Contents

Introduction: Evolution of Epidemiological Concepts and Design

Evolution of concept, design, and method in CVD epidemiology has been driven by the cardiovascular researchers who began to work among populations and to confront major issues with unfamiliar approaches. They had to deal with imprecise measures and unaccustomed masses of data. They clamored for help to analyze multiple related variables, calling on experts to address issues of reliability and validity of measurement, of bias and confounding.

By error and by intent the early investigators stimulated the evolution of guidelines to causal inference from statistical associations. Through their early industry and insistence, both general and particular issues of epidemiological concept and method were advanced.

Theorists provided them tools for handling measurement variation, regression-dilution bias, regression toward the mean, multivariable relationships, and statistical summaries of evidence. The “place” of ecologic correlations was elevated through greater understanding of their complementary role to individual correlations. The rigid hierarchy among observation and experiment was eroded as study design and methods improved. The requirements of meta-analysis and systematic summaries were built into initial designs rather than struggled with as afterthoughts.

Stepping outside the laboratory and clinic among newly minted epidemiologists strengthened rather than diminished the search for mechanisms. And it became not only essential that epidemiology be a part of any successful organizational admixture of research, but that disciplines interact and cross-fertilize. Increasingly accepted as a basic science, epidemiology became as well the estimator of generalized effects, attributable risk, and the final arbiter of effectiveness, and applications of medical research. Intellectual equity established, co-equal administration and funding became a battle cry.

Major steps in the evolution of concepts, design, and method for CVD epidemiology are simply listed, along with a unique story about the development of qualitative guidelines to causal inference from statistical associations, which greatly strengthened epidemiological analysis in the 1960s.

Evolution in CVD epidemiological concept, method, and mission

- From case series, to case-control, survey, cohort-prospective observations, and trials
- From single to multi-factor risk prediction and prevention strategy
- From single to multi-disciplinary research on causality and prevention
- From individual to cross-cultural comparisons of causes and risk of disease
- Advance of causal inference in analysis of observed associations (by Koch-Henle, Yerushalmy, Palmer, Sartwell, Lilienfeld, Hammond, Hill, Popper, Evans, Rothman, et al.), approaching the strength of experiments
- Novel methods, strategies, and populations are enlisted for carrying out mass surveys and preventive trials
- Statistical summaries and meta-analysis of evidence are advanced and planned, a priori
- CVD epidemiology becomes the strength and handmaiden of genetics in the modern drive toward "personalized medicine" and away from studies related to the public health
- Epidemiology develops the life course concept and study of disease and health
- It evolves from secondary to primary to 'primordial' prevention of high risk in the first place
- It complements the medical strategy with a population-wide approach to prevention, from a focus on affluent populations to a global epidemiology of health inequities

Among the stories of the evolution in concept and method in chronic disease epidemiology is a dramatic one about development of the guidelines to causal inference from observed associations. Their practical logic remains useful in epidemiological thinking about causality, though it they are now complemented by more quantitative strategies.

We begin with a truism from Henry James: "The whole of anything is never told."

Guidelines to Causal Inference from Epidemiological Associations: Origins and Early Evolution, 1957-1965

The guidelines to causal inference from statistical associations widely used in chronic disease epidemiology today sought to adapt the shortcomings for observational science of the classic Henle-Koch causal postulates (1). The criteria are attributed to two sources: the formulation published in Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service in 1964 (2), which helped that committee conclude that cigarette smoking caused lung cancer, and that published the following year as Environment and Disease. Association or Causation?, a more extensive but similar version from the President's Address to the Occupational Health Section of the Royal Society of Medicine by pioneer British statistician-epidemiologist, Austin Bradford Hill (3).

