

2003

Causation: An Epidemiologic Perspective (In Five Parts)

Douglas L. Weed

Follow this and additional works at: <https://brooklynworks.brooklaw.edu/jlp>

Recommended Citation

Douglas L. Weed, *Causation: An Epidemiologic Perspective (In Five Parts)*, 12 J. L. & Pol'y (2003).

Available at: <https://brooklynworks.brooklaw.edu/jlp/vol12/iss1/3>

This Article is brought to you for free and open access by the Law Journals at BrooklynWorks. It has been accepted for inclusion in Journal of Law and Policy by an authorized editor of BrooklynWorks.

CAUSATION: AN EPIDEMIOLOGIC PERSPECTIVE (IN FIVE PARTS)

*Douglas L. Weed, M.D., Ph.D.**

I. THERE ARE TWO SIDES TO EVERY QUESTION—*PROTAGORAS*

How much scientific evidence does it take to claim causation? What kinds and characteristics of evidence are needed to claim that an exposure causes a disease?

Epidemiology appears to be uniquely positioned to answer these questions. Causation, after all, is an integral part of this key public health discipline. As the first half of a common definition states: “Epidemiology is the study of the distributions and (causal) determinants of disease in populations.”¹

Epidemiology’s search for disease causation has been a long one. For nearly two centuries, we have examined why populations suffer from cholera and tuberculosis, pellagra and scurvy, heart disease and cancer, dementia, suicide, and AIDS, to name a few examples. Significant progress has been made in our understanding of why populations get sick and how their health can and has been improved.² Indeed, the main reason epidemiologists study

* Douglas L. Weed, M.D., Ph.D., is Dean, Education and Training, and Chief, Office of Preventive Oncology, and Director, Cancer Prevention Fellowship Program, Division of Cancer Prevention, National Cancer Institute, Rockville, Maryland.

¹ DICTIONARY OF EPIDEMIOLOGY 62 (John M. Last ed., 4th ed. 2001).

² ROBERT BEAGLEHOLE & RUTH BONITA, PUBLIC HEALTH AT THE CROSSROADS: ACHIEVEMENTS AND PROSPECTS 13-43 (1997); CHRONIC DISEASE EPIDEMIOLOGY AND CONTROL 1-26 (Ross C. Brownson et al. eds., 2d ed. 1998); COMM. FOR THE STUDY OF THE FUTURE OF PUBLIC HEALTH, INSTITUTE OF MEDICINE, THE FUTURE OF PUBLIC HEALTH 1-18 (1988); JOHN DUFFY, THE SANITARIANS: A HISTORY OF AMERICAN PUBLIC HEALTH (1992).

causation is to use scientific knowledge to control disease (and injuries) through preventive interventions. As the full definition in the *Dictionary of Epidemiology* states, “Epidemiology is the study of the distributions and (causal) determinants of disease in populations *and the application of this study to control health problems.*”³ Identify causes, remove them from the environment, and prevent disease: this is the time-honored central mission of epidemiology and all other public health disciplines.

What epidemiologists do *not* do is study disease causation in order to assign responsibility for harm caused to individuals; *specific causation* is not a traditional problem for epidemiologists.⁴ For judges, legal scholars, and others involved in toxic tort litigation, however, the problem of specific causation is paramount. Binding together these two views—one from the world of epidemiology, the other from the law—is that both require an answer to the problem of general causation. Put another way, there are two sides to the questions of causation posed above, two very different reasons for answering the same question: one for public health decisions, the other for legal decisions.

My purpose in this paper is to describe how epidemiologists make claims about general causation, how they *practice* causal

³ DICTIONARY OF EPIDEMIOLOGY, *supra* note 1, at 62 (emphasis added). Definitions of epidemiology have undergone significant transformation in the past thirty years. In the 1970s and 1980s, epidemiology was often defined strictly as a science without reference to the public health application of the knowledge gained by scientific study; in the 1990s, definitions of epidemiology have emphasized the dual role of professional practitioners in scientific investigation and in public health interventions. DOUGLAS L. WEED, EPIDEMIOLOGY AND ETHICS IN EPIDEMIOLOGY, *in* ETHICS AND EPIDEMIOLOGY 76-94 (Steven S. Coughlin & Tom L. Beauchamp eds., 1996); Douglas L. Weed & Robert E. McKeown, *Science and Social Responsibility in Public Health*, ENVTL. HEALTH PERSP., 1804, 1805 (2003); Douglas L. Weed & Pamela J. Mink, *Roles and Responsibilities of Epidemiologists*, ANNALS EPIDEMIOLOGY 67, 68 (2002).

