

## ON THE METHODOLOGY OF INVESTIGATIONS OF ETIOLOGIC FACTORS IN CHRONIC DISEASES

J. YERUSHALMY, PH.D., AND  
CARROLL E. PALMER, M.D., PH.D.,  
BERKELEY, CALIF., AND  
WASHINGTON, D. C.

*From the Division of Biostatistics, School of Public Health, University of California at Berkeley, and the Tuberculosis Program, Division of Special Health Services, U. S. Public Health Service, Washington, D. C.*

(Received for publication Jan. 12, 1959)

MUCH of our knowledge concerning etiologic factors in chronic diseases has been derived from uncontrolled, or poorly controlled, observations. Experimentation on human beings is difficult if not impossible in investigations of causes of diseases and conditions of long duration, especially those requiring strict randomization. Occasionally a set of circumstances occurs which enables the alert investigator to utilize observations in such a manner as to approximate and simulate planned experimentation. The classic and oft-quoted observations on cholera by Snow represent an example of an occurrence of this type. Such an advantageous combination of circumstances, however, is indeed rare.

The search for meaningful associations, in the main, will have to be guided by observations of phenomena as they occur, with little, if any, interference on the part of the observer. Progress in the identification of causative factors in the important chronic diseases will depend to a great extent on our ability to utilize such observations and on the development of more rigorous methods and rules of analysis which will increase the likelihood of correct interpretation and will minimize the effect of extraneous factors inherent in uncontrolled observations.

The multiple approach, i.e., the investigation of the effect of a specific factor on a disease through a number of different independent types of observations, if each independently indicates the same relationship, reduces the chance that results are due to extraneous factors and increases the significance of the findings. However, the observations must be indeed different in type and not merely different manifestations of the same set of phenomena all subject to the same limitations. Moreover, each observation must be accurate and reliable, and the relationship implied in each of the different types of observation must be valid in itself.

---

Presented in part at the Fourth International Biometric Conference, Ottawa, Canada, Sept. 2, 1958.

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 associations as are observed may, to a greater or lesser degree, be due to factors other than those under study. Thus, whether the investigation is based on a single type or a multiplicity of types of observation, it is necessary to evaluate each for its soundness and validity.

Unfortunately, the methodology and criteria for evaluation have not yet been adequately developed. It is, therefore, perhaps not surprising that some data derived from uncontrolled observations are accepted as conclusive by some investigators and rejected by other equally competent observers.

In this connection it may be instructive to compare the current approach in investigations of etiologic factors in many chronic diseases with the more rigorous methods long in use by bacteriologists in implicating a living organism as the causal agent for a specific disease. Almost from the very beginning, when bacteria were first found to cause disease, bacteriologists felt the need for a set of rules to act as guideposts in investigations of bacteria as possible causal agents in disease. The formulation by Koch of these postulates and their utilization in the field of bacteriology has contributed greatly to the orderly and systematic identification of causative organisms in many diseases.

It is the purpose of this paper to develop an elementary parallelism between investigations of etiologic factors in certain chronic diseases and those of bacterial diseases. This is done with a full realization of the inherent complexities involved. We do not pretend to provide a formulation which embraces all the relevant considerations. We are making an attempt to provide a basis for discussion with the hope that it will be useful in the eventual development of guideposts for the more systematic investigation of causative factors in chronic disease.\*

#### MULTIPLE CAUSATION

In many of the important chronic diseases we are not yet at the stage of attempting to identify a definite, final, and single entity as a causal agent. Rather we are concerned with the investigation of conditions, often environmental, which may be involved in the causation of a given disease. These conditions, however, may, at best, be looked upon only as vectors or vehicles which may contain the specific causative agent. It is perhaps this distinction between the testing of a single definite organism as a possible causal agent and the investigation of characteristics or conditions as cause-carrying vectors which is at the root of the complexities with which we are concerned.

---

\*From the comments of several very competent investigators who read this paper in manuscript, it appears that the exploratory nature of the discussion may be easily overlooked. We wish, therefore, to emphasize that we do not claim to provide a definitive formulation of any of the points presented. Rather, we wish to raise certain issues which appear to us to be important in contemporary epidemiology, with the hope that they may stimulate discussion which may lead to the clarification of those issues. We also wish to make it clear that the examples used in this paper, in particular that of the relation between smoking and lung cancer, are only to illustrate methodology and not to express a judgment on the relative value of such other evidence as may bear on the problem.