Our story recounts a little-known but critical series of formative events and exchanges among experts who, in fact, created the guidelines employed by the U.S. Advisory Committee, a story constructed from the serial published dialogue of U.S. thinkers in which the criteria were honed, and from personal

experience with some of the principals involved. In our pursuit of the history of CVD epidemiology, we noted and were puzzled that Bradford Hill's criteria, published the year following the widely quoted U.S. Surgeon General's Report, made attribution neither to the elegant formulation by the Surgeon General's Advisors nor to the prior published exchanges among the American scientists who developed those criteria between 1957 and 1964. Then we noticed with equal surprise that the Advisory Committee to the Surgeon General also mysteriously "produced" the criteria they used to such advantage in their systematic deliberations on smoking, also without acknowledgment of their origins and evolution to the state they found so useful in their causal deliberations about cigarette smoking.

These constitute the main parts of our story, in which we bring a ray of light from original sources about this joint UK and U.S. omission of professional courtesy and proper attribution. This tale intends to complement more scholarly accounts of how the guidelines have since evolved to a more quantitative and "unified concept" for causal analysis, applicable both to communicable and non-communicable diseases (4,5). It ignores subsequent debate about rationale and logic and whether the guidelines are subjective and "unscientific" (6,7). And it also illustrates the intimate participation in the formative process by several pioneer actors and thoughtful leaders in the development of modern epidemiology, among which were our long- term and close colleagues.

Ancel Keys incited the ire

The sequence of events on the North American side seems to have originated with a January, 1953 lecture given by Minnesota physiologist, Ancel Keys, on "A Newer Public Health." It was delivered before an apparently sympathetic audience of pioneers in CVD epidemiology, including Frederick Epstein and Ernst Boas, at the Mt. Sinai Hospital of New York City, and published the same year in that hospital's Journal (8). In this early presentation, Keys reviewed the clinical, bench, and population evidence supporting the diet-blood lipid-atherosclerosis relation and proposed a hypothesis about modifiable causes of the newly recognized epidemic of heart attacks.

Keys's extensive, multi-causal argument included a curvilinear ecologic correlation he found among 6 countries for national fat consumption data from the Food and Agriculture Organization of the United Nations (FAO) and data on cardiac death rates from the World Health Organization (WHO) (Figure 1, below). He chose data available from nations in which he, as an expert in international nutrition and food data, had most confidence in their quality as "fully comparable dietary and vital statistics data" (8, pg. 133). The crude ecologic association was a relatively minor piece of his wide-ranging review and reasoned argument, followed by modest claims: "that dietary fat somehow is associated with cardiac disease mortality at least in middle age" (8, pg.134).

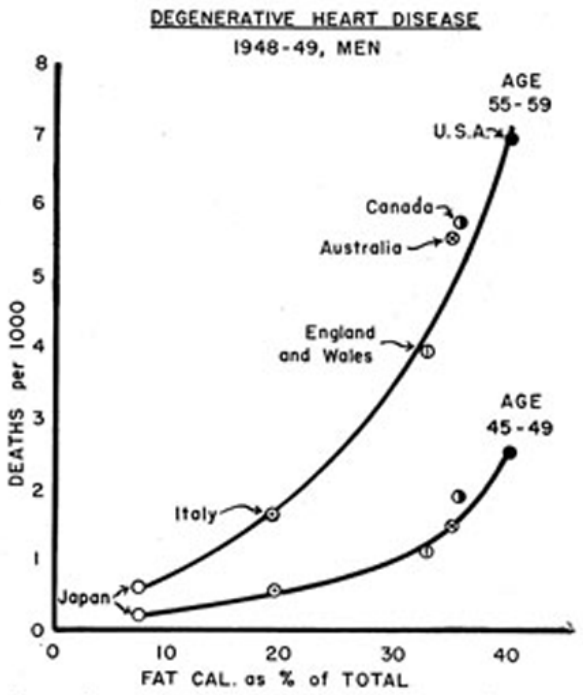

Figure 1: "Mortality from degenerative heart diseases (categories 93 and 94 in the Revision of 1938, categories 420 and 422 in the Revision of 1948), International List. National vital statistics from official sources. Fat calories as percentage of total calories calculated from national food balance data for 1949 supplied by the Nutrition Division, Food and Agriculture Organization of the United Nations." From the Journal of Mount Sinai Hospital (4). Reproduced with permission.

