Causation, bias and confounding: a hitchhiker's guide to the epidemiological galaxy
Introductory comments

Samuel Shapiro

Evolution of epidemiology

In the evolution of modern medicine epidemiology has from the beginning played an important role in exploring the causes of disease. To give just a few historical examples,[1] epidemiological methods have shown that lime juice prevents scurvy; that washing one's hands (especially after performing an autopsy) prevents puerperal sepsis; that water contaminated with sewage causes cholera; that thiamine deficiency causes beriberi; and that injections with unsterile needles cause serum hepatitis. Sometimes epidemiology has identified 'causes' long before the mechanisms were understood. Epidemiology has also interacted with related disciplines such as clinical medicine, pathology, toxicology or microbiology in the elucidation of the aetiology of diseases as diverse as leukaemia[2] or malaria.[3]

Yet, despite many such achievements, and despite gradual improvements in methods, in the early years of the 20th century, epidemiology – or more broadly, public health – began to fall out of favour. That decline occurred largely under the impact of a paradigm, propagated by authorities such as Flexner, that decreed that causality can best be elucidated by the basic sciences,[4] and that any work done away from the laboratory bench qualifies, at best, as second class. To some extent that paradigm remains pervasive today. Despite it, however, in the past half-century the application of epidemiology to the exploration of causality has grown exponentially, and a great many findings now enjoy wide acceptance. This change has come about for several reasons.

First, it is now appreciated that any hypothesised causal (or protective) effect, whether generated in the laboratory or elsewhere, must be shown to be valid in humans.

Second, as the 20th century progressed, it became clear that new epidemics were taking place (e.g. coronary heart disease, lung cancer, breast cancer) and that they had environmental components to their aetiology that needed to be understood.

Third, following the advent, mainly after the Second World War, of an unprecedentedly large array of drugs, devices and new medical procedures, there was a need to develop methods to assess both the benefits and risks associated with their use. The outcomes at issue were clinical outcomes, and clinical judgment was central to their evaluation. However, clinical judgment, alone, was insufficient, and susceptible to error.

Fourth, as epidemiological methods began to enjoy wider application, it was appreciated afresh that bench science, unguided by research in human populations, could sometimes fail to identify important causal relationships (e.g. smoking and lung cancer).[5] Indeed, in that instance it took many years for basic research to catch up.

Fifth, some of the most important diseases that now needed to be investigated (e.g. coronary heart disease) appeared to have multiple causes (roughly synonymous terms are 'multifactorial aetiology' or 'web of causation').[6] To cope with the complexity, appropriate statistical methods were needed. Multivariate methods that could to a greater or lesser extent allow for and discriminate among multiple risk factors ('risk indicators', 'third variables', 'covariates') already existed, or were being developed or refined. Initially multivariate methods had been used mainly in the social sciences, and when early attempts were made to apply them in epidemiological research they commonly proved to be infeasible, mainly because they were too time-consuming and expensive. Fortunately, however, rapid advances in computer technology soon enabled epidemiologists to undertake far more complex analyses than had previously been possible.

Epidemiology today and in the future

Today epidemiology is established as a leading discipline in causal research, and now it is also appreciated that such research is best conducted by close collaboration among all relevant disciplines – basic and applied – and that causality (or protection) is most strongly supported when several lines of investigation all converge on the same relatively invariate association. Striking recent examples are the association of *Helicobacter pylori* infection with peptic ulcer disease and gastric cancer[7] and HIV infection with Kaposi's sarcoma.[8] When appropriate, a further important function of epidemiology has been to demonstrate with reasonable confidence when a given exposure is not associated with an alleged increase in risk (e.g. the allegation that calcium channel blocking drugs cause cancer).[9]

There can be no doubt that in the future epidemiology will continue to play a prominent role in causal research. However, its very success has at the same time engendered certain paradoxes. One paradox is that under the impact of technical advances in epidemiological and statistical methods, and in the methods used in controlled trials, the limitations intrinsic to those methods are often being lost sight of. Causal research in epidemiology was originally conducted mostly by clinicians who were evaluating clinical phenomena, and who recognised that the inherent uncertainty of clinical judgment, as applied to any individual patient, inevitably translated to some degree of uncertainty when applied to quantitative clinical evidence in a group of patients. Today much of the research is carried out by persons with little or no clinical insight, sometimes with unfortunate, and biased, results.