As an example, in the study of the relationship of smoking to cancer of the lung, the immediate purpose is to determine whether smoking is implicated as a possible vector. If this were shown, further intensive analysis would be in order to determine how smoking causes the disease or what specific causative agent it contains. If such a specific agent were found, the demonstration that it causes lung cancer would not differ essentially in methodology from the demonstration that a certain microorganism causes a specific bacterial infection.

It is likely that for many of the chronic conditions there may be more than one agent causing the disease, each of which may be found in more than one vector. Thus even if smoking is found to be a cause of lung cancer, it is obvious that the smoking per se is not the specific cause, but that tobacco smoke may contain the proximate causal agent. The same agent might eventually be found in other environmental phenomena, such as automobile exhausts or fumes given off by various industrial processes. It is possible, also, that there are several different agents each of which may cause lung cancer and each of which may be found in one or more vectors.

Whether there are different specific causes for a given chronic disease or a single specific cause involved in different vectors, the practical effects on the conduct of the investigation are the same, namely, that in many of the chronic diseases the investigator is confronted with the problem of "multiple causation." With present knowledge, or ignorance, of causal factors in chronic disease, the search often starts with the broad characteristic—the vector—and continues through many steps to the identification of the specific cause. The problem of multiple causation is more prominent during the early stages—the investigation of vectors—than at the last step of testing the specific causal agent.

The present status of the search for etiologic factors, with respect to many of the important chronic diseases, is not unlike that which existed with the bacterial diseases prior to the discovery of microorganisms. At that time, the problem of multiple causation must have been more disturbing than in the later stages.

Thus, for example, once the typhoid bacillus has been identified the demonstration that it is the cause of typhoid fever is accomplished by a much more direct type of investigation than is the type of study of polluted water or polluted milk, both of which may be involved as vectors for the same bacillus.

In either case, whether the investigation relates to a vector or to a definite causal agent, there is the further complication that, even if either is causally related to the disease, the presence of the vector, or of the causal agent, does not necessarily imply the presence of the disease. In the case of a vector or a characteristic that is being investigated, it is obvious that even if it is causally related to a given disease, it is neither necessary nor sufficient for the disease. It is not necessary because the characteristic is usually not the sole cause and therefore the disease may exist in its absence. It is not sufficient because the characteristic may exist without the presence of the disease. The lack of sufficiency is often encountered also in investigations of single specific causative agents, such as microorganisms as causes of certain diseases. In most of the

investigations of etiologic factors in chronic disease, however, many of the suspected characteristics are *prima facie* neither necessary nor sufficient.

To use again the example of smoking, even if it were shown to cause lung cancer, the large number of smokers who do not have cancer of the lung testify to the insufficiency of smoking as a cause of the disease, and the existence of lung cancer patients who have never smoked clearly indicates that smoking is not a necessary cause. This situation is in contrast to that of diseases where a single specific agent is under suspicion, for in these cases the single agent, although not sufficient for the disease, is nevertheless necessary. Consequently, an essential element in the demonstration of the agent as a cause is that when adequately investigated it can be shown to be present in every case of the disease. Because the characteristics currently under study are neither necessary nor sufficient, investigations of etiologic factors in many chronic diseases lack the advantage of a one-to-one correspondence between cause and effect, in either direction.

The importance of any characteristic as a possible cause or cause-carrying vector for a given disease can then be revealed not by its presence in every case, but only by an increase in the relative frequency of its occurrence among persons with the disease. Consequently, the statistical method must play an important role in investigations of causative factors in chronic diseases. On the other hand, because of the nature of uncontrolled observations, conventional statistical techniques cannot be utilized without modification, because the fundamental requirement of group comparability, ordinarily achieved through randomization, is not satisfied. Perhaps it would be useful to employ a special word, such as "epidemic," to emphasize the fact that in these investigations special techniques must be utilized, techniques which are basically quantitative (metric) but a kind specially designed for investigating the etiology of human disease.

#### KOCH'S POSTULATES

While there is no single formulation of Koch's postulates uniformly quoted in all texts on bacteriology, they can be stated as consisting essentially of the following:

- I. The organism must be found in all cases of the disease in question.
- II. It must be isolated from patients and grown in pure culture.
- III. When the pure culture is inoculated into susceptible animals or man, it must reproduce the disease.

