3

Types of Epidemiologic Research

CHAPTER OUTLINE

3.0 PREVIEW

In this chapter we will employ two design criteria to create a basic typology of epidemiologic research. The criteria are (1) whether the study factor is artificially *manipulated* by the investigators or others—i.e., whether dif-

Table 3.1 Types and Objectives of Epidemiologic Research

Type	*Subtype*	*Objectives*
Experimental (artificial manipulation of study factor with randomization)	Laboratory	Test etiologic hypotheses and estimate acute behavioral and biological effects
		Suggest efficacy of intervention to modify risk factors in a population
	Clinical trial	Test etiologic hypotheses and estimate long-term health effects
		Test efficacy of interventions to modify health status
		Suggest feasibility of population intervention
	Community intervention	Identify persons at "high risk"
		Test efficacy and effectiveness of clinical/societal interventions to modify health status within a particular population
		Suggest public health policies and programs
Quasi-experimental (artificial manipulation of the study factor without randomization)	Clinical/ laboratory	Same as clinical trial or laboratory experiment
	Program/policy	Evaluate extent to which public health goals are achieved
		Determine unanticipated problems or consequences of implementation and reasons for success or failure of the intervention
		Compare costs and benefits of intervention
		Suggest changes in current health policies or programs
Observational (no artificial manipulation of the study factor)	Descriptive	Estimate disease frequency and time trends and identify diseased individuals
		Generate etiologic hypotheses and suggest rationale for new studies
	Analytic (etiologic)	Test specific etiologic hypotheses and estimate chronic health effects
		Generate new etiologic hypotheses and suggest mechanisms of causation
		Generate preventive hypotheses and suggest potential for disease prevention

ferent categories of the study factor are allocated to subjects; and (2) if manipulated, whether categories of the study factor (i.e., treatments) are randomly allocated—i.e., *randomized*—to all subjects. Together, these two criteria delineate three mutually exclusive types of studies, as outlined in the following sections and summarized in Table 3.1.

3.1 EXPERIMENTS

Studies in which randomization is used to allocate subjects are called *experiments*. While many professionals in the health field believe that epidemiologists do not conduct "true" experiments, a well-defined tradition of experimental research has grown in epidemiology, largely in response to the need for testing the efficacy of new chemotherapies and vaccines (Harris and Fitzgerald, 1970; Comstock, 1978). More recently, population researchers have increased their use of experimental designs to examine the impact of other medical practices (Saracci, 1979; Wennberg et al., 1980) and behavioral programs intended to reduce morbidity and mortality in the community (Syme, 1978; Farquhar, 1978).

In the simplest type of experiment, one group of subjects is given an experimental *treatment* and another group (often called the *controls*) is given either no treatment, a sham treatment (e.g., a placebo drug), or, less preferably, the customary treatment under nonexperimental conditions (e.g., an available drug). After a period of follow-up, the change in the response variable (e.g., disease status) is measured for every subject, and the two groups are compared with respect to estimated summary values of the response. A difference in response between groups suggests an effect—beneficial or adverse—of the new treatment.

Experimental designs are specifically adapted to each study situation by modifying and/or restricting the randomization process. Most important among such features is the procedure of *double blinding*, in which neither the investigator nor the subjects know the treatments to which they have been allocated. In a clinical drug trial, double blinding is usually accomplished by having a third party assign groups and not informing either the examiners or the subjects about their treatments until after the trial. Double blinding is very important to experimental design because it helps to ensure that neither the investigator nor the subject influences the response in ways that promote bias. For more thorough discussions of experimental designs, see Harris and Fitzgerald (1970), Peto et al. (1976, 1977), and Weddell (1979).

3.1.1 Subtypes and Objectives

We can define three subtypes of epidemiologic experiments; the classification depends on the approximate duration of the study and the selection of subjects.

The *laboratory experiment* has the shortest duration—usually, a matter of hours or days—and, therefore, is used to estimate acute biological or behavioral responses that are believed to be risk factors for the disease.

Typically, the study population is very highly restricted or self-selected so that it is seldom representative of the target population. For example, we might test whether a particular type of acute stressor (e.g., an unsolvable puzzle) increases catecholamine response among healthy subjects. A positive result would suggest a biological explanation for the well-known relationship between emotional stress and CHD.

