RE: “INVITED COMMENTARY: RESPONSE TO SCIENCE ARTICLE, ‘EPIDEMIOLOGY FACES ITS LIMITS’”

In Ernst L. Wynder’s “Invited Commentary: Response to Science Article, ‘Epidemiology Faces Its Limits’” (1), two related points must be clarified regarding references and attribution of concepts. The first is a mistake in formal referencing. Sir Austin Bradford Hill’s frequently cited reference (2) is used instead of the legitimate but more cumbersome reference by the US Public Health Service (3). The Advisory Committee’s causal criteria clearly antedate Hill’s in date of publication and in conception. In June 1963, the criteria were first scribbled on the inside of a Lucky Strike cigarette package by Dr. Reuel Stallones as an attempt to summarize a 3-day epidemiologic brainstorming retreat (4). They were aired in subsequent deliberations of the Committee and were modified, fleshed out, and formally published in January 1964 (1). Hill’s criteria paper (2) was published in May 1965, more than a year later than ours.

As formal guidelines, we wanted something like Koch’s postulates, but the Henle-Koch postulates pertained to the causation of infectious disease. They were not applicable to the broader question of inference-making regarding the most probable causative factors in multifactorial chronic diseases. The Committee’s charge was to thoroughly evaluate all evidence for all the health effects of tobacco, namely, bronchitis-emphysema, cardiovascular disease, oropharyngeal problems, etc., as well as lung cancer and “if possible, to reach some definitive conclusions on the relationship between smoking and health in general” (3).

Wynder’s statement that “in 1957, Bradford Hill reported criteria on causation that became the framework of the first Surgeon General’s Report on Smoking and Health in 1964” (1, p. 748) is mistaken. The date, 1957, is not verifiable—not in the literature, nor by personal knowledge and recollections of those most intimate to the study—and we had frequent contact with Sir Richard Doll about their extensive data. In Hill’s paper, published in May 1965, there is no evidence, either internal or by formal referencing, that he had propounded the “Bradford Hill Criteria” in 1957 or on any date earlier than 1965 (2). In fact, Hill’s 1965 paper contained only one reference regarding causal criteria, namely the report of the advisory committee (3).

The most important point is that we propounded and articulated these criteria de novo during the progress of our deliberations. At the time, we did not consider them hewn in stone or intended for all time and all occasions, but as a postulates, but the Henle-Koch postulates pertained to the causation of infectious disease. They were not applicable to the broader question of inference-making regarding the most probable causative factors in multifactorial chronic diseases. The Committee’s charge was to thoroughly evaluate all evidence for all the health effects of tobacco, namely, bronchitis-emphysema, cardiovascular disease, oropharyngeal problems, etc., as well as lung cancer and “if possible, to reach some definitive conclusions on the relationship between smoking and health in general” (3).

Wynder’s statement that “in 1957, Bradford Hill reported criteria on causation that became the framework of the first Surgeon General’s Report on Smoking and Health in 1964” (1, p. 748) is mistaken. The date, 1957, is not verifiable—not in the literature, nor by personal knowledge and recollections of those most intimate to the study—and we had frequent contact with Sir Richard Doll about their extensive data. In Hill’s paper, published in May 1965, there is no evidence, either internal or by formal referencing, that he had propounded the “Bradford Hill Criteria” in 1957 or on any date earlier than 1965 (2). In fact, Hill’s 1965 paper contained only one reference regarding causal criteria, namely the report of the advisory committee (3).

The most important point is that we propounded and articulated these criteria de novo during the progress of our deliberations. At the time, we did not consider them hewn in stone or intended for all time and all occasions, but as a formal description of how we drew our most important epidemiologic conclusions from the totality of tobacco-related materials extant. This common misapprehension and misreferencing for the past 30 years has many possible factors:

1. A. B. Hill’s reference (2) is short and sweet, while the title of the Surgeon’s General’s “Smoking Report” (3) is very cumbersome and not easy to find or to get exactly right.
2. Hill’s presentation is more gracefully stated and is easier reading. Abe Lilienfeld preferred it as “logically more elegant” (5).
3. Some authors have not read either of the original sources and use someone else’s referencing for convenience.
4. The reference to Hill has become much more popular in the literature, thus perpetuating its more frequent use. I had thought Abe Lilienfeld’s broadside in 1983 against Burch (5) would have corrected the situation, but apparently it has not. It would seem that the clearest and most accurate referencing would be something like: “… first propounded by the Surgeon General’s Advisory Committee on Smoking and Health (3) and later expanded and refined by A. B. Hill (2)”.