Herman Hilleboe and Jacob Yerushalmy lit the fire

As we learned from participants in the first meeting of experts on atherosclerosis called by the World Health Organization (WHO) in Geneva in fall 1955, Keys presented there a less complete and more bluntly self-assured version of his argument, including the fateful figure of the 6-country ecologic correlation. He was challenged in Oxfordian debate style by Regius Professor George Pickering, who skillfully set the debater's trap, asking Keys to cite the single most important piece of evidence in support of his hypothesis. Keys fell into the trap. His one piece, rather than a whole coherent body of evidence, was readily diminished by the wily debater.

Keys presentation also irritated others of the WHO panel: Herman Hilleboe, then New York State Commissioner of Health, and Jacob Yerushalmy, professor of statistics at the University of California, Berkeley. They later wrote, with acerbic, third-person formality: "At that meeting, statements were made about the association between heart disease mortality and fat in the diet. As it later developed, these were based on a few selected countries and were of questionable validity. On their return to the United States the authors reviewed the available data carefully, and the results indicated that the subject required further study" (9, pg. 2344).

Hilleboe and Yerushalmy brought home the vulnerability of Keys's arguments and made the 6-country display an example of bias; of "how the impression that there is a strong association between the mortality from heart disease and proportion of fat available in the diet in different countries assumed the stature of a proved fact" (ibid). They were to use Keys's crude correlation in a polemic about bias,

requirements for supplemental evidence, and causal criteria applicable to such epidemiological associations

Two years later, at the 1957 meeting of the prestigious American Epidemiological Society, Yerushalmy and Hilleboe unleashed a dramatic critique of Keys's by-then-remote presentation. They had found the diet fat-heart disease association greatly reduced when tested in data available, without regard for quality, among all 22 nations reported by the FAO. They also found that the association with diet fat consumption was not "specific;" it was equally strong for dietary protein and heart disease and for diet fat and other diseases than cardiac. But they treated the 6-nation ecologic correlation as the sole basis of Keys's diet-heart argument, effectively demolishing its validity for that purpose, and implying that the 6 countries were selected with bias. They soon published this critique as a primer on the limitations of observational evidence for causal inference, a lecture to the naïve (9). (Keys's 6-country correlation has since been widely conflated, both in the medical literature and lay press including Wikipedia, with the much-later cross-cultural prospective study of traditional lifestyles and heart attack by Keys and international colleagues, the Seven Countries Study (10)).

Yerushalmy's and Hilleboe's logic in respect to the ecologic correlation was, nevertheless, impeccable, while their focus on that association, ignoring the whole of Keys's other evidence and argument, was unbalanced. By the time their article appeared, Keys and his scientific colleagues, and much of the research world, had moved on to improved methods and new evidence, with broader vistas about lifestyle, biology, and heart attacks.

The sharp and patronizing tone of the Yerushalmy-Hilleboe critique brought a cloud over the mood of those in the Laboratory of Physiological Hygiene at the University of Minnesota on the day the journal arrived:

> It is well known that the indirect method merely suggests that there is an association between the characteristics studied and mortality rates and, further, that no matter how plausible such an association may appear, it is not in itself proof of a cause-effect relationship (9, pg. 2343) . . . At the present time it is possible to conclude only that international statistics on diet and mortality are not sufficiently sensitive to contribute materially to our knowledge concerning their relationship" (9, pg. 2352).

A dialectic began

From this contentious beginning, a rich and protracted "conversation" arose about the interpretation of associations found in epidemiological observations, with a series of editorial articles and letters-to-the-editor replies. In the ensuing process, guidelines for causal inference in epidemiology were progressively honed.