It is of immediate importance to note that the bacteriologist is not satisfied with the demonstration of the concurrent presence of the bacteria and the disease as proof that the organism causes the disease. He does not stop with the first postulate. This must indicate that he is aware that it is possible for an organism to be present in all cases of a certain disease\* and yet not cause it. The presence

---

\*This, of course, implies that it is not present in all "noncases."

of the organism in such cases must be looked upon either as the effect, rather than the cause, of the disease,\* or as a mere accompaniment to another organism which is the true cause, a "satellite," as it were.

The second and third postulates were thus formulated to differentiate between organisms which are specifically responsible for a disease and those which, although present, do not themselves cause it. Postulates II and III must therefore be considered as tests for specificity of effect, i.e., the organism is the "essential" cause of the disease. The requirement that the organism be isolated in pure form and reproduce the disease is the equivalent of establishing that the suspected organism has the capability of causing the disease in question.

The third postulate serves also to satisfy a sequential requirement. In order to implicate an organism in a cause and effect relationship, it is necessary to show not only simultaneous occurrence of organism and disease, but appearance in the correct sequence. The organism must be introduced into the body first and the disease developed subsequently. Postulate III accomplishes this in addition to satisfying the requirement for specificity of effect.

Not in every case is it possible to satisfy all of Koch's postulates. Bacteriologists, however, are aware of the uncertainties and the opportunities for error when the evidence is not complete as may be judged from the following quotation:

. . . it must be recognized that any omission in the complete chain of evidence involves a risk of error; and much confusion has been caused by uncritical attempts to support the claims advanced on behalf of numberless bacteria isolated from different parts of the body in various diseases.<sup>2</sup>

From the point of view of their possible application to the current status of investigations of etiologic factors in chronic diseases, the Koch postulates may, in the light of the above, be thought to imply two essential types of evidence necessary for the implication of an organism as a cause of disease. These may be stated to relate to:

- A. The simultaneous presence of organism and disease and their appearance in the correct sequence, and
- B. The specificity of effect of the organism on the development of the disease.

In current investigations of etiologic factors in chronic diseases, attention has been focused primarily on the equivalent of A. Little if any concern has been expressed with respect to the second basic type of evidence, namely, the specificity of effect.† Even with respect to the demonstration of the simultaneous occurrence of characteristic and disease, no set procedures have been established. It is, therefore, desirable to discuss the implications of these two types of evidence to the methods of searching for etiologic factors in chronic disease.

---

\*That is, the organism has a predisposition for diseased tissue.

†The term "specificity of effect" is not entirely satisfactory for the chronic disease situation. It is used here only in lieu of a more suitable word to express the type of evidence which increases the conviction that the suspected factor is indeed directly related to the disease. It is appreciated that the term's connotation, in this instance, is not the same as when applied to the bacterial disease situation.

*A. The Simultaneous Presence of Characteristic and Disease and Their Appearance in the Correct Sequence.*—Koch's postulates are directly applicable only when a specific single organism is under investigation, so that the presence of the organism may be demanded in every case of the disease. With diseases in which broad characteristics are under study, simultaneous occurrence can be required only in terms of an increase in the relative frequency of occurrence. This major difference, as was pointed out previously, stems from the fact that broad characteristics, such as those currently studied in connection with chronic diseases, appear to be neither necessary nor sufficient for the development of the disease.

As a consequence, a number of complications often arise: first, difficulty of measurement; second, selection of controls against whom an increase in frequency may be gauged; and third, the question of the correct sequence of events. The equivalent, for chronic diseases currently under investigation, of Koch's first postulate must take one or the other of the following forms, which are, of course, not independent:

1. The suspected characteristic must be found more frequently in persons with the disease in question than in persons without the disease, or
2. Persons possessing the characteristic must develop the disease more frequently than do persons not possessing the characteristic.

In general, studies of the retrospective type are utilized to establish statement 1, and those of the prospective type are needed to establish statement 2. It is not necessary in all instances to satisfy both statements, but it should be pointed out that statement 1 does not involve testing for the correct sequence. In many studies the question of sequence does not arise because it is often quite apparent that the disease cannot cause the characteristic.

Occasionally, however, the problem of sequence does arise. Thus, if the blood cholesterol level is found to be higher in persons who have experienced a coronary attack than in appropriate controls, it is not certain whether the high cholesterol level is a cause or an effect of the coronary attack. Similarly, when, as is often the case, knowledge of the presence of the disease may influence the discovery by the investigator or the recalling and reporting by the patient of the presence of the characteristic, the correct sequence is of great importance.\* Statement 2 is then the more meaningful, and the prospective type of study is indicated.

