UNIT 4 DESIGN STRATEGIES IN RESEARCH: ANALYTIC STUDIES

Structure

4.1 Introduction

4.2 Analytic Studies

4.3 Observational Studies

- 4.3.1 Cohort Study
- 4.3.2 Case-Control Study
- 4.3.3 Analytic Cross-Sectional Studies

4.4 Experimental/Intervention Studies

- 4.4.1 Issues in the Design and Conduct of Clinical Trials

4.5 Let Us Sum Up

4.6 Glossary

4.7 Answers to Check Your Progress Exercises

4.1 INTRODUCTION

In the last unit we provided a brief overview of the various design strategies used in epidemiological research, with a major focus on descriptive studies. Here in this unit we will dwell on analytic studies, their types, characteristics and strength and limitations. Unlike descriptive studies, which seeks to delineate the magnitude of the problem in different segments of the population, say in terms of prevalence and incidence, and to establish normal or abnormal levels of measurements, the analytic studies aim to evaluate the relationship between a cause and an effect, an exposure (a risk factor or a protective factor) and a disease, of an outcome. There are a number of specific analytic study design options that can be employed. These include the observational epidemiological research designs (namely the cohort studies, the case-control studies and the cross-sectional studies) and the experimental study designs. In this unit we will learn about these analytic study designs their characteristics and strength and limitations and about the issues in the design and conduct of experimental studies.

Objectives

After studying this unit, you will be able to:

- differentiate between the analytic and descriptive epidemiological study designs,
- classify the different analytic design strategies in epidemiological research,
- define and discuss the basic characteristics of analytic studies in epidemiological research,
- describe the different types of analytic studies which can be conducted at the population and individual level,
- enumerate the strengths and weaknesses of the analytic studies, and
- elaborate on the issues to be considered in the design and conduct of experimental studies.

4.2 ANALYTIC STUDIES

What are analytic studies? How do they differ from descriptive studies? Can you suggest! Well let us try to understand this with the help of an example. In the last unit, you may recall studying that case study/reports and case series' (a type of descriptive study) are the clinical route to definition and recognition of disease entities

and to the *formulation of hypotheses*. We explained this with the help of the autoimmune deficient syndrome (AIDS) example. The case series', we learnt, led to an initial AIDS case definition for the purposes of identifying additional cases and inaugurating surveillance. Now with the case definition in hand and the hypothesis have been advanced, analytic studies are the next epidemiologic recourse. It was possible to conduct case-control studies (a type of analytic study) in which persons with the disease could be compared with persons without the disease and characteristics associated with the condition identified. Comparisons of AIDS cases to apparently healthy male homosexual controls indicated that the cases had higher numbers of partners, had greater involvement in certain sexual practices, and more exposure to drugs used to enhance sexual pleasure. These findings led to analytic studies to test these and other exposure hypotheses.

Now then having gone through the example presented above, can you now suggest how analytic studies differ from descriptive studies? Yes, analytic epidemiological studies focus on the *determinants of a disease* by *testing the hypothesis formulated from descriptive studies*, with the ultimate goal of judging whether a particular exposure causes or prevents diseases. Analytic study, therefore, is a *comparative study* (uses comparison groups) intended to *identify and quantify associations*, *test hypotheses* (in contrast to hypothesis formulation), and *identifies causes* (determinants).

There are various types of analytic study design options which you as a researcher can adopt. These options are primarily divided into two broad design strategies, namely *observational* and *experimental/intervention* as you may have already seen earlier in Figure 3.1 in Unit 3, which illustrates the epidemiological research designs. Figure 4.1 highlights the two major groups of analytic studies.

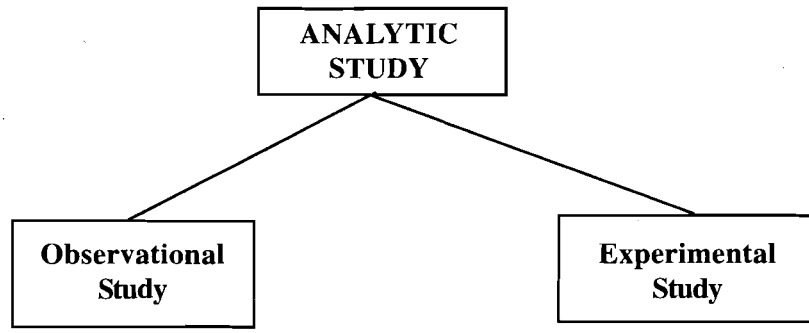

Figure 4.1: Two major groups of analytic studies

The term *observational study* is used for an analytic study which investigates the natural course of event/disease. Opposed to these are *interventions or experimental studies*, which necessarily entail an intervention. Further, the major difference between the two lies in the role played by the investigator and/or researcher. In observational study, the researcher simply observes the natural course of event(s) as mentioned above, noting who is exposed/non-exposed and who has/has not developed the outcome of interest. In contrast, in experimental studies, the investigator(s) themselves allocate the exposure and then follow-up the cases for the subsequent development of the disease/event. Note, if an intervention is "assigned" through the research, it is an experiment. If it is chosen, then the study type is observational.

