

The major grounds for criticism of retrospective
and prospective studies.

Such studies are subject to biases from various sources. We only discuss here those biases which may affect an accurate assessment of association, leaving until later biases which may affect association in comparison with causation. The main sources of bias are these:

- (1) dependence on memory for information about past or current smoking habits,
- (2) effect of the patient's smoking habits on the reported cause of death,
- (3) effects of selection of respondents
- (4) effects of knowledge of diagnosis at time of interview.

These deserve brief separate treatment.

Almost all retrospective and prospective studies have gathered information on past smoking habits. (Doll and Hill's prospective study of doctors is the outstanding exception.) Experience with family expenditure surveys shows how uncertain such information about consumer purchases can be. Clearly presence of "smoker's cough", of respiratory disease, or of diagnosed lung cancer, could affect memory for amount smoked and thus bias the comparison among non-smokers and various grades of smokers. Exactly the same situation arises to

a lesser degree, when present smoking habits are inquired for. The events to be reported are more recent, but memory is still involved. (cp. example of dish purchases within the last week).

The possible effect of knowledge of the patient's smoking habits on the reported cause of death has been considered in the literature. Evidence as to such a source of bias must inevitably be indirect. Most direct are such items as Wynder's statement at the New York City ASA meeting (Xmas 1955?) that he considered smoking habits in diagnosing lung cancer, and the 5% of the doctors involved in Doll and Hill's special inquiry (1956, page 1076) who admitted that it had or might well have affected their diagnosis. (The % admitting may confidently be expected to be an underestimate.) Next less direct comes the evidence from the comparison of smoking behavior for non-microscopically confirmed diagnoses of lung cancer with that for microscopically confirmed cases. This evidence is quite hard to assess fairly, since quite different results are reached depending on whether the bias is regarded as due to an excess of false lung cancer diagnoses for smokers to a deficit of true lung cancer diagnoses among non-smokers. If the former were the case, then positive bias would be reflected by a lowered apparent

dependence on smoking for confirmed cases. Doll and Hill (1956 page 1076) rely on the absence of such an effect as a proof of the absence of positive bias. If bias, however, arises from the existence of a number of non-smokers dying from lung cancer whose deaths are reported under other causes of death, then it is not unreasonable that the rate of necropsy and operation is larger for a smoker with a fixed intensity of disease than for a nonsmoker. In such a situation a trend from microscopically diagnosed cases to otherwise diagnosed cases may be expected to continue toward undiagnosed cases. In both Doll and Hill's study and Hammond and Horn's studies microscopically diagnosed cases show a greater apparent association with smoking than do otherwise diagnosed cases. This clearly suggests that undiagnosed cases would show an even lower gradient, which would mean that the present overall gradient is falsely high - - that the association of lung cancer with smoking is exaggerated by such an effect.

The possible effects of self-selection of respondents in consequence of varying smoking habits, varying states of health, etc. have been clearly discussed by Berkson, who points out that if non-respondants are allowed to exclude themselves from the sample at their own will, the results will be exposed to a wide variety

of bias. One of these, which is discussed in detail by Berkson, leads to a fallacious impression of positive association between smoking and lung cancer, and to a reduced overall death rate. Such low overall rates are observed in Hammond and Horn's study and in the early years of Doll and Hill's study. The existence of such lowered rates is suggestive of a bias of the type proposed by Berkson, but is not proof thereof. The absence or smallness of such a deficit in death rates is evidence against the specific sort of bias discussed by Berkson, though not against other possible biases of the same general type. Thus the gradual decrease in the death rate deficit of the Doll and Hill is evidence against the importance of the specific bias mechanism selected by Berkson, but does not in any way rule out other biases due to differential non-response or other causes of selection of respondents.

The retrospective studies, and, as well, the earliest years of each prospective study, are endangered by bias due to the knowledge, by those concerned in the interview or questionnaire-filling out situation - - the patient, the interviewer, the relatives of the deceased - - of the then current diagnosis of illness or reported cause of death. Such knowledge can influence, subconsciously, such matters as, for example, extent of

interviewer pressure for recollection and interpretation by interviewer or questionnaire-completer of unspecific statements and knowledge. The active discussion of this source of bias in the lung cancer-smoking literature is an indication of its possible seriousness.

