relationship in real life is greater than is revealed by the data. In the above example the misclassifications of patients regarding their blood pressure or stroke status reduced an actual twelvefold increase of stroke in hypertensives to an observed twofold increase.

Nevertheless, the study of large groups allows one to detect important relationships, using poor data that are intolerable in conscientious patient care. This, then, is the explanation to the clinician of the seeming tolerance of epidemiology for inadequate data.

REFERENCES

Chapman, J. M., V. A. Clark, and A. H. Coulson. 1966. Problems of measurement in blood pressure surveys: Inter-observer differences in blood pressure determinations. *Am. J. Epidemiology*, **84**:483–494.

Cluff, L. E., G. F. Thornton, and L. G. Seidl. 1964. Studies on the epidemiology of adverse drug reactions: I. Methods of surveillance. *J. Am. Med. Assoc.*, **188**:976–983.

Collen, M. F., J. L. Cutler, A. B. Siegelaub, R. L. Cella. 1969. Reliability of a self-administered questionnaire. *Arch. Intern. Med.*, **123**:664–681.

Feinstein, A. R. 1968. Clinical epidemiology. II. The identification rates of disease. *Ann. Intern. Med.*, **69**:1037–1061.

Friedman, G. D., C. C. Seltzer, A. B. Siegelaub, R. Feldman, and M. F. Collen. 1972. Smoking among white, black and yellow men and women: Kaiser-Permanente Multiphasic Health examination data, 1964–1968. *Am. J. Epidemiology*, **96**:23–35.

Madow, W. G., Interview data on chronic conditions compared with information derived from medical records. National Center for Health Statistics Report, Ser. 2, No. 23, U.S. Department of Health, Education, and Welfare, 1967.

Reid, D. D., and J. G. Evans. 1970. New drugs and changing mortality from non-infectious disease in England and Wales. *Brit. Med. Bull.*, **26**:191–196.

Wilcox, J. 1961. Observer factors in the measurement of blood pressure. *Nursing Research*, **10**:4–20.

Yerushalmy, J. 1969. The statistical assessment of the variability in observer perception and description of roentgenographic pulmonary shadows. *Radiol. Clin. N. Amer.*, **7**:381–392.

Chapter 4

Basic Methods of Study

In the two preceding chapters the reader has been introduced to the data employed in epidemiology and the basic measurements that are used to describe groups of persons. It is now appropriate to consider the major types of epidemiological investigation. Each type of study uses these tools in a particular way and has a unique logical framework. In addition, each type of study is especially appropriate for the unique circumstances surrounding any particular investigation—the aims of the investigation, the populations available for study, and the human and financial resources that can be brought to bear on the problem.

Relationships

Much of the effort of medical scientists in understanding the etiology of disease and developing appropriate therapies involves a study of

the relationship of one type of event or characteristic or "variable" to another. Consider the following questions as examples:

Does exposure to cold wet weather predispose to the common cold?
What is the influence of the serum potassium concentration on the contractility of the heart?
Is obesity related to the occurrence of gallstones?
What is the effect of vitamin C deprivation on wound healing?
Which part of the hemoglobin molecule carries the oxygen?
Does BCG vaccination provide protection against pulmonary tuberculosis?

In Table 4-1 these questions are listed together with the relationship that should be studied to help answer each. In a two-variable relationship one is usually considered the *independent* variable, which affects the other, or *dependent*, variable.

The relationships that are studied need not be only between one variable and a second. Often, the investigator must be concerned with the interrelationship of three or more variables. For example, in order to better understand the relationship of potassium to the force of cardiac contraction, calcium concentration must also be taken into account. Whether or not obesity is related to gallstone occurrence may depend on racial characteristics and the type of diet eaten, both of which must be considered and assessed as additional independent variables.

Observational versus Experimental Studies

There are two basic approaches to investigating the relationship between variables. In *observational* studies, nature is allowed to take its course and changes or differences in one characteristic are related to changes or differences in the other, if any. In *experimental* studies, the investigator actually intervenes and makes one variable change and then sees what happens to the other. In doing so he tries, as much as possible, not to allow other important variables to affect the outcome. By controlling the experimental situation, he may conclude that the intervention or manipulation of the independent variable actually affected, or caused the change in, the dependent variable.

Table 4-1 Examples of Relationships Studied in Order to Answer Certain Questions

Question	Suggests study of the relationship between variables	
	Independent variables	Dependent variables
Does exposure to cold wet weather predispose to the common cold?	Daily weather conditions	Incidence of common cold
What is the influence of the serum potassium concentration on the contractility of the heart?	Serum potassium concentration	Stroke output of the heart
Is obesity related to the occurrence of gallstones?	Skinfold thickness	Prevalence of gallstones
What is the effect of vitamin C deprivation on wound healing?	Vitamin C content of the diet	Tensile strength of healing wounds
Which part of the hemoglobin molecule carries the oxygen?	Portion of hemoglobin molecule	Affinity for oxygen
Does BCG vaccination provide protection against pulmonary tuberculosis?	Presence or absence of vaccination	Incidence of tuberculosis

Epidemiology includes both observational and experimental studies. An example of an epidemiologic experiment was the large-scale field trial of poliomyelitis vaccine in which two large groups of children, comparable in all important respects (e.g., age, health, socioeconomic status, and likelihood of exposure to poliomyelitis

virus) received vaccine and placebo, respectively, with follow-up to measure the subsequent incidence of poliomyelitis (described in Chap. 9).