We will learn more about these two categories of analytic studies in the subsequent sections, sub-sections. We shall begin with on overview of observational studies.

4.3 OBSERVATIONAL STUDIES

Consider the following examples:

- In a research study, 33 health care workers who became seropositive to HIV after percutaneous exposure to HIV-infected blood were compared to 665 health care workers with similar exposure who did not become seropositive.
- Subjects (80,082 women) between the ages of 34 and 59 years were followed for 14 years to look for instances of non-fatal myocardial infarction or death from coronary heart disease. These women were divided into low, intermediate, and high groups on the basis of their consumption of dietary fat.

Having read these examples, can you suggest what type of analytic study are the two epidemiological studies highlighted above? Yes, in the first example, since the researchers did not control who became seropositive, this is an observational study. Similarly, since the women themselves controlled their diets, rather than having a diet imposed on them by the researchers in the second example, this also represents an observational design. Thus, you may have realized by now that a study is observational *when the natural course of event is monitored without any intervention.*

Observational studies may thus be defined *as an analytic epidemiological study in which observations are made but investigators do not control the exposure or intervention and other factors.* Changes or differences in one characteristic are studied in relation to changes or differences in others, without the intervention of the investigator. There are three ways in which analytical observational studies can be carried out. Refer to Figure 4.2, which highlights the three observational study designs.

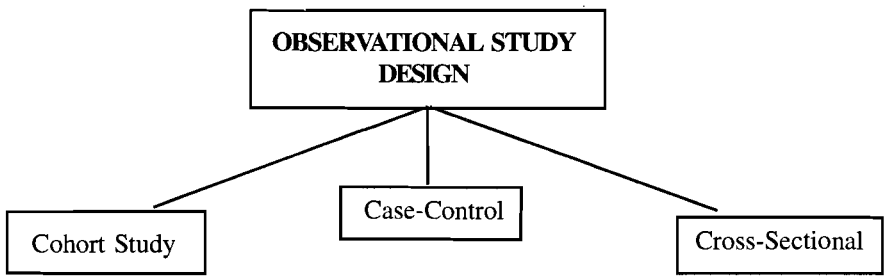

Figure 4.2: The three observational study designs

A detail review of the three types of observational studies follows in the subsequent sub-sections, but here a brief orientation to the three types follows. Consider the example given herewith. Suppose we have undertaken an observational study for studying the relationship between maternal anaemia and birth weight. Now there can be three ways in which this study can be carried out as indicated in Table 4.1, and the discussion which follows.

Table 4.1: The three experimental study designs for studying the relationship between maternal anaemia and birth weight

Birth Weight	A (Cohort)			B (Case-control)			C (Cross-sectional)		
	Maternal Anaemia			Maternal Anaemia			Maternal Anaemia		
	Yes	No	Total	Yes	No	Total	Yes	No	Total
< 2.5 kg			?			n1			?
≥ 2. 5 kg			?			n2			?
Total	**n1**	**n2**	**n**	**?**	**?**	**n**	**?**	**?**	**n**

Source: Indrayan A, Satyanarayana L. Designs of Medical Studies. Indian Pediatrics 1999; 36:691-696.

Study A: In Table 4.1, under the category A, you notice that group of individuals (women) are defined on the basis of presence or absence of exposure (anaemic and non-anaemic) and the subjects are then followed-up over a period of time to assess the occurrence of the outcome (i.e. birth weight of the infants). A study following such a design is called a *cohort study*. Note data related to the number of children in various birth weight categories (i.e. the outcome) born to these women become known only after the data is collected. This is a natural course since outcome in any case occurs afterwards. Such a study may be referred to as a *prospective cohort study*.

Study B: In this design, the relationship between maternal anaemia and birth weight is investigated by first enlisting infants (n1 and n2) of different birth weights (as can be seen in Table 4.1, column B) and then their mothers are compared with respect to the proportion having a history of an exposure i.e. anaemic or non-anaemic. You would have realized that in such a design, what we have done is that from a known outcome (birth weight), we have tried to assess the exposure (i.e. maternal anaemia status). A study using this design is called a *case-control study*.

Study C: In the third case (refer to Table 4.1 under column C), the method adopted is to take a sample of total women without any consideration of anaemia status (i.e. anaemic or non-anaemic). This sample is then surveyed over a period of time and the outcome i.e. the number of women with and without anaemia and with babies of different birth weight would be known only after the survey is completed. Such a study is called observational *cross-sectional study*.

Having gone through the example presented above, surely you would have got a brief insight into the types of observational studies used in epidemiological research. Next, we shall dwell further and elaborate on these study designs highlighting their strengths and limitations. We begin our study with cohort studies.