Yerushalmy fanned the flame

This was led off in 1959, again, by Yerushalmy, and soon joined by a string of his mentors, colleagues, and relatives. First, with his Johns Hopkins colleague, Carroll Palmer, came a systematic elaboration of what they termed the "proper handling" of the classical causality issues evoked by observational associations: selection bias and confounding:

The major weakness of observations on humans stems from the fact that they often do not possess the characteristic of group comparability, a basic requirement which in experimentation is accomplished by conscious effort through randomization. The possibility always exists, therefore, that such association as observed may. . . be due to factors other than those under study (11, pg. 28) .

Their guidelines at this early stage of formality proved important in the overall development, particularly by inclusion of these elements:

> The suspected characteristic must be found more frequently in persons with the disease in question than in persons without the disease; or persons possessing the characteristic must develop the disease more frequently than do persons not possessing the characteristic; an observed association between a characteristic and a disease must be tested for validity by investigating the relationship between the characteristic and other diseases and, if possible, the relationship of similar or related characteristics to the disease in question. . . In general, the lower the frequency of these other associations the higher is the specificity of the original observed association and the higher the validity of the causal inference (11, pg. 39).

The new article by Yerushalmy and Palmer referred to Bradford Hill's earlier thinking on observation versus experiment in determining causation, from his 1953 Cutter Lecture (12), and to Abraham Lilienfeld's discussion of smoking as a possible cause of lung cancer (13). And once again in this seminal article, Yerushalmy made a particular example of Keys's precipitating event: the ecologic correlation of diet fat and heart disease, writing without qualification: "this original association [of diet fat and cardiac mortality] must be considered nonspecific and cannot be used in even partial support of a supposed causal relationship" (11, pg. 38).

Abe's wise questions turned the damper down

Later that same year, 1959, in correspondence to the same journal, Abraham Lilienfeld, an established academician at Johns Hopkins, responded to Yerushalmy and Palmer, questioning their insistence on specificity as a guide. He advised that cases occurring without the characteristic under consideration do not necessarily invalidate a hypothesis though may weaken it. And he added that greater frequency of the characteristic in those without the disease also fails to invalidate the hypothesis, due to "accessory factors" affecting susceptibility (14).

For example, Yerushalmy's and Palmer's specificity criterion would, erroneously, not allow for a characteristic that induces multiple diseases, as later was found true, for example, for smoking and for

diet and alcohol. This argument Lilienfeld later used powerfully in an editorial: "The Case Against the Cigarette" (15).

Lillienfeld's thinking on the evidence required to arrive at causal inference is thought to have dated from his and Palmer's collaboration on childhood mortality issues (16). During the 1950s and '60s, Lillienfeld refined his causal thinking with his brother-in-law Yerushalmy. "Abe and Yak" (for Jakob Yerushalmy) debated the cigarette smoking-lung cancer relationship heatedly, "usually in the haze of smoke from Yak's cigarette and Abe's pipe" (16, pg. 512).

Sartwell fine-tuned the flame

The next printed response to this on-going discussion appeared in the same journal in 1960. Here Philip Sartwell, a Johns Hopkins colleague of Lilienfeld, added characteristics he felt would strengthen causal inference: replication of the association with different investigators and populations; finding a graded effect, a quantitative relationship between the intensity or frequency or exposure and the frequency of the disease; a chronologic relationship in which the characteristic must precede the disease; and, finally, "biologic reasonableness" of the association. The latter was always to be considered but left suspect, because judgmental (17).

Hammond buttressed the rationale

Unreferenced in this exchange, but contemporaneous with, and likely familiar to all the U.S. participants, were the thoughts of Cuyler Hammond, chief statistician of the American Cancer Society and Yale professor of biometry. Citing R.A. Fisher, he plumped for the suitability in medicine and biology of a probability approach to cause and effect and argued for taking into account the total situation of "causative factors" that increase the probability of developing a disease (18).

