute to human disease but is also the basis from which preventive approaches are being launched: Its ultimate goal is the eradication of those diseases affected by specific causative factors.

Modern day epidemiology has continued to make major inroads on our understanding of disease and to suggest appropriate preventive measures. Efforts with respect to smoking and lung cancer, alcoholism and cirrhosis of the liver, and dietary fat and atherosclerosis serve as good examples where there has been agreement in findings and in public health positions—in part because the associations were of large magnitude. At the same time, epidemiology has problems when the associations are of a low order of magnitude. In such instances, findings in the literature are, in general, inconsistent. There are many situations in the present literature where weak associations have been reported or where a weak association might have been missed that involves a potentially important public health question requiring confirmation.

When risks are small, and especially when effects occur many years after their causes, detecting them, estimating their magnitude, and assessing their importance for the community in light of other relevant factors pose problems of study design, data collection, analysis, and interpretation which can be exceedingly difficult.

This report is concerned with weak associations between some factor and disease and how we should interpret them. It is not easy to define quantitatively what is meant by “weak,” but Cornfield’s view that any relative risk of under 3 might be considered weak is reasonable. The important point to note, however, is that the closer the risk of some association comes to unity, the more likely it is that choice of the comparison standard, bias, confounding, or inappropriate analysis may explain it and the greater the need for thorough understanding of the underlying biological mechanisms.

The emphasis of the workshop is on observational rather than experimental aspects of epidemiology. We should not, however, ignore the light that the application of preventive measures can throw on disease etiology. The history of beri-beri, pellagra, goiter, and lead poisoning provide classic historical examples, while the elimination of the risk for nasal and lung cancer found in the nickel refining process by cleaning up the industry provides a more modern illustration.

Our aim is to present a series of guidelines which, if conscientiously followed, would prevent the drawing of false conclusions from epidemiological data. The main principles of design and conduct of epidemiological studies are described by Szklo. Of the two analytic types of study—case-control and cohort—our focus stresses the former. Comparison of diseased persons with suitable controls is a simple efficient, and economical way of learning about the causes of determinants of disease and establishing risks. Such studies will often be conducted by physicians and other persons who may be relatively inexperienced in epidemiological methods.

One of the first questions that must be answered by anyone intending to conduct a case-control study concerns the source of cases and controls. Hospitals provide a convenient source of cancer patients. Most cancer cases reach hospitals eventually; thus hospitals offer a reasonably representative sample of the disease. Hospitals also provide an accessible source of patients with other diseases who can be selected as controls. In many case-control studies of cancer, both cases and controls are drawn from hospital patients. An alternative source of controls often has been suggested. Patients in hospitals are not a representative sample of persons living in the community. It has been argued, therefore, that it would be better to compare hospital patients not with hospital controls, but with a more representative group of controls drawn from the general community. Controls drawn from the neighborhood in which the cases live have, therefore, often been selected, as they have been seen as providing a more satisfactory comparison group. While there may be some substance to the view that “neighborhood” controls are more representative of persons in the general community than are patients in hospitals, a case-control study is concerned more with the comparability of cases and controls than with their representativeness. There is good reason to believe that hospital controls are more comparable with hospital cases in attitude, and therefore in the way they may be expected to respond to questions, than are neighborhood controls. Patients in hospitals have time on their hands to reflect on the origins of their illnesses. They are also the subjects of frequent questioning by physicians and others about their habits, behavior, past exposures, dietary peculiarities, and so on. Persons engaged in the everyday activities of life in the community have no cause to speculate on these topics and are not repeatedly questioned about them. Thus, there is a serious risk that by comparing hospital cases with neighborhood controls, ascertainment bias may be introduced into the recording of past experience. The fact of being admitted to hospitals is likely to make hospital controls more similar to the cases than healthy persons living in the general community. When interpreting the findings, it is particularly important to consider the kind of ascertainment bias that might have been introduced into a study by the way in which the controls were selected.
Other forms of bias that may occur in case-control and cohort studies are discussed by Feinleib.

Confounding, whereby a suspect factor is correlated with a true causative factor and also with the disease under investigation, represents a problem in many epidemiological studies. Despite every effort to allow for confounding, one may be left with the possibility that one has failed to allow adequately for those factors one knows to be possible confounders. Moreover, one needs to recognize that there may be other confounding factors of which one may not be aware. Confounding is reviewed by Stellman.

One of the most prevalent practices in epidemiology is to continue to explore negative results. A study is done to test an hypothesis. No surprising results emerge: The study is negative. However, by analyzing sufficient subgroups after the fact, the investigator may succeed in finding one where there is an association between a factor and disease. There is, of course, nothing wrong in such an exploration of data. However, mistakes can arise in the subsequent presentation of results which may be published as significant new findings instead of tentative suggestions that further studies are needed. The question of use and abuse of subgroups is covered by Stallones.

Epidemiology is, above all, concerned with the integration and interpretation of the evidence. This means that the literature must be comprehensively reviewed and the evidence carefully weighed. Selective review of the literature may result in the citation of only those papers that support the investigators' hypothesis. Studies reaching a contrary conclusion may be ignored. Although the criteria for determining cause and effect were well presented and discussed in the first Surgeon General's Report on Smoking and Health in 1964 under "Criteria of Judgment" and are taught in every course in elementary epidemiology, they are frequently not pursued as an investigator describes a particular association. The intellectual process of integrating the data is presented by Schlesselman.

We hope that this symposium will provide a guide to future epidemiological studies. If our advice is heeded and our recommendations followed, such studies should establish associations where these exist and reject them where they do not with more success in the future than has sometimes been the case in the past. Epidemiology is a key science contributing to the understanding of disease causation and, if properly exercised, can also provide the opportunity to achieve disease prevention.