4.3.1 Cohort Study

Let us begin by understanding what is a cohort. *A cohort is a group of people who share a common characteristic or experience within a defined time period* (e.g., are born, leave school, lose their job, are exposed to a disease, drug or a vaccine, etc.). In our example presented above, women were defined on the basis of presence or absence of exposure (anaemic and non-anaemic) and followed-up to study the relationship between maternal anaemia and birth weight. This group of women can be called a cohort.

Cohort study, therefore, *is an epidemiology study that observes a large group of people over a period of time*. A group or groups of people/individuals are defined on the basis of presence or absence of exposure to a suspected risk factor for a disease and eligible participants are then followed-up over a period of time to assess the occurrence of the outcome. Figure 4.3 illustrates the cohort study design.

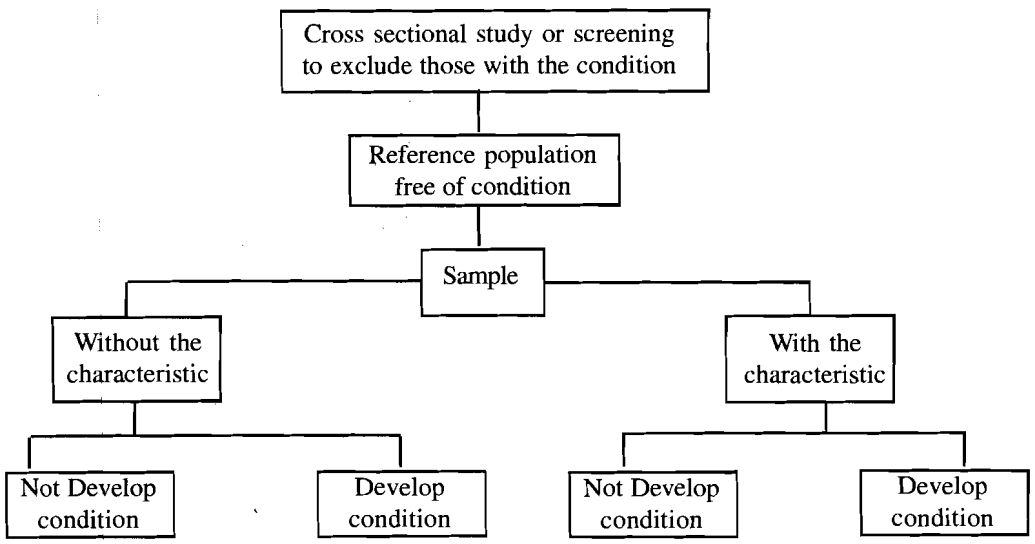

Figure 4.3: Cohort study design

From Figure 4.3, it must be evident to you that:

a) The group of person (cohort) to be studied are defined in terms of characteristics manifest prior to appearance of disease under study, and

b) These defined study groups are observed over a period of time to determine and compare the frequency of study disease among them.

The objective of cohort studies is to test the hypothesis regarding the causation of the disease. The Framingham Heart Study (USA) is a landmark investigation of *cohort* epidemiological research undertaken to identify the common factors or characteristics that contribute to coronary vascular disease (CVD) by following its development over a long period of time in a large group of participants. The researchers recruited 5,209 men and women between the ages of 30 and 62 from the town of Framingham, Massachusetts, USA and began the first round of extensive physical examinations and lifestyle interviews in 1948 that they later analyzed for common patterns related to CVD development. In 1971, the study enrolled a second-generation group – 5,124 of the original participants' adult children and their spouses – to participate in similar examinations. A Third Generation (the children of the Offspring Cohort) were recruited (2002) and examined, seeking to further understand how genetic factors relate to cardiovascular disease. Monitoring of the Framingham Study population has led to the identification of the major CVD risk factors – high blood pressure, high blood cholesterol, smoking, obesity, diabetes, and physical inactivity – as well as, a great deal of valuable information on the effects of related factors such as blood triglyceride and HDL cholesterol levels, age, gender and psychosocial issues.

Cohort studies, it is important to note, can be classified as either *prospective* or *retrospective*, depending on the relationship between the initiation of the study and the occurrence of the disease. In *retrospective cohort study*, the relevant events (both the exposure and the outcome of interest) have already occurred when the study is initiated. In *prospective study*, the relevant exposure may not have occurred at the time the study was initiated, but the outcome has certainly not yet occurred. Thus, after the selection of the cohort, the study group must be followed into the future to assess outcome, which may be the incidence rate of disease or for that matter the birth weight as illustrated in the example presented in Table 4.1 above.

As a consequence of this design, cohort studies offer a number of advantages. These include:

- can assess multiple outcomes (effects) of single exposure i.e. can provide information on the full range of health effects of a single exposure,
- can demonstrate temporal relationship between exposure and disease,
- are logical and allows direct measurement of incidence of disease in the exposed and the non-exposed population,
- it is suitable for the study of rare exposures, and
- if prospective, it minimizes the potential for selection bias and biases in the ascertainment of exposure.

Since cohort studies often involve following large number of individuals over a large period of time, they are generally very time consuming and expensive. Other limitations of