Because of the difficulties of performing well-controlled experiments on human populations and the availability of an abundance of observational data, epidemiologists have tended to concentrate on observational studies. In doing so they have tried to "control" the important extraneous variables by their data-analysis methods. Also, they are always on the lookout for "natural experiments"—spontaneous occurrences which approximate experiments by virtue of a change in only one independent variable that is apparently unaccompanied by changes in other important variables. An example might be the sudden graded exposure to ionizing radiation received in 1945 by the population of Hiroshima, which has permitted the study of the relationship of different doses of radiation exposure to the subsequent development of a variety of diseases.

Such natural experiments are rare (thank goodness) and the observational epidemiologist has to rely on other techniques and criteria for determining the possible effects of additional variables.

Observational studies fall into two main categories, descriptive and analytic. These studies, in turn, may be subdivided into cross-sectional or prevalence studies, case-control studies, and incidence or cohort studies, depending on the groups of persons investigated and the time relationships involved. (Case-control studies are probably best included only in the analytic category.) These will be described subsequently. Attention will also be paid to defining and clarifying the relationship between prospective and retrospective studies due to the confusion that revolves around this distinction.

Descriptive versus Analytic Studies

There are two fundamental objectives of observational epidemiologic studies. One is to *describe* the occurrence of disease or disease-related phenomena in populations. The other is to *explain* the observed pattern of occurrence of disease. Seeking the latter objective involves the identification of causal or etiological factors.

Descriptive studies usually involve the determination of the incidence, prevalence, and mortality rates for diseases in large population groups, according to basic group characteristics such as

age, sex, race, and geographic area. In this way, the general distribution of disease in the population is described.

Studies attempting to explain disease are often referred to as *analytic* studies. The starting point for an analytic study is often a descriptive finding that raises certain questions or suggests certain hypotheses that require further investigation. With analytic studies the investigator has a specific question or group of questions in mind that he sets about to answer.

The distinction between descriptive and analytic studies is not clear-cut. A large-scale descriptive study may (perhaps unexpectedly) provide abundant and impressive data that give a clear answer to a specific question. In an analytic study, designed to answer specific questions, data collected incidentally may be of great descriptive interest and raise further questions for investigation.

Despite this fuzziness, it is often useful to categorize epidemiologic studies in this manner. Descriptive studies usually involve a more diffuse, superficial, or general view of a disease problem. Analytic studies narrow down on a specific question and may require a more rigorous study design and data analysis.

Prevalence or Cross-Sectional Studies

Prevalence, or *cross-sectional*, studies examine the relationships between diseases and other characteristics or variables of interest as they exist in a defined population at one particular time. The presence or absence of disease and the presence or absence of the other variables (or, if they are quantitative, their level) are determined in each member of the study population or in a representative sample at one particular time. The relationship between a variable and the disease can be examined in two ways, either (1) in terms of the prevalence of disease in different population subgroups defined according to the presence or absence (or level) of the variables or, conversely, (2) in terms of the presence or absence (or level) of the variables in the diseased versus the nondiseased.

Case-Control Studies

Case-control studies are similar to prevalence studies in that they assess the relationship of *existing* disease to other variables or

attributes. After the initial identification of cases, that is, location of persons with the disease of interest, a suitable control group or comparison group of persons without the disease is identified. The relationship of an attribute to the disease is examined by comparing the diseased and nondiseased with regard to how frequently the attribute is present or, if quantitative, what the levels of the attribute are in the two groups.

Incidence or Cohort Studies

Instead of measuring the relationship of attributes to existing disease, as do prevalence and case-control studies, *incidence*, or *cohort*, studies look more directly at attributes or factors related to the *development* of disease. A study population free of the disease under investigation is identified at a particular time. The attributes of interest are measured initially in this group of persons, known as a *cohort*. Then, these persons are followed up over a period of time for the development of the disease being studied. The relationship of an attribute to the disease is examined by dividing the population into subgroups according to the presence or absence (or level) of the attribute initially and comparing the subsequent incidence of disease in each of the subgroups.

An Illustrative Example

Prevalence, case-control, and incidence studies are discussed in detail in Chapters 6, 7, and 8, respectively. At this point an example may help to clarify the distinction among these study plans. Suppose we wish to learn whether obesity predisposes to degenerative arthritis of the knees. In a prevalence study we would x-ray the knees of a defined population, perhaps all the adults in a community, and determine degree of obesity by measuring height and weight or skinfold thickness. We would then compare the prevalence of osteoarthritis in population subgroups showing various degrees of obesity. Or, we may wish to contrast the mean skinfold thickness or other obesity measure in those with osteoarthritis and those without.