The *clinical trial* has a longer duration, ranging from days to years, and usually is restricted to a highly selected population, such as a group of screened subjects, diagnosed cases, or other volunteers. This type of experiment has the longest tradition in epidemiology and has acquired a substantial literature describing all aspects of study planning, implementation, analysis, and interpretation (e.g., Harris and Fitzgerald, 1970; Peto et al., 1976, 1977; Weddell, 1979). The major objective of the clinical trial is to test the possible effect (i.e., efficacy) of a therapeutic or preventive intervention. For example, we may wish to assess to what extent, if at all, a new type of chemotherapy prolongs the life of children with acute lymphatic leukemia. Additionally, the clinical trial may be employed to examine specific etiologic relationships having no immediate therapeutic or preventive component, if such trials are ethically feasible.

Lastly, the more recently developed *community intervention* has a long duration (at least six months) and differs from the previous types in that it is conducted within a particular sociopolitical context of a naturally formed population (community). Thus, the objectives of the community intervention usually pertain to implementation and assessment of interventions aimed at primary prevention through risk factor modification in a well-defined population. Generally, we wish to determine the potential benefits of modifying certain individual behaviors, biological characteristics, or aspects of the environment.

3.1.2 Advantages and Limitations

In general, experiments afford the most control over the study situation because they best enable the investigator to isolate the observed effect of the treatment (i.e., study factor). With a large enough sample size, a well-designed experiment can be expected to control nearly all distorting effects from extraneous risk factors, including those that are unmeasured. The control is accomplished either by holding these extraneous factors fixed for all subjects or by randomization—so that all treatment groups have, on the average, the same distribution of the extraneous factors. However, even well-designed experiments are not free of all distorting influences since our procedures are subject to human errors and chance. Thus, the investigator may have to control for the potentially distorting influences of extraneous risk factors in the analysis of an experiment, in much the same way that control is accomplished in nonexperimental studies (Rothman, 1977) (to be discussed in Chapter 16). In fact, control of extraneous risk factors through randomization alone is made particularly difficult in the community setting, where it may be practically impossible to manipulate certain

environmental or social factors (the treatment) while controlling for the effects of other risk factors. One way of handling this problem efficiently (in addition to analytic techniques) is to use a grouped design in which subjects are aggregated into relatively homogeneous groups before randomization. Thus, the experimental unit is actually the group, which may be any convenient aggregation, such as families, work groups, patient groups, or geographic areas (Sherwin, 1978; Cornfield, 1978).

One weakness of most experimental designs is related to their excessive control over the study situation—a weakness that, paradoxically, is also their strength. Because the experiment often takes place in an artificial setting among a selected sample, the study population may differ from the larger target population on several characteristics. If the true effect of the treatment depends on these characteristics, the observed effect in the study population could differ from the effect that exists in the target population. This difference may have important implications to public health activities, which are based on experimental findings.

The most often cited limitations of experiments are certain practical issues of implementing the design. Specifically, randomization may not be ethical if an arbitrary group of subjects must be denied an experimental treatment that is regarded as beneficial by clinicians or patients. In addition, the design may not be feasible, such as the artificial manipulation of a psychosocial attribute or the use of double blinding with a nonpharmacologic treatment. Solutions to these problems are sometimes offered by the use of special experimental options (e.g., Armitage, 1960; Zelen, 1979) or quasi-experimental designs.

3.2 QUASI EXPERIMENTS

A study in which the study factor has been artificially manipulated but for which randomization has not been used is called a *quasi experiment*, a term borrowed from the social sciences (Campbell and Stanley, 1963) (see Table 3.1). These designs may involve one-group comparisons, multiple-group comparisons, or a combination of these.

With a *one-group* (or internal) *comparison*, each experimental unit serves as its own control by observing the response variable before and after one or more interventions. For example, we could observe the motor vehicle fatality rate in a state before and after the enforcement of a new speed limit. (Note that the experimental unit here is the state.)