⁴ LEON GORDIS, EPIDEMIOLOGY 286-87 (2000); Michael D. Green, D. Mical Freedman & Leon Gordis, *Reference Guide on Epidemiology*, *in* FED. JUDICIAL CTR., REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 333 (2d ed. 2000), available at [http://www.fjc.gov/public/pdf.nsf/lookup/sciman00.pdf/\\$file/sciman00.pdf](http://www.fjc.gov/public/pdf.nsf/lookup/sciman00.pdf/$file/sciman00.pdf).

PROVING CAUSATION

45

inference. I will identify some important problems that exist in that practice, and what the future holds for solving them.

II. HABIT IS THE ENORMOUS FLYWHEEL OF SOCIETY—*WILLIAM JAMES*

What follows is a brief description of the practice of causal inference in epidemiology, with the following simplifying assumptions:

1. Scientific evidence to be assessed has been made available through a systematic literature review.
2. A statistical association between the exposure (the purported cause) and the disease (the purported effect) has been established at a level of significance of $p < 0.05$.
3. All epidemiologic studies examined have measured all *known* confounders (an unreasonable assumption in many situations, but helpful for the purposes of this brief discussion).
4. Evidence from a randomized prevention trial is not available for the exposure-disease association under scrutiny.
5. A quantitative meta-analysis could be carried out, but will not be.

This situation well approximates what Austin Bradford Hill—of Hill's causal criteria fame—faced in his now-classic paper on causal inference,⁵ published one year after the 1964 Surgeon General's *Report on Smoking and Cancer*, another key paper on causal inference as it has been conceptualized and practiced in epidemiology.⁶

From a systematic literature review, different types of scientific evidence would emerge, including but not limited to laboratory-based “biological” studies, as well as several types of epidemiologic studies, e.g., case-control and cohort. Epidemiologists typically consider the potential biases in the

⁵ Austin Bradford Hill, *The Environment and Disease: Association or Causation?*, 58 PROC. ROYAL SOC'Y MED. 295, 295-300 (1965).

⁶ SURGEON GENERAL'S ADVISORY COMM. ON SMOKING & HEALTH, U.S. DEP'T OF HEALTH EDUC. & WELFARE, PUB. NO. 1103, SMOKING AND HEALTH (1964) [hereinafter 1964 SURGEON GENERAL'S REPORT].

results of such studies as well as the potential role of *unknown* confounders.⁷ If we can assume, however, for the sake of brevity in this description, that bias is not likely to be responsible for the results of these studies, then the practitioner of causal inference moves to the next step in the process: the examination of the summarized evidence in terms of Hill's causal criteria. Excluding "experimentation" (see assumption number 4 above), there are eight such considerations:⁸

1. Consistency
2. Strength of association
3. Dose response (or biological gradient)
4. Biological plausibility
5. Coherence
6. Temporality
7. Specificity
8. Analogy

These so-called causal criteria—of which only "temporality" is the only true criterion—are then "applied to" or "considered in the light of" the evidence.

The first decision to be made by the user is the *selection of the*

⁷ Jennifer L. Kelsey, Diana B. Petitti & Abby C. King, *Key Methodologic Concepts and Issues*, in APPLIED EPIDEMIOLOGY 35 (Ross C. Brownson & Diana B. Petitti eds., 1998); Kenneth J. Rothman & Sander Greenland, *Precision and Validity in Epidemiologic Studies*, in MODERN EPIDEMIOLOGY 115-134 (2d ed. 1998).