The most important consideration in either the prospective or retrospective study relates to the selection of controls. What is meant by "persons without the disease" in statement 1 and "persons not possessing the characteristic" in statement 2?

It is obvious that in each case these statements must refer to persons who are in all other respects similar to those having the disease or the characteristic.

---

\*As an example, knowledge of the presence of a congenital malformation in a newborn infant may lead to more thorough questioning of the mother and to her recalling certain infections during pregnancy which might well have gone unnoticed had the birth been normal.

Consequently, the method of selection of the study and control groups is of paramount importance. In prospective studies, the problem is related to the question of specificity of effect and will be discussed later.

In retrospective studies the usual method consists of the selection of a group of persons possessing a certain disease and the selection of an appropriate control group. The two groups are then compared to determine differences between them in the frequencies of one or more characteristics. The validity of the comparison rests in large measure on the composition of the control group in relation to that of the study group. A probability sample of the general population appears on the surface to be the most appropriate group for use in such comparisons. In many types of investigation, however, a probability sample may not be the most appropriate.

Such a control is no doubt desirable when little is known of the possible influence which the common demographic characteristics have on the occurrence of the disease and investigation is performed for the purpose of initial identification of the effects of such unsuspected characteristics. But when these are known and one or more other specific characteristics are under suspicion, the probability sample may not be the most economical or efficient mechanism for study.

The purpose in the latter case is obviously to focus observations on the factors under suspicion, and therefore it is necessary to make the two groups as nearly comparable as possible in all respects, except that one group has the disease and the other does not. A control group consisting of a probability sample of the general population may not be appropriate for this purpose, even when stratified for a number of characteristics.

Since many diseases are known to concentrate their attacks on certain age, sex, race, socioeconomic, and other groups, it becomes necessary to stratify a probability sample for these factors if it is to serve as an adequate control. The result is, first, that exceedingly large groups must be selected under these circumstances, even if the control group is to be stratified to the study group in only a few factors. More important, perhaps, is the fact that once it is so stratified and matched to the study group the major advantages of having been derived from a probability sample may be lost, for the resulting control group derived after such stratification and matching may no longer be a representative sample of any identifiable population.

Because of these difficulties it is perhaps simpler to begin with a matched sample in which for each individual in the study group a paired control is selected who is as near as possible to the study individual in as many characteristics as possible. Unfortunately, we have not yet learned the best methods of selecting matched controls. Matching is often performed in a haphazard fashion and limited to the selection of persons who are also patients in the study group institution, but in a different ward and suffering from a different condition.

Serious investigations of the development of optimal methods for selecting controls are needed. It is indeed very likely that in most investigations several groups of controls selected in different ways should be utilized. More important is the need for a different approach in selections of controls. Instead of using

individual single characteristics as a basis for selection, it is, perhaps, possible to devise mechanisms for selection which encompass groups of characteristics even if these are sometimes ill defined. The reason for this is that in many of the studies in which an etiologic factor is sought, we are concerned with environmental factors which, when taken together, form the vague concept of "mode of life."

For example, if one were to examine the effect of diet on a specific chronic disease, one would be impressed with the fact that diet is much influenced not only by socioeconomic factors, but also by many other mode-of-life conditions. In order to pinpoint the effect of diet, it is necessary to compare two groups who are nearly alike in socioeconomic and mode-of-life factors to determine whether differences in dietary intake are meaningful. For if precautions are not taken to equalize the two groups for these characteristics, it is not possible to determine whether any differences in diet between the study and control groups are attributable to the disease or to differences in the mode of life which may characterize the two groups.

Consequently, we must explore different means of selecting controls. For example, it is possible that a person's residential area might be a cross-sectional mechanism for identifying people in terms of the mode-of-life characteristic. In other words, the best control for a person with coronary heart disease, for the purpose of studying etiologic factors, might be a person of nearly the same age who lives near him—his neighbor. Other possibilities might be a person's friends—his bridge partner, the one who drinks an occasional beer with him at the local tavern, or the one who works next to him at a desk or assembly line.