With a simple *multiple-group* (or external) *comparison*, treatment or intervention groups are compared with each other, as they are in a simple experiment without randomization. In this design, treatment groups are formed from convenience (e.g., geographic areas) or according to the voluntary behavior of subjects (e.g., elective surgery). Using the previous illustration, we could compare the motor vehicle fatality rates in 1975 among several states, some of which had the new speed limit enforced in 1974 and some of which did not.

A mixed design combines elements of both internal and external comparisons, thereby enhancing the potential for making a causal inference. For example, we could observe the absolute changes between 1973 and 1975 in the motor vehicle fatality rates for states in which the new speed limit was first enforced in 1974 and compare these changes to comparable trends in states for which the new speed limit was not enforced.

As we suggested above, a quasi-experimental study can be conducted in a variety of ways. For a more thorough understanding of important principles and applications, see Nagel and Neef (1979), Campbell and Stanley (1963), Isaac and Michael, (1971), Alwin and Sullivan (1975), and Campbell and Cook (1979).

3.2.1 Subtypes and Objectives

We may distinguish two subtypes of quasi experiments on the basis of how the manipulation or intervention has been done: (1) clinical or laboratory studies and (2) program or policy studies.

In *clinical/laboratory studies,* the investigators or their colleagues execute the intervention themselves with a specific group of subjects. Essentially, the clinical/laboratory quasi experiment is a clinical trial or laboratory experiment without randomization and has objectives similar to those of its experimental counterparts (see Table 3.1). This type of quasi experiment is often done when an experiment would be too costly, infeasible, or unethical. For example, we might test the efficacy of treating mild hypertension with dietary modifications by comparing the periodic blood pressure levels of a voluntary test group of mild hypertensives who receive special instruction on diet with the levels of another group of mild hypertensives who do not receive the instruction.

In a *program/policy study,* the intervention typically is planned and implemented by others not involved in the immediate investigation, generally with the intent of alleviating a social problem (which need not be health-oriented). This subtype of quasi experiment is analogous to a community intervention (experiment) without randomization and is closely connected to the process of health planning. The major objectives of a program/policy study are to evaluate the effectiveness of a planned intervention and to suggest program or policy changes in response to this evaluation. An experimental design often cannot be used in this situation, because the intervention has already been initiated when the study begins. Moreover, even when the evaluation is planned prior to the onset of the new program or policy, it is not generally consistent with program/policy objectives to deprive a portion of the target population of the treatment. Thus, the program/policy study is not used primarily to test the efficacy of a therapeutic or preventive measure (which is often assumed); rather, it is used to evaluate the extent to which public health goals have been achieved (i.e., the effectiveness of the intervention).

Because a program/policy study usually is directed toward a large target population, the intervention may not involve the direct manipulation of the known or suspected risk factor of the outcome event. Sometimes, the strategy is to modify or regulate the sociopolitical or physical environment in such a way as to produce desired changes among targeted individuals, which will, in turn, have an impact on the health status of the population. For example, since we know that blood pressure level is an important determinant of cardiovascular disease, we might wish to evaluate a statewide program aimed at educating the populace about the etiology, effects, and treatment of hypertension.

3.2.2 Advantages and Limitations

As already implied, the chief advantage of the quasi experiment over the experiment is the smaller number of practical obstacles. Thus, in general, quasi experiments are likely to be more feasible and less expensive for conducting large studies—and, occasionally, they may be the only alternative.

On the other hand, the lack of randomization also means that the investigator has less control over the influence of extraneous risk factors. Actual control, however, can vary considerably among quasi-experimental designs. On one extreme, failure to randomize can introduce serious distortion in the results that cannot be corrected in the analysis. For example, a nonrandomized comparison of a new drug with traditional therapy for treating diagnosed cases might inadvertently favor the new drug if that test group consisted of cases with better prognoses. Conceivably, this distortion could result if a physician who advocates the new drug is the person who diagnoses the cases, allocates patients into groups, or measures the response.