In both prime approaches of science, observation and experiment, Hammond considered these characteristics important: the quantity and duration of exposure, the "degree" of association, the consistency of the relation, and the need for supplemental evidence according to strength of the relation. He also dealt with parallel time trends between exposure and disease and the potential for multiple disease effects of a given causal factor. And he emphasized the remoteness of opportunities for experiment in humans, hence the need for systematic evaluation of observational evidence.

Hammond was early to argue that despite lack of certainty, "in human affairs, important decisions must necessarily be based upon the preponderance of evidence" (18, pg. 194).

Thus, from the heat of debate a burnished set of guidelines had emerged.

Then, in the early 1960s, due to the urgency of advice needed for the Report of the U.S. Surgeon General on Smoking and Health, causal criteria were taken up as a national priority by other, independent and presumably unbiased minds.

Stallones provided order to the spectrum

In 1963, Reuel "Stony" Stallones, then Dean of Public Health at the University of California, Berkeley, was named consultant to President Kennedy's initiative establishing the Advisory Committee to the Surgeon General on the role of tobacco in health. For evaluation of epidemiologic evidence in his assignment on smoking and CVD, Stallones presented the committee a draft of the causal criteria (19), and the following preamble to the national report:

> Statistical methods cannot establish proof of a causal relationship in an association. The causal significance of an association is a matter of judgment which goes beyond any statement of statistical probability. To judge or evaluate the causal significance of the association between the attribute or agent and the disease, or effect upon health, a number of criteria must be utilized, no one of which is an all-sufficient basis for judgment. These criteria include: the consistency of the association; the strength of the association; the specificity of the association; the temporal relationship of the association; [and] the coherence of the association (2, pg. 20).

Stallones' draft to the Advisory Committee ranked the criteria differently and had this caveat:

> The application of these rules may be modified by knowledge of the precision with which the variables may be measured and by whether a factor is considered to be necessary or contributory and sufficient or insufficient. All of these considerations must be involved in the judgment of the importance of smoking as a causative factor . . . (19, pg. 4-5).

Stallones' daughter, epidemiologist Lorann Stallones, told us in a 2008 reminiscence of her father: "It is my understanding that Mickey LeMaistre [former President of the University of Texas and member of the Advisory Committee] has a napkin that Dad wrote these ['rules'] down on while they were working on the Surgeon General's Report."

A more colorful, possibly fanciful, depiction of Stallones' off-the-cuff summation of the guidelines was quoted, without attribution, in Richard Kluger's book "Ashes to Ashes," in a scene from the June 1963 meeting of the Committee in Saratoga Springs: "Stallones said, 'This is what I think we've been talking about;' taking an empty pack of Luckies from his pocket and tearing it apart, scrawled on the inside surface of the wrapper: 'The consistency of the statistical association, the strength of the association, specificity of the association, and the coherence of the association'" (20, pg. 252).

Neither records of that meeting, archival drafts of the Report, nor references cited in the Report make attribution to origins of these characteristics from outside the Committee. When we asked Stallones in the late 1960s where he got the causal criteria, his main contribution to the Report, he responded: "They just came into my head. I probably read them somewhere."

In fact, the guidelines were not universally applied within the Report. John Hickam, in a letter to Len Schuman, scribe for the Report, requested they be highlighted elsewhere and not be included in his section on CVD and respiratory diseases (19). In the end, the guidelines were presented early, in

Chapter 3, and elaborated in the Report's main Chapter 9, on smoking and cancer. In contrast to their absent consideration for smoking and CVD, each cancer type was run systematically through the gamut for its match to "Stallones' criteria" (2).

Today, the Report is widely credited in the U.S. with more than its due for the formulation of the guidelines, but also, properly, for their first application to a major health issue. Through them, the Report was enabled to credibly recognize the main cause of lung cancer, launching an international public health effort against tobacco.