It is very likely that in many studies different types of control groups must be selected, depending upon the nature of the characteristics under suspicion. For example, if interest is focused on specific occupations as suspected etiologic factors, it would not do to select as a control a person in the same office or factory. In this case a neighbor might be a better control. On the other hand, if one is interested in the effect of housing as a suspected etiologic factor, the neighbor would obviously be the worst person to select as a control.

*B. Specificity of Effect of the Characteristic on the Development of the Disease.*—The demonstration of a greater frequency of the characteristic in the study group than in an adequately selected control group satisfies no more than the equivalent of the first of Koch's postulates, for this indicates only that the characteristic and the disease tend to occur together more often than expected.

The requirements of Koch's second and third postulates, that the organism be isolated in pure form and shown to reproduce the disease, serve the purpose of ruling out cases in which a suspected organism is not the real cause, but merely accompanies it.

The similar requirement of establishing that a characteristic has a specific effect on the development of a chronic disease involves extremely complex considerations. The more serious of these stem from the problem of self-selection. The groups under observation and comparison are not selected by the investigator in a random procedure, but the individuals in the two groups have made

for themselves the critical decision of acquiring the characteristic. As a result, great doubt often exists as to the comparability of the two groups. In many cases a question arises as to whether the characteristic under study may not in fact be a mere axis of classification—an index which differentiates the two groups in terms of many important factors, and that these rather than the characteristic under investigation are causally related to the disease.

For example, when a sample of the population (even if randomly selected) is divided into smokers and nonsmokers and the two groups are observed for the occurrence of cancer of the lung, the habit of smoking may serve only as a differentiating index for the two groups. Persons who become smokers may differ in many respects from those who elect to be nonsmokers. It is perhaps these unknown factors which may be related to the occurrence of cancer of the lung, much as in the situation where an organism may accompany the real causative agent but be harmless in itself (like *sarcinae* in gastric cancer).

A recent study in Buffalo, for example, indicates that smokers differ significantly from nonsmokers in a number of characteristics: they are more neurotic, are more frequently hospitalized, and they change jobs and spouses more often than nonsmokers.<sup>3</sup> Had the two groups been compared for other characteristics, it is quite possible that they would have been found to differ in these also. It is conceivable that some of these other characteristics, rather than smoking, are responsible in whole or in part for the increase in lung cancer among smokers.

To use a more familiar illustration, only a decade or so ago many evaluations of immunizing or therapeutic agents utilized volunteers as the study group, while persons who refused to volunteer for treatment were used as controls. Few today would be willing to accept the latter as controls for the volunteers.

Bradford Hill expresses the general lack of confidence in this method of selecting controls in commenting on an investigation of the effects of vaccination against influenza which used volunteers and in which the difference in influenza rates between the study and control groups was reported to be statistically significant with a  $P$  of 0.00657:

. . . and yet does this ritual and do all these decimal places mean anything at all? Admittedly the technical test says that the two groups had experiences that differed by more than one would expect to occur by chance; equally it tells nothing else. As it stands, I do not myself believe that it gives any support whatever for the author's conclusion that 'here is evidence strongly in favor of immunization of large groups in industry.'<sup>4</sup>

On the other hand, many investigators, including Professor Hill, feel that recent prospective studies lend strong support to the thesis that smoking is not only associated with, but is a cause of cancer of the lung. Yet it is difficult to perceive a basic difference in methodology in the two situations. Both suffer from the same fundamental weakness—the factor of self-selection.

The reason that nonvolunteers are not acceptable as controls for volunteers must be the belief that volunteering per se represents an index—an axis of classification—which differentiates the two groups according to many factors which

may have a bearing on the occurrence of the disease. The volunteer may be of the type of person who takes better care of his health and who may be a more careful person in general. As a consequence, he might be likely to experience less of the disease in question than the nonvolunteer, even if he were not immunized. Differences in morbidity from the disease between volunteers and nonvolunteers thus may not be due to the effectiveness of the vaccine being tested, but to differences in the personalities and modes of life of the two groups.

It follows by similar reasoning that the experiences in lung cancer mortality of smokers and of nonsmokers may not logically be compared to assess smoking as the cause of the disease. Smoking, like volunteering, may represent an index which differentiates the two groups in many aspects of mode of life and perhaps also on constitutional grounds. Consequently, smokers may experience risks of death from cancer of the lung different from those of nonsmokers, even if they never smoked at all. That is, people who eventually become smokers may possess certain characteristics which make them more vulnerable to certain diseases than persons who are in the category of "nonsmokers." Smoking by itself, as an etiologic factor in lung cancer, may be the analogue of the "satellite," the accompanying organism of other disease situations.