Some *natural experiments* are quasi experiments,* and they may afford as much control as a true experiment. The design is called a natural experiment because the allocation process appears to be random, although no deliberate attempt was made to randomize. For example, consider the work done by John Snow in London during the middle of the nineteenth century (Snow, 1855). Snow believed that the cholera outbreaks of that period were due to contaminated water supplies, and he took advantage of a natural experiment to test his hypothesis. Shortly before the 1854 epidemic, one of the companies that supplied water to London residents changed its source of intake to the upper part of the Thames, where the water was noticeably cleaner. Fortunately, the houses supplied by that company were randomly distributed throughout London and were not set off from the other houses by divisions of economic status, factory affiliation, or other factors. In fact, very frequently, the residents did not know from which companies they got their water. Using the records of the water companies, census information, and the exact residential location of each new case fatality, Snow was able to compare the mortality rates for cholera

*Natural experiments may also be observational studies.

among consumers of the various companies. As predicted, he found a lower mortality rate within homes that received the cleaner water.

3.3 OBSERVATIONAL STUDIES

Epidemiologists most often use the *observational study,* in which there is no artificial manipulation of the study factor. While an observational study can take many forms, as we will see in Chapter 4, the simplest design closely resembles an experiment or quasi experiment. For example, we could identify a group of elderly hypertensives and a group of elderly normotensives from a baseline exam and follow both cohorts for several years. If the observed relative frequency of senile dementia is greater in the former group, we might infer that blood pressure is a risk factor for the disease. In addition, we could observe the association between blood pressure change or variability (without intervention) and the subsequent development of senile dementia.

Because observational studies are central to epidemiologic research, a more thorough presentation of these designs will be undertaken in Chapters 4 and 5. You may also wish to consult MacMahon and Pugh (1970), Susser (1973), Lilienfeld and Lilienfeld (1980), Cochran (1965), and McKinlay (1975).

3.3.1 Subtypes and Objectives

Observational studies are commonly divided into two subtypes on the basis of the degree of a priori knowledge regarding the disease.

A *descriptive study* usually is conducted when little is known about the occurrence, the natural history, or the determinants of a disease. The objectives are to estimate the disease frequency or time trend in a particular population and to generate more specific etiologic hypotheses. For example, since we have limited knowledge of the mental health of the elderly, we might plan a community study to estimate the relative frequency of specific mental disorders. Suppose we found that a particular type of depression was very common in this age group, compared to the general population. We might initiate future investigations to explain the relative excess of this disorder and to discover other social and health implications.

An *analytic* (or etiologic) *study* is conducted when enough is known about the disease before the investigation so that specific a priori hypotheses can be tested. The objectives are to identify risk factors for the disease, estimate their effects on the disease, and suggest possible intervention strategies. An example of this subtype is the hypothetical senile dementia study mentioned above, for which high blood pressure was hypothesized to be a risk factor for the disease.

While descriptive and analytic studies are often treated as two mutually exclusive categories, they are, in fact, opposite ends of a continuum. At one end, we know very little about the disease and are searching for clues; at the other end, we know a great deal about the disease and are testing specific hypotheses. Very often, as our understanding of a disease

increases, an investigation planned as a descriptive study will become more analytic after data collection begins.

3.3.2 Advantages and Limitations

Observational research is often the most practical or feasible to conduct because the study factor is not manipulated. However, for reasons that will be discussed in Chapter 4, these studies are not uniformly less expensive nor do they always take less time than other types of studies. Another potential advantage is that observational studies are often carried out in more natural settings, so that the study population is more representative of the target population. As we suggested in Section 3.1.2, this feature has important implications to health planners and policymakers who base their decisions partly on the results of epidemiologic investigations.

The major limitation of observational designs is that they afford the investigator the least control over the study situation; therefore, results are generally more susceptible to distorting influences. Consequently, a greater burden is placed on the investigator, particularly in the analysis stage, to deal with these potential sources of error that threaten the validity of his or her findings.

Because the investigator achieves relatively little control in the design stage, observational studies tend to be unique, making them very difficult to replicate and, therefore, making scientific generalizations less secure (Jamison, 1980). Yet epidemiologists can take advantage of this replication problem by using apparently inconsistent results to refine their hypotheses.

3.4 CONCLUDING REMARKS

In this chapter, we have briefly outlined three broad categories of empirical research: experiments, which involve randomization of study factor categories (or treatments) among all subjects; quasi experiments, which involve artificial manipulation of the study factor without randomization; and observational studies, which involve no artificial manipulation of the study factor. Since epidemiologists, overall, conduct more observational than nonobservational studies, the next two chapters will deal with the specific options and types of observational study designs. Throughout the remainder of this book, in fact, all principles and methods of epidemiologic research are presented in terms of observational studies. Nevertheless, most of the topics and discussions that follow, including validity considerations in Part II and analysis in Part III, apply also to nonobservational studies.