Sir Austin etched the tablet

In his President's Address to the Section on Occupational Medicine of the Royal Society of Medicine in 1965, Austin Bradford Hill, head of epidemiology at the London School of Hygiene and Tropical Medicine and professor emeritus of statistics at the University of London, asked: "What aspects of [an] association should we especially consider before deciding that the most likely interpretation of it is causation?" (3, pg. 295)

Hill had long pondered his findings of an association of cigarette smoking and lung cancer from the British Doctors Study. In research and teaching he had displayed a clear priority of ordered thinking about causality. For example, Hill had vigorously engaged, on the American platform of his 1953 Cutter Lecture, the issue of causal inference from observational versus experimental evidence. There he was reacting in defense of observations to the preceding lecture by Oxford nutritionist Hugh Sinclair, who maintained bombastically that: "the use of the experimental method has brilliant discoveries to its credit, whereas the method of observation has achieved little" (12, pg. 995).

Pioneer, as was Hill, in the design of case-control and cohort studies (e.g. the British Doctor's Study), as well as early designer and advocate of randomized clinical trials (RCT), he pointed out in this early lecture how so few lifestyle effects on health are amenable to experiment. Accordingly, he proposed ways to sharpen data collection and reduce bias, so that observations might "be made in such a way as to fulfill, as far as possible, experimental requirements" (12, pg. 995).

It seems strange, therefore, that in his 1965 address he failed even to allude to the elegant layout of the guidelines found in the 1964 U.S. Report, or to the considerable American thinking found in the documented exchange of views that led up to that report.

In any case, because of the clarity and color of his language, we quote "Hill's Guidelines" verbatim (with italics and comments added):

"First upon my list I would put the strength of the association. (3, pg. 295)."

Hill illustrates strength of association by his own finding of 10- to 20-fold greater rates of lung cancer in heavy smokers of cigarettes.

"Next on my list of features to be specially considered I would place the consistency of the observed association" (3, pg. 296).

To illustrate consistency, Hill mentions--in one of only two citations of the U.S. Surgeon General's Report of the year before--its findings of the cigarette smoking-lung cancer association in 29 retrospective and 7 prospective inquiries.

"Specificity of the association [is] the third characteristic which invariably we must consider" (3, pg. 297).

On this point of contention, Hill considered that if specificity exists, we may be able to draw conclusions without hesitation; [but] if it is not apparent, we are not thereby necessarily left sitting irresolutely on the fence" (3, pg. 297).

"My fourth characteristic is the temporal relationship of the association—which is the cart and which is the horse?" [And this is] "particularly relevant with diseases of slow development" (3, pg. 297).

"Fifthly, if the association is one which can reveal a biological gradient, or dose-response curve, then we should look most carefully for such evidence" (3, pg. 298).

"It will be helpful if the causation we suspect is biologically plausible." {He qualifies:] "What is biologically plausible depends upon the biological knowledge of the day" (3, pg. 298).

"Coherence: . . the cause-and-effect interpretation of our data should not seriously conflict with the generally known facts of the natural history and biology of the disease" (3, pg. 298).

Here, Hill cites a whole paragraph from the U.S. Advisory Committee report about the coherence of the temporal rise in cancer and in cigarette smoking, indicating that he was intimately acquainted with the earlier report by the time he gave his President's Address.

"Experiment: Occasionally it is possible to appeal to experimental or semi-experimental evidence." [If people stop smoking cigarettes] "is the frequency of the associated events affected? Here the strongest support for the causation hypothesis may be revealed" (3, pg. 298).

"Analogy: With the effects of thalidomide and rubella before us we would surely be ready to accept slighter but similar evidence with another drug or another viral disease in pregnancy" (3, pg. 299).

Hill concluded that, "No formal tests of significance can answer those questions" (3, pg. 299) and closed his discourse in the tradition of John Snow, with a powerful message about taking public action based on congruent, if forever-incomplete, evidence for cause:

> All scientific work is incomplete--whether it be observational or experimental. All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a

> freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time (3, pg. 300).