⁸ These so-called "Hill's criteria" emerged from a conversation in the medical and public health literature that can be traced back at least as far as the early nineteenth century in the works of Jakob Henle and Robert Koch, both of whom discussed the nature of causation in terms of infectious diseases. See Alfred S. Evans, *Causation and Disease: a Chronological Journey*, 108 AM. J. EPIDEMIOLOGY, 249, 249-58 (1978); ALFRED S. EVANS, CAUSATION AND DISEASE (1993). The extension of this discussion to the causation of chronic diseases—cancer, heart disease, mental illness, and diabetes, to name a few examples—began around 1950 and continues through today. The practice of causal inference, however, remains firmly rooted in the criteria proposed by Austin Bradford Hill in 1965. DOUGLAS L. WEED, *Causal and Preventive Inference*, in CANCER PREVENTION AND CONTROL 285 (Peter Greenwald, Barnett S. Kramer & Douglas L. Weed eds., 1995); Douglas L. Weed & Lester S. Gorelic, *The Practice of Causal Inference in Cancer Epidemiology*, 5 CANCER EPIDEMIOLOGY BIOMARKERS & PREVENTION 303, 303-311 (1996).

PROVING CAUSATION

47

criteria. In cancer epidemiology, for example, the most likely choice involves: consistency, strength, dose-response, and biological plausibility, leaving behind coherence, specificity, analogy, and (interestingly) temporality. In each individual application, the user will select those he believes to be most relevant.

After some written narrative discussion, typically a paragraph or two for each of the criteria selected, the user of this method then makes a claim about the extent to which the exposure and the disease under question are causally related. Sometimes recommendations regarding preventive interventions are also included.

This is a bare-bones—but reasonably accurate—account of the epidemiologic approach to causal inference. This is our habit, our way of solving a very important professional problem.

III. ONLY THE WEARER KNOWS WHERE THE SHOE PINCHES—*OLD ENGLISH PROVERB*

This description of the use of causal criteria could have included the following, more complete, set of steps:

1. Selection of the Criteria (as mentioned above)
2. Prioritization of the Criteria Selected
3. Assigning a Rule of Inference to each Criterion

Studies of the practice of causal inference have shown that epidemiologists rarely pay attention to the second and third steps above.⁹ It is unfortunate but true that a practitioner can undertake this practice precisely as described above¹⁰ without mentioning to those who review, edit, and eventually read the causal assessment

⁹ See Weed & Gorelic, *supra* note 8, at 303-11; Douglas L. Weed & Stephen D. Hursting, *Biologic Plausibility in Causal Inference: Current Method and Practice*, 147 AM. J. EPIDEMIOLOGY 415 (1998); Douglas L. Weed, *Epidemiologic Evidence and Causal Inference*, 14 HEMATOLOGY/ONCOLOGY CLINICS OF NORTH AMERICA 797 (2000); Douglas L. Weed, *Methods in Epidemiology and Public Health: Does Practice Match Theory?*, 55 J. EPIDEMIOLOGY & COMMUNITY HEALTH 104 (2001) [hereinafter Weed, *Methods in Epidemiology*].

¹⁰ See *supra* Part II.

in the peer-reviewed scientific literature, how these criteria are being prioritized (other than some are included and others are not). Similarly, it is exceedingly unusual for a practitioner to describe the rules of inference assigned to each criterion prior to their application. By “rules of inference” I mean the conditions under which one will accept or not accept the criterion as having been satisfied, or more *likely* satisfied. If, for example, the user of the method believes that causation is extremely unlikely if the summarized risk estimate—the relative risk estimate across the studies collected—is less than 2.0, then a reasonable rule of inference for that criterion, for that user, in that particular circumstance, would be that “relative risk estimates less than 2.0 will be considered unlikely causal” or something along those lines. The rules of inference used for the criterion of biological plausibility are especially mysterious. Other criteria can be similarly described.

In sum, the current user of causal inference methods in epidemiology can select the criteria they wish, prioritize them in any manner they wish, and assign rules of inferences to them (implicitly) without ever mentioning them. While it is possible to infer these various choices by careful reading, they need not be stated anywhere in the paper.