This suggestion is no more unreasonable than is the one which implies that the act of volunteering renders a person incomparable with a nonvolunteering control. In fact, one may be in a much stronger position in the case of volunteers, for the volunteers and nonvolunteers are often recruited from the same group, such as nurses in a hospital. The two groups are thus equalized with respect to many characteristics, including age, sex, and the general environment of the hospital.

The demonstration of high relative frequencies in the study group is thus only a first step in the process of searching for etiologic factors. The investigation must proceed to the second and more crucial consideration (which, for want of a better term, is denoted here as that of specificity of effect), i.e., to the demonstration that the difference in relative frequencies reflects a specific and meaningful relationship between the characteristic under suspicion and the disease under consideration.

Problems of specificity arise not only in connection with the characteristic but also with the disease entity. Difficulties in diagnosis and in grouping related disease entities may often blur and complicate the understanding of observed associations.

More serious are tendencies to use an easily obtainable operation or measurement as a substitute for the disease itself. Thus in studies of the relationship of environmental factors to coronary heart disease, the blood cholesterol level is substituted for the disease entity. An investigation may show a relationship between a characteristic such as dietary fat and the blood cholesterol level, but the results are often presented with the implication that a relationship between fat intake and coronary heart disease has been demonstrated.

Similarly, in investigations of the value of an immunizing agent, reaction to an antigen is often substituted for immunity. The value, for example, of

B.C.G. as a protection against tuberculosis has been "demonstrated" in a number of investigations by the postvaccination changes in tuberculin reactions.

The use of substitutes of these kinds is at times desirable and often unavoidable. They have a definite role to play as an interim step in many an investigation. However, they cannot be used as proof of a relationship unless the specificity of the substitute has been demonstrated to be of high order and approaches a one-to-one relationship with that of the disease.

#### TESTING FOR SPECIFICITY OF EFFECT

Although relationships demonstrated by means of observations on humans can never be accepted with the confidence that a well-planned experiment can supply, much can be done to sharpen the tool of epidemetric investigation. At this stage of our knowledge and experience with epidemetric investigations, it is not possible to formulate rules which can be used as definite guides to separate causal from spurious relationships. As a first step, however, it may be stated on an intuitive basis that when a given characteristic is found to be associated with one, or at most a few, diseases, then the evidence for a causal relationship is more convincing, taken by itself, than when the characteristic is found to be associated with a large number of diseases.

Thus, while it is not possible to satisfy the equivalent of Koch's second and third postulates, it may be possible in many investigations to attempt a first approach to the attainment of the major objective underlying these postulates. This objective is taken to be that of establishing that the organism is specifically related to the disease and not in an incidental fashion as a reflection or accompaniment of other organisms.

It is in this sense that the term "testing for specificity of effect" is utilized for epidemetric investigations. The basic assumption of such a test is that if the characteristic is not related to the disease in a causal way, then the relationship should not be restricted to the disease under study but should also be present with other disease entities. If the characteristic can be shown to be related only or mostly to the disease under study and not to many other disease entities, then our confidence that it is a cause-carrying vector for that disease is greatly increased.

If, on the other hand, it is found that the characteristic is also related to numerous other diseases, including those without obvious physiologic or pathologic connection with the characteristic in question, the relationship must be assumed—until further proof—to be nonspecific. Any meaningful association of the characteristic with the disease under study must depend upon other evidence. The statistical relationship need not be entirely overlooked, but it must be considered as secondary evidence only.

The testing of the relationship of the characteristic to other disease entities, therefore, must constitute an essential part of the investigation. The relevance of the association with the original disease will depend to a great extent on the results of such investigation. The fewer the additional diseases which are associated with the characteristic, the greater the weight which may be assigned to the relationship of the characteristic to the specific disease under study.

Occasionally it is possible to investigate the specificity of a relationship with respect to the characteristic as well as the disease, by substituting an associated characteristic for the one under study and investigating the relationship of the substituted characteristic to the disease.