Hill's Predecessors

Absent the accrual of characteristics seen in the American dialogue about the guidelines, or direct evidence of Hill's exposure to those developments, Morabia attributed Hill's synthesis to his exposure to the 18th century thinking on causality of Scots philosopher, David Hume. He puts Hume, Hammond, Yerushalmy, Palmer, Lilienfeld, Sartwell, and the Surgeon General's Advisory Committee, all, in a chain of predecessors (if unacknowledged) to Hill (21, pg. 369).

According to Labarthe and Stallones, Hill's intent was pragmatic, to provide humble guidelines to individual judgment rather than fixed rules and commandments from on high (22). Berlivet added the opinion that Hill's expanding the guidelines from 5 in the U.S. Report to 9 in his, was intended to reach: "a stricter regulation of causal inference in epidemiology… to avoid the traps of any epistemological debate on causality. . . Bradford Hill always pledged against a strict interpretation of his own formalization" (23, pg. 58).

On the point of Hill's failure to credit either the documented thinking of American experts prior to the 1964 U.S. Report, or the Report itself, we have the following statement found in his personal memoir written to Brian Furner, librarian of the London School of Hygiene and Tropical Medicine in 1988:

> In 1964 Richard Shilling persuaded me to become the first President of the new section of Occupational Medicine of the Royal Society of Medicine. My presidential address was "The Environment and Disease – Association or Causation." There is here a question of priority. I had been thinking on these lines for a long time and at about that time someone wrote a similar article in the U.S. Surgeon General's Report on smoking and lung cancer. Had I seen it--I think I probably had--but I elaborated on it and presented it in a talk that I gave to the 100th meeting of the Medical Section of the Royal Statistical Society (held in the [London] School). This was not published so I took it up again, further elaborated it and used it as my Presidential Address. And later I transferred it to my short textbook as a final chapter. The question of priority (it does not matter a damn to me) turns on whether I had seen the Surgeon General's report or not. I would guess I had, made a mental note of it and developed it. This is what I wrote to some U.S. prof. who wrote to ask me about it (1985/86) [Italics added] (24).

In epidemiology today, these thoughtful "joint" (U.S.-UK) guidelines to causal inference from statistical associations-are enshrined. Though rejected as subjective or not logical by a few (6), debated and supplemented by leading theorists (4,5, 25), they are, nevertheless, widely observed in common usage (25,26).

Thus, our anecdote about a fire lit in the mind of gruff skeptic and statistician, Jacob Yerushalmy, by the naïve bluster of physiologist Ancel Keys, may provide a little color to the otherwise staid origins of

these classic criteria. The subsequent exchanges, moreover, assembled in series probably for the first time here, show clearly the guidelines evolving among American scholars between 1957 and 1963.

Among the historical contingent that has thought the Surgeon General's Report the true Act of Creation, and accepted the Advisory Committee's self-anointed title, views of the Report may now change. Another contingent, which has considered "Hill's Criteria" delivered truth, the Nine Commandments, may now ponder what preoccupation led Bradford Hill not to acknowledge the scholarship of others, at home or over the seas. We find that original sources—the words of Stallones and Schuman in the U.S., and Hill's memoir in a London archive—reveal no evidence to dispel the mystery of misplaced priority and provenance. We suspect, however, neither pride nor intrigue. *(Henry Blackburn and Darwin Labarthe)* (27)