It is also true—and at least as unfortunate—that in many scientific journals, these sorts of causal assessments can occur without a systematic review of the literature. The studies selected for review may not, in these circumstances, represent those culled from a larger set using stated inclusion and exclusion criteria.¹¹

Add to this ever-accumulating pile of subjective features the uncertainties stemming from loosening the overly simplistic assumptions regarding statistical significance, confounding, and bias, along with the potential (and documented) role of personal, social, moral, and political values in decision making, and it is fair

¹¹ Douglas L. Weed, *On the Use of Causal Criteria*, 26 INT’L J. EPIDEMIOLOGY 1137 (1997); Douglas L. Weed, *Methodologic Guidelines for Review Papers*, 89 J. NAT’L CANCER INST., 6-7 (1997); Rosalind A. Breslow, Sharon A. Ross & Douglas L. Weed, *Quality of Reviews in Epidemiology*, 88 AM. J. PUB. HEALTH 475 (1998).

PROVING CAUSATION

49

to say that the current practice of causal inference is, at best in trouble, and at worst in shambles.¹²

IV. ALL PROGRESS IS PRECARIOUS, AND THE SOLUTION OF ONE PROBLEM BRINGS US FACE TO FACE WITH ANOTHER PROBLEM—
MARTIN LUTHER KING

Given the current situation, what is most impressive about epidemiology's role in the identification of potentially preventable causes of illness and injury is that so much scientific and public health progress *has* been made. Smoking is indeed a cause of lung cancer, laryngea cancer, esophageal cancer, and bladder cancer. Human Papillomavirus does cause cervical cancer and HIV causes AIDS. The list of chemical carcinogens—*asbestos, arsenic, aniline dyes, diethylstilbestrol, and cadmium to cite a few examples*—is long. Radiation of many types is responsible for—*causes skin cancer, breast cancer, and other diseases*. Put another way, I do not want to give the reader of this brief paper the impression that the methods of causal inference are irremediable. But serious problems we do have.

What is to be done? Two approaches for improving the situation can be identified: one empirical, the other theoretical.

A. An Empirical Approach to Improving the Current Practice of Causal Inference

There are examples—call them “case studies”—of causal associations in the historical record (i.e., the peer-reviewed scientific literature) about which we can all agree on the outcome. For each, we can describe the evidence—the studies and their evidentiary characteristics—that existed at the time a causal claim was first made or, alternatively, at the time a consensus about causation was reached.

Smoking and cancer represents an excellent example. In 1964, a committee of scientists organized by the Surgeon General of the

¹² Douglas L. Weed, *Underdetermination and Incommensurability in Contemporary Epidemiology*, 7 KENNEDY INST. ETHICS J. 107 (1997).

United States carefully collected, summarized, and examined the evidence (in a manner remarkably more systematic than much of what is published today in 2003!).¹³ They concluded that lung cancer and laryngeal cancer were caused by smoking cigarettes. Esophageal cancer and bladder cancer, however, were spared this conclusion. In 1982, after eighteen years of additional research, a new committee was formed, again under the auspices of the U.S. Surgeon General. The same causal criteria from 1964 were applied to a new (expanded) body of evidence with somewhat different evidentiary characteristics. In the judgment of the committee, esophageal cancer joined the ranks of those caused by smoking.¹⁴

How can this type of analysis assist us in improving the practice of causal inference? Here's just one example: Careful study of the 1964 decision on esophageal cancer will allow us to describe what evidence—as reflected in the causal criteria and their rules of inference—was *insufficient* to make a causal claim. Careful study of the 1982 decision on esophageal cancer (in which the committee changed its mind about causation) will provide an estimate of the cumulative amount and minimum characteristics of evidence required to make a causal claim (for that committee).

In any such example of what we now consider to be a case of “known” causation, the extent to which the observed level of evidence—the kinds and characteristics of evidence—is representative of other causal associations is a fair question. Perhaps it would have been reasonable to claim causation with less (perhaps much less) evidence. Nevertheless, such an approach can provide empirical examples of the minimum level of evidence for causation aligned with a particular exposure-disease causal combination.