As an illustration, Keys<sup>5</sup> used an association between the amount of fat in the diet and mortality from heart disease in several countries as support for his theory that fat is a causal factor in atherosclerosis. However, Yerushalmy and Hilleboe showed that a similar and, indeed, stronger relationship can be demonstrated between protein and heart disease and that a negative association exists between these components of the diet and mortality from diseases other than those of the heart. They concluded that the original association must be considered nonspecific and cannot be used in even partial support of a supposed causal relationship.<sup>6</sup>

To return to the smoking—lung cancer illustration, if smoking were shown to be related to lung cancer only or restricted to lung cancer and several related and physiologically explainable diseases, the association would have specificity and significance in suggesting a causal relationship. If, on the other hand, similar relationships can be shown with a variety of diseases, some of which cannot reasonably be thought to be influenced by smoking per se, then the association with cancer of the lung lacks specificity. The support which the statistical association provides for a causal relationship between smoking and lung cancer is correspondingly reduced.

It follows, therefore, that investigations based primarily on observations on humans, in addition to demonstrating an increase in relative frequency as in statements 1 and 2 mentioned previously, must consider and study relationships of the suspected characteristic with other diseases, and, if possible, associations of related characteristics with the disease in question. In other words, for an association to be valid in the sense that it may have causal implications, it must satisfy the requirement for specificity. The formulation of such a requirement is difficult in view of the many complexities encountered in chronic disease investigations.

Thus, it is not possible to indicate in precise terms what is meant by "relationships which may be expected on physiologic and pathologic grounds." Nevertheless, because it seems important that these issues be discussed, the following statement is presented for purposes of such discussion:

3. 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. The suspected characteristic can be said to be specifically related to the disease in question when the results of such investigation indicate that similar relationships do not exist with a variety of characteristics and with many disease entities when such relationships are not predictable on physiologic, pathologic, experimental, or epidemiologic grounds. 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.

#### SUMMARY AND CONCLUSIONS

Investigations of etiologic factors in chronic diseases are often based on uncontrolled observations which are subject to considerable limitations. In bacterial diseases, the utilization of Koch's postulates as criteria has contributed to the orderly and systematic identification of causative organisms in many diseases.

In this paper, an elementary parallelism is developed between current investigations of etiologic factors in a chronic disease and investigations of bacterial disease outlined by Koch. It is suggested that two essential types of evidence are involved in Koch's postulates:

- A. The simultaneous presence of organism and disease, and their appearance in the correct sequence, and
- B. The specificity of effect of the organism on the development of the disease.

Because in chronic diseases multiple causation is often encountered, and especially because of the disturbing factor of self-selection, problems arise in measurement, in selection of controls, and in testing for specificity of effect.

For purposes of discussion the following statements are suggested as a first approach toward the development of acceptable guideposts for the implication of a characteristic as an etiologic factor in a chronic disease:

1. The suspected characteristic must be found more frequently in persons with the disease in question than in persons without the disease, or
2. Persons possessing the characteristic must develop the disease more frequently than do persons not possessing the characteristic.
3. 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. The suspected characteristic can be said to be specifically related to the disease in question when the results of such investigation indicate that similar relationships do not exist with a variety of characteristics and with many disease entities when such relationships are not predictable on physiologic, pathologic, experimental, or epidemiologic grounds. 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.

These statements are not intended to provide a formulation which embraces all the relevant considerations involved. They are presented with a full realization of the inherent complexities, especially the arbitrary nature of some parts

of statement 3 above; for example, when are "relationships predictable on physiologic grounds"? It is hoped, nevertheless, that they might be useful as a basis for discussion and in the eventual development of a more systematic methodology for the investigation of causative factors in chronic diseases.

## REFERENCES

1. Snow, John: *On the Mode of Communication of Cholera*, London, 1855, John Churchill.
2. Wilson, G. S., and Miles, A. A.: *Topley and Wilson's Principles of Bacteriology and Immunity*, Baltimore, 1955, The Williams & Wilkins Company, Vol. II, p. 1141.
3. Lilienfeld, A. M.: Emotional and Other Selected Characteristics of Cigarette Smokers and Nonsmokers as Related to Epidemiological Studies of Lung Cancer and Other Diseases, *J. Nat. Cancer Inst.* **22**:259, 1959.
4. Hill, A. Bradford: Observations and Experiment, *New England J. Med.* **248**:995, 1953.
5. Keys, A.: Atherosclerosis: A Problem in Newer Public Health, *J. Mt. Sinai Hosp.* **20**:134, 118, 1953.
6. Yerushalmy, J., and Hilleboe, Herman E.: Fat in the Diet and Mortality From Heart Disease, *New York J. Med.* **57**:2343, 1957.