A related paradox is that while refinements in methods have improved precision, the findings yielded by those methods have also come more and more to be represented as being more definitive than is justified by the evidence. In addition, some epidemiologists seem to have lost sight of the principle that good science is sceptical science (not nihilist, but sceptical), that good scientists have a duty to do

Department of Epidemiology, University of Cape Town, Cape Town, South Africa
Samuel Shapiro, FCP(SA), FRCP(E), *Visiting Professor of Epidemiology (also Emeritus Director, Slone Epidemiology Center, Boston University, Boston, MA, USA)*

Correspondence to: Professor Samuel Shapiro, Department of Public Health and Family Medicine, University of Cape Town, Cape Town, South Africa. E-mail: samshap@mweb.co.za

everything they can to falsify their own hypotheses, and that they should only reluctantly embrace them, and then only tentatively, when they fail to do so.[10] They should also always be at pains to draw attention to the defects in their own research.

And still a further paradox is that epidemiological methods are now represented as being so complex, and opaque, as to be beyond the comprehension of the uninitiated: only high priests initiated into the arcana can interpret the evidence for the laity. And a consequence of that representation is that a large volume of poor research, camouflaged by the arcana, is now published.

These developments are unfortunate. One of the strengths of epidemiology is its simplicity – and with all the advances that have been made, fundamentally it remains a simple discipline. When evidence is presented that seems at first sight to be impenetrable, there is no reason why a clinical audience should be intimidated. Clinicians should be able to decipher the evidence for themselves. My purpose in this series of articles is to demystify some of the arcana, and to assist the Journal's readers in doing likewise. The series of articles is organised as follows: following these introductory comments, Part 1 (in this issue of the Journal) commences with definitions and methods relevant to causal research, and then moves on to principles of causality.[11] Part 2 (July 2008 issue) will cover additional principles of causality: confounding, effect modification and strength of association.[12] Part 3 (October 2008 issue) will cover further principles of causality: statistical stability, dose- and duration-response effects, internal consistency, external consistency, analogy and biological plausibility, and will put forward some conclusions.[13]

References
1. Sigerist HE. *A History of Medicine.* New York, NY: Oxford University Press, 1999.
2. International Agency for Research on Cancer (IARC). *Ionizing Radiation, Part 1: X- and Gamma Radiation and Neutrons (IARC Monographs on the Evaluation of Carcinogenic Risk to Humans Vol. 73)*. Lyon, France: IARC, 1999.
3. Ross R. *The Prevention of Malaria.* London, UK: John Murray, 1910.
4. Flexner A. *Medical Education in the United States and Canada Bulletin Number Four (The Flexner Report)*. Stanford, CA: The Carnegie Foundation for the Advancement of Teaching, 1910.
5. International Agency for Research on Cancer (IARC). *Tobacco Smoke and Involuntary Smoking (IARC Monographs on the Evaluation of Carcinogenic Risk to Humans Vol. 44)*. Lyon, France: IARC, 2004.
6. MacMahon B, Pugh TF. *Epidemiology. Principles and Methods.* Boston, MA: Little Brown, 1970.
7. Goodman K, Cockburn M. The role of epidemiology in understanding the health effects of *Helicobacter pylori. Epidemiology* 2001; **12**: 226–271.
8. International Agency for Research on Cancer (IARC). *Immunodeficiency Viruses and Human T-cell Lymphocyte Viruses (IARC Monographs on the Evaluation of Carcinogenic Risk to Humans Vol. 83)*. Lyon, France: 1996.
9. Rosenberg L, Rao RS, Palmer J, *et al*. Calcium channel blockers and the risk of cancer. *JAMA* 1998; **279**: 1000–1004.
10. Popper K. *The Logic of Scientific Discovery.* London, UK: Routledge, 1959.
11. Shapiro S. Causation, bias and confounding: a hitchhiker's guide to the epidemiological galaxy. Part 1. Principles of causality in epidemiological research: time order, specification of the study base and specificity. *J Fam Plan Reprod Health Care* 2008; **34**: 83–87.
12. Shapiro S. Causation, bias and confounding: a hitchhiker's guide to the epidemiological galaxy. Part 2. Principles of causality in epidemiological research: confounding, effect modification and strength of association. *J Fam Plan Reprod Health Care* 2008; **34**: (in press).
13. Shapiro S. Causation, bias and confounding: a hitchhiker's guide to the epidemiological galaxy. Part 3: Principles of causality in epidemiological research: statistical stability, dose- and duration-response effects, internal and external consistency, analogy and biological plausibility. *J Fam Plan Reprod Health Care* 2008; **34**: (in press).