B. A Theoretical Approach to Improving the Practice of Causal Inference

Alternatively, we may approach the research problem by

¹³ 1964 SURGEON GENERAL'S REPORT, *supra* note 6.

¹⁴ DEP'T OF HEALTH & HUMAN SERVICES, PUB. NO. 82-50179, THE HEALTH CONSEQUENCES OF SMOKING: CANCER (1982).

PROVING CAUSATION

51

examining the causal criteria, their rules of inference, and their prioritization of the criteria, in terms of connections to causal definitions, hypotheses, and theories. Simply put, for a given causal hypothesis or definition, what evidentiary conditions would be expected? Very little work has been done on this question in epidemiology.

On the other hand, setting out causal conditions based on what we believe to be reasonable assumptions is a time-honored approach in epidemiology and can be traced to the early discussion of this topic in the late 1950s and early 1960s. Hill's 1965 classic paper is an excellent example of this approach.¹⁵ Each criterion is considered separately. For example, the reasoning for consistency goes like this: a causal association should be observed in different study populations, using different methods, examined by different investigators; it would be *consistent*, in other words. Here's another example: a causal association is a *plausible* association; to put it another way, if a purported relationship goes against what we know is biologically possible, then we would be less inclined to call it causal. Similarly, a cause must precede its effect in time, the essence of temporality. This approach is extremely popular in epidemiology. Nearly everyone has their opinion on how the criteria should be used. And these opinions are reflected in the high level of subjectivity or personal preference discussed earlier in this paper.

Greater objectivity can be achieved if the criteria can be linked with general causal hypotheses (or general theories of causation or causal definitions).¹⁶ Currently, we have several definitions of causation to work with.¹⁷ Finally, the theoretical approach can be linked with the case-based empirical approach described above. Together, these two approaches are our best hope for progress in this difficult yet critical arena.

V. MORNING COMES WHETHER YOU SET THE ALARM OR NOT—

¹⁵ Hill, *supra* note 5, at 295-300.

¹⁶ Weed, *Methods in Epidemiology*, *supra* note 9, at 104-10.

¹⁷ Mark Parascandola & Douglas L. Weed, *Causation in Epidemiology*, 55 J. EPIDEMIOLOGY & COMMUNITY HEALTH 905 (2001).

URSULA K. LEGUIN

In the end, my comments on general causation have been as much about the fuzzy future as they are about the dizzy present or the hazy past.

Epidemiologists have only recently recognized that their practice of causal inference is seriously ill. For nearly fifty years, ever since the so-called chronic diseases—cancer, heart disease, diabetes, and mental illness emerged as foci for “modern” epidemiology—we have (almost systematically) ignored how we might improve the practice of causal inference. A few recent studies have shown that this practice, after years of neglect, suffers from a variety of ailments not uncommon for methods more qualitative than quantitative, yet which aspire to provide more-certain-than-not results.

The methods of causal inference are often used uncritically and are subject to unacceptable levels of subjectivity. Their results—the causal claims and preventive recommendations—are susceptible to the whims of personal preference, what philosophers call “values.” Precisely opposite claims have emerged from investigators using the same causal criteria on the same evidence.

It is no exaggeration to say that any epidemiologist who claims he is an expert—that he can reliably make claims about causation—is either hopelessly naïve or a flagrant prevaricator. As noted earlier, I do not mean to suggest that prior claims about what factors or exposures cause illnesses are incorrect. We—the public health community—have made the right call in many situations. What I *am* suggesting, on the other hand, is that we have failed to use the past record of achievement in general causation to the public’s advantage. Couple our reluctance to look back and gain from our experience—our successes and failures—with our well-known aversion to theoretical development, and it is not surprising that we have made so little progress on a problem so central to our discipline.

What is needed for causation in epidemiology—and the cliché is unintended—*is* more research. Two approaches have been described: one, like both legal and moral reasoning, emerges from careful empirical study of recorded case studies; the second is a

PROVING CAUSATION

53

theoretical approach both more speculative and potentially more generalizable.

The future comes to everyone at the same pace: sixty minutes per hour. To best meet that future for the problems of general causation, we must have the power to shape it through research and its application, what epidemiologists do best.