References

1. Koch, R. About the Present State of the Bacteriological Diagnosis of Cholera. Journal for Hygiene and Infectious Diseases. 1893;(14): 319-333 (in German).
2. Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service. US Department of Health, Education, and Welfare, Public 3Health Service Publication No. 1103. 1964.
3. Bradford-Hill, A. President's Address. The Environment and Disease: Association or Causation? Proceedings of the Royal Society of Medicine. 1965;58(5): 295-300.
4. Evans, A.E. Causation and Disease: The Henle-Koch Postulates Revisited. Yale Journal of Biology and Medicine 1976;49:175-95.
5. Susser, M. Causal Thinking in the Health Sciences: Concepts and Strategies in Epidemiology 1973. London, Oxford University Press.
6. Burch, P.R.J. The Surgeon General's "Epidemiologic Criteria for Causality." A Critique. Journal of Chronic Diseases. 1983. 36:821-836.
7. Lilienfeld, A.M. The Surgeon General's "Epidemiologic Criteria for Causality." A Critique of Burch's Critique. Journal of Chronic Diseases. 1983;36: 837-845.
8. Keys, A. Atherosclerosis: a problem in newer public health. Journal of the Mount Sinai Hospital. 1953;20 (2):118-39.
9. Yerushalmy, J. and Hilleboe, H.E. Fat in the diet and mortality from heart disease. New York State Journal of Medicine. 1957;57(14): 2343-2354.
10. Keys, A, ed. Seven Countries: A Multivariate Analysis of Death and Coronary Heart Disease. Cambridge: Harvard University Press;1980.
11. Yerushalmy, J. and Palmer, C.E. On the methodology of investigations of etiologic factors in chronic diseases. Journal of Chronic Diseases. 1959;10(1): 27-40.
12. Bradford-Hill, A. Observations and experiment. New England Journal of Medicine.1953; 248(24): 995-1001.
13. Lilienfeld, A.M. Emotional and other selected characteristics of cigarette smokers and nonsmokers as related to epidemiological studies of lung cancer and other diseases. Journal of the National Cancer Institute. 1959;22(2):259-282.

14. Lilienfeld, A.M. On the methodology of investigations of etiologic factors in chronic diseases. Some comments. Journal of Chronic Diseases. 1959;10(1):41-46.
15. Lilienfeld, A.M. The case against the cigarette. The Nation. March 31.1962.
16. Lilienfeld, D. Abe & Yak, the interactions of Abraham Lilienfeld and Jacob Yerushalmy in the development of modern epidemiology, 1945-1973. Epidemiology 2007; 18:507-514.
17. Sartwell, P.E. On the methodology of investigations of etiologic factors in chronic diseases. Further comments. Journal of Chronic Diseases. 1960;11: 61-63.
18. Hammond, E.C. Cause and effect. In, Wynder, E.L. (ed.) The Biologic Effects of Tobacco. 1955. Boston. Little, Brown and Company.
19. Stallones, R.A. 1963. The association between tobacco smoking and coronary heart disease. Draft of June 28 to the Surgeon General's Advisory Committee on Smoking and Health. J.Hickam's September 16, 1963 letter of response to L. Schuman. Leonard M. Schuman papers. Box 52. "Cardiovascular." University of Minnesota Archives. <http://special.lib.umn.edu/uarch/>
20. Kluger, R. Ashes to Ashes. America's Hundred-Year Cigarette War, the Public Health, and the Unabashed Triumph of Philip Morris. New York, N.Y. Vintage Books; 1997.
21. Morabia, A. On the Origin of Hill's Causal Criteria. Epidemiology. 1991; 2: 367-369.
22. Labarthe, D. and Stallones, R.A. Epidemiologic inference. In, Rothman, K.J. (ed.) Causal Inference. 1988. Epidemiology Resources, Inc.
23. Berlivet, L. 'Association or Causation?' The Debate on the Scientific Status of Risk Factor Epidemiology, 1947–c.1965. Clinical Medicine 2005;75:39-74.
24. Hill, A.B. 1988. Memoirs of Bradford Hill's time at the School. GB 0809. Archives of the London School of Hygiene and Tropical Medicine.
25. Rothman, K.J. Causes. American Journal of Epidemiology 1976; 104:587-592.
26. Parascandola,M., Weed, D.L., Dasgupta, A. Two Surgeon General's reports on smoking and cancer: a historical investigation of the practice of causal inference. Emerging Themes in Epidemiology. 2006; 3:1.
27. Blackburn, H. and D. Labarthe. Stories From the Evolution of Guidelines for Causal Inference in Epidemiologic Associations: 1953–1965. Am. J. Epidemiol. 2012; 176 (12): 1071-1077.