The distinction between association and causation.

The problem of the rotation and distinction between proof of association and proof of causation is an old one. Much has been written about it. The actions of those leading the progress of physical, biological and medical science have been consistent with a single point of view toward this problem, a point of view taken implicitly by today's leaders in putting statistics to good use and a point of view consistent with Anglo-American legal principles.

Explicitly summarized, the view is the following:

(1) empirical evidence may show association, and may show its degree of closeness - - enough evidence will make this association, and its closeness, as certain as may be desired,

(2) empirical evidence alone cannot show causation,

(3) causation is shown by empirically showing association and theoretically showing that the association could not reasonably have arisen in any other way than by causation in a particular direction.

The analogy with the legal doctrine of "proof beyond a reasonable doubt" is clear. So long as there remains a reasonable mechanism to explain the observed association other than the proposed causation, there remains a reasonable doubt, and causation has not been established.

Experimentation, in which the experimenter compares the results of different actions applied to humans, animals or objects according to his choice (today more and more often guided by randomization procedures to avoid even subconscious bias) has always been regarded as the best sort of evidence relating to causation. Why is this so? It is logically possible that the results to come - - which mouse is to die of disease, which plot to grow the most corn - - might in some unknown way influence the experimenter in choice of actions - - he might always give the new, possibly worthless remedy to the mouse who will resist the disease anyway, he might always apply the new, possibly worthless fertilizer to the plots which will anyway yield the most corn. All this is logically possible, but not reasonable (especially where randomization is carefully practiced). It is this lack of reasonableness, in a situation where putative cause both precedes putative effect and is controlled by a human or randomized choice, which gives to experiment its high evidential validity concerning causation.

In areas where experiment has not been feasible, as in much of astronomy, and in assessing of the possible effects of smoking on lung cancer, the same standard of "proof beyond a reasonable doubt", here expressible as

"association without a reasonable alternative explanation" must be applied whenever causation is to be established.

The standards of scientific evidence are closely similar to those of legal evidence - - proof is required beyond a reasonable doubt. Many applications of physical, biological and medical science - - in some fields most applications - - are based on something less than proof. Just as good driving of an automobile may call for slowing down when it is not certain that an accident is about to happen ahead, so good medical practice may call for treatment of, or advice to, patients based on the doctor's best judgment, and not on what has been proved. The same is true in many fields of engineering. It will always be important to distinguish between good scientific practice and what is scientifically proved.

Reasonable explanations for association
without causation between smoking and
lung cancer.

When, and especially if, it becomes established that there is association between smoking and lung cancer in the general U.S. male population, it will be necessary to consider reasonable explanations of association not involving causation, and only after all such alternative explanations have been disposed of can causation be regarded as having been proved.

The most important class of reasonable alternatives is made up of alternatives which operate because smoking is a voluntary choice - - voluntary, but influenced by external, genetic and personality factors which may be associated with the actual cause of the lung cancer. There are a number of such alternatives, and no single one need be responsible for the association between smoking and lung cancer which we presume, for the sake of discussion, to have eventually been established. Each of a number might contribute its small quota to the overall association. Thus evidence to prove such alternatives unreasonable must do more than merely show each one incapable of providing the whole of any observed association, it must instead show each one's contribution to be so small that the total could not account for the observed association, supposed established.

It is known that smoking habits differ in amount between rural and urban residents, and between occupations. It is reasonable to assume that other environmental factors, including some so far not recognized, also influence smoking habits. Studies purporting to establish the relation between smoking and lung cancer have given clear indication (Hammond and Horn, 1955) that lung cancer rates differ between rural and urban residents in the same smoking class (i.e. smoking the same amount of tobacco). This is prima facie evidence that rural-urban differences exert an effect on lung cancer other than through smoking. It is reasonable to suppose that other environmental factors, some which may well be associated with differences in smoking habit, also effect the incidence of cancer of the lung in ways not involving smoking. It is, at the present, very hard to set an upper limit to the degree of association between smoking and lung cancer which might arise from such environmental influences.

It has been suggested, and no contrary evidence has been adduced, that genetic, constitutional and psychological differences between persons might influence both the incidence and the progression of cancer of the lung. Such effects could well contribute to the observed association.