REFERENCES


Peter V. V. Hamill
1001 Whitehall Cove
Annapolis, MD 21401

THE AUTHOR REPLIES

Peter Hamill’s letter (1) describes the history of establishing the criteria of judgment as they, in fact, took place. He is correct that the five criteria were first published in the Surgeon General’s Report on Smoking and Health in 1964 (2) and that the attribution of the criteria has often been given to Bradford Hill, who restated them as nine points given to Bradford Hill, who restated them as nine points

1. The widespread increase in lung cancer rates in many countries paralleled the increase in tobacco consumption in those countries.
2. Lung cancer rates and tobacco consumption were higher in men than in women, and lung cancer rates and tobacco consumption were higher in urban than in rural populations.
3. A long latency exists between the onset of smoking and the peak age groups for lung cancer (late 50s).
4. Epidermoid bronchogenic cancer is rarely found in nonsmokers.
5. Clinicians had long used smoking history as a differential diagnosis for lung cancer.
6. Tobacco is a proven animal carcinogen.
7. No other factor can explain the observed statistical association.
8. Each one of these points taken alone may not suffice as proving that smoking causes lung cancer, but taken together, they appear to establish proof that tobacco is a carcinogen to the human bronchial epithelium.

The five criteria for causality in smoking and lung cancer established in the Surgeon General's Report were similar: 1) the consistency of the association in retrospective and prospective studies; 2) the strength of the association, as measured by relative and absolute risk; 3) the specificity of the association in that lung cancer is rare in nonsmokers; 4) the temporal relation of the association, in that a long induction period is noted; and 5) the coherence of the association, in that the rise in lung cancer mortality rates parallels the increase in cigarette consumption, that a sex differential exists, that an urban-rural differential exists, and that a dose-response relation exists. Listed separately, but important in supporting causality, was biologic evidence in the form of animal experimentation and bronchial histopathologic damage in human smokers.

As we look at the history of this issue, it is clear that the criteria of causation that were prepared from the expert panel of the Surgeon General's Report would get more attention than that of a resident in medicine expressing the same view in a lesser publication. Regardless of the historic perspective, it is important that we recognize the value of the criteria of judgment. One can only wish that these criteria would receive more attention by current investigators, particularly those involved in the study of weak associations.

REFERENCES

Ernst L. Wynder
American Health Foundation
New York, NY 10017

RE: “TIME-TO-EVENT ANALYSIS OF LONGITUDINAL FOLLOW-UP OF A SURVEY: CHOICE OF THE TIME-SCALE”

Korn et al. (1) recommend using age rather than length of follow-up (time-on-study) as the time scale in Cox proportional hazards regression models when analyzing data from longitudinal studies of healthy populations. We concur with their recommendation and consider their paper to be a significant contribution. However, we feel that it is important to respond to their statement that models using length of follow-up as the time-scale are "incorrect." We also discuss some other issues.

If time-on-study modeling is incorrect, do the many published analyses that have used that type of model need to be reanalyzed? Korn et al. (1, p. 74) assert that analyses performed using time-on-study models "may not be seriously in error" if 1) the underlying hazard is exponential or 2) the covariates and age are statistically independent. To get a sense of how often these conditions hold and how similar the results from the two models are when they hold, we reanalyzed four studies with different outcome variables involving NHEFS I Epidemiologic Followup Study (NHEFS) data that originally used time-on-study models (2). We chose a variety of outcomes in an effort to find a nonexponential hazard: all-cause mortality, hip fracture incidence, coronary heart disease incidence, and arthritis incidence. In all four studies, we found that the hazard was approximately exponential and that the coefficients and their standard errors from the time-on-study models were similar to those from the age models (stratified on 5-year birth cohort and not stratified on birth cohort). These findings support the conclusion that for outcomes typically examined in epidemiologic studies, the underlying hazard is approximately exponential and the results from time-on-study and age models are similar. Therefore, while modeling with age as the time-scale may be theoretically preferable, it generally makes little or no substantive difference.

A practical issue pertaining to the use of age as the time-scale is the paucity of software that is currently available. The only software of which we are aware that incorporates age as the time-scale and can take into account a complex survey design is the program of Korn et al. (1).

Another issue to consider when using age as the time-scale is what the appropriate model is when the effect of a risk factor changes with age. This situation occurs frequently in epidemiologic studies of chronic disease as the effects of some risk factors weaken with age. In a time-on-study model, age-dependent risk factors have been analyzed by carrying out separate analyses for various age-at-baseline or age-at-risk groups or by including age interaction terms in the model. Korn et al. did not discuss age-dependent risk factors in the context of the age model. Some adjustments to the model would need to be made in this situation for substantive reasons as well as for statistical reasons, since the hazards would be nonproportional.

Korn et al. used data from the NHEFS to illustrate their points. The first National Health and Nutrition Examination Survey (NHEFS I), which serves as the baseline for NHEFS, is a complex survey. We would like to stress that age must be modeled with particular care in analyses of the NHEFS because certain age groups were oversampled at baseline and because the sample weights are highly variable (3). Thus, for this survey, modeling of age-dependent risk factors must also take into account the age-related effects of the oversampling.