REFERENCES

ALWIN, D. R., and SULLIVAN, M. J. 1975. Issues of design and analysis in evaluation research. *Sociol. Meth. & Res.* 4: 77–100.

ARMITAGE, P. 1960. *Sequential medical trials*. Springfield, Ill.: Thomas.

CAMPBELL, D. T., and COOK, T. D. 1979. *Quasi-experimentation: Design and analysis for field settings*. Chicago: Rand McNally.

CAMPBELL, D. T., and STANLEY, J. C. 1963. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

COCHRAN, W. G. 1965. The planning of observational studies of human populations. *J. R. Stat. Soc. A* 138: 234–265.

COMSTOCK, G. W. 1978. Uncontrolled ruminations on modern controlled trials. *Am. J. Epidemiol.* 108: 81–84.

CORNFIELD, J. 1978. Randomization by group: A formal analysis. *Am. J. Epidemiol.* 108: 100–102.

FARQUHAR, J. W. 1978. The community-based model of life style intervention trials. *Am. J. Epidemiol.* 108: 103–111.

HARRIS, E. L., and FITZGERALD, J. D., eds. 1970. *The principles and practice of clinical trials*. Edinburgh: Livingstone.

ISAAC, S., and MICHAEL, W. B. 1971. *Handbook in research evaluation*. San Diego: EDITS.

JAMISON, J. R. 1980. The use of inferential statistics in health and disease: A warning. *South Afr. Med. J.* 57: 783–785.

KEMPTHORNE, O. 1978. Logical, epistemological and statistical aspects of nature-nurture data interpretation. *Biometrics* 34: 1–23.

LILIENFELD, A. M., and LILIENFELD, D. E. 1980. *Foundations of epidemiology*. 2nd ed. New York: Oxford University Press.

MCKINLAY, S. M. 1975. The design and analysis of the observational study—A review. *J. Am. Stat. Assoc.* 70: 503–523.

MACMAHON, B., and PUGH, T. F. 1970. *Epidemiology: Principles and methods*. Boston: Little, Brown.

NAGEL, S. S., and NEEF, M. 1979. *Policy analysis in social science research*. Beverly Hills, Calif.: Sage; App. 3.1.

PETO, R.; PIKE, M. C.; ARMITAGE, P.; BRESLOW, N. E.; COX, D. R.; HOWARD, S. V.; MANTEL, N.; MCPHERSON, K.; PETO, J.; and SMITH, P. G. 1976. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I: Introduction and design. *Br. J. Cancer* 34: 585–612.

———. 1977. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II: Analysis and examples. *Br. J. Cancer* 35: 1–38.

ROTHMAN, K. J. 1977. Epidemiologic methods in clinical trials. *Cancer* 39: 1771–1775.

SARACCI, R. 1979. Controlled studies. In *Health care and epidemiology*, edited by W. W. Holland and L. Karhausen, Chap. 13, pp. 194–207. Boston: Hall.

SHERWIN, R. 1978. Controlled trials of the diet-heart hypothesis: Some comments on the experimental unit. *Am. J. Epidemiol.* 108: 92–99.

SNOW, J. 1855. *On the mode of communication of cholera*. 2nd ed. London: Churchill.

SUSSER, M. 1973. *Causal thinking in the health sciences*. New York: Oxford University Press.

SYME, S. L. 1978. Life style intervention in clinic-based trials. *Am. J. Epidemiol.* 108: 87–91.

WEDDELL, J. 1979. Experimental studies. In *Health care and epidemiology*, edited by W. W. Holland and L. Karhausen, Chap. 14, pp. 208–225. Boston: Hall.

WENNBERG, J. E.; BUNKER, J. P.; and BARNES, B. 1980. The need for assessing the outcome of common medical practices. *Annu. Rev. Public Health* 1: 277–295.

ZELEN, M. 1979. A new design for randomized clinical trials. *N. Engl. J. Med.* 300: 1242–1245.