In a case-control study of this question, we might collect a group of persons with osteoarthritis of the knees hospitalized at a local hospital during the past year. For a control group, we might

select for each osteoarthritis case, a person of the same sex and similar age, admitted to the same hospital during the same week for minor elective surgery such as herniorrhaphy or hemorrhoidectomy. We would then compare the recorded heights and weights of the case group with those of the control group to see if, indeed, the osteoarthritis cases were more obese.

To approach this problem by an incidence study, we would go back to a defined adult population and x-ray their knees to exclude persons with existing osteoarthritis. We would then measure skinfold thickness or height and weight in order to divide the population without osteoarthritis into the obese and nonobese or, preferably, some finer gradations of fatness. We would call them back 10 years later for repeat knee x-rays, which would demonstrate new cases of osteoarthritis. Then we would compare the incidence of osteoarthritis in the various fatness groups.

Remembering our original question, "Does obesity *predispose* to osteoarthritis?" the incidence study approach seems to provide the most direct answer, since we looked for obesity *before* the osteoarthritis developed. The prevalence and case-control studies provided only indirect evidence, since they looked at obesity at the same time as disease. However, the time sequence can often be taken into account in the prevalence and case-control studies. In addition to measuring current weight in persons with and without osteoarthritis we could also have inquired about their weight 10 years ago, or at age 25, or before their knees started to hurt, thus investigating their weight prior to the development of osteoarthritis. The information obtained may not be as accurate as that derived from weighing the subjects initially in an incidence study, but time sequence can be considered in prevalence or case-control studies.

Prospective and Retrospective Studies

The question of time sequence leads naturally into a consideration of the much-discussed *prospective* and *retrospective* studies. It is almost a matter of faith that investigations are unsatisfactory if they are retrospective. One often hears such comments as, "Of course, this study was retrospective, so we can not be confident of the findings."

Before discussing the merits of prospective versus retrospec-

tive studies, it is important to clarify their meaning. Actually, much confusion has resulted because the terms are used in two different ways leading to such semantic horrors as "retrospective-prospective" studies.

One of the meanings of prospective versus retrospective has to do with the time period over which the data were recorded in relation to the time the decision was made to do the study. In this sense, retrospective studies involve a decision to carry out an investigation with observations that have been recorded in the past. In contrast, prospective studies involve the collection of observations after the decision is made to carry out the investigation.

The other meaning of prospective versus retrospective studies is related not to the time sequence of the observations and the decision to do the study but, rather, to the time sequence of observations of study variables and the occurrence of disease. In this sense, prospective studies are analogous to incidence studies, and retrospective studies are analogous to prevalence, or case-control, studies. Prospective or incidence studies measure characteristics and wait for disease to develop, while retrospective or prevalence studies measure the characteristics in persons already diseased.

It is strongly suggested that this second set of definitions be discarded, since better terms are available, as noted. The advantages and disadvantages of prevalence, incidence, and case-control studies will be discussed in Chaps. 6, 7, and 8. The following discussion of prospective versus retrospective studies will consider only the first pair of definitions, relating to when the data were collected.

In prospective studies the investigator can plan and control the methods for making and recording observations, keeping in mind their purpose. In retrospective studies the already-recorded data may have been collected for an entirely unrelated purpose. Therefore these data may well be incomplete and recorded in a manner not appropriate for the present study.

Consequently, there often are severe problems involved in retrospective studies. Consider a study of changes in the outcome of treatment of congestive heart failure in a particular hospital over a period of several years. In carrying out a retrospective study, the

investigator would be plagued by the fact that the criteria for the diagnosis of congestive heart failure vary over the years and vary from doctor to doctor. The recent advent of central venous pressure measurements may have improved the ability to diagnose the condition. Cases diagnosed many years ago may differ in character and severity from those diagnosed last year. Therefore observed changes in outcome may be related more to differences in initial severity than to the effects of treatment. If one of the criteria for improvement were weight loss, the investigator would find, to his frustration, that admission and discharge weights were not recorded for many patients over the years, ruling them out of this aspect of the study.

If this study were carried out prospectively, the investigator could initially establish criteria for the diagnosis of congestive heart failure and set up objective measures of severity and improvement. In addition, he could establish procedures to ensure that all the needed measurements were made uniformly on all patients. Thus, the superiority of a prospective study of this question is obvious.

Not all retrospective data need be of poor quality. If we again consider retrospective studies using hospital charts, a variety of data come to mind that would probably have been recorded accurately and consistently. Examples are time of admission, number of days spent in the hospital, sex of the patient, whether the patient died, and whether he received any blood transfusions.

REFERENCE

MacMahon, B., and T. F. Pugh, *Epidemiology: Principles and Methods.* (Boston: Little, Brown, 1970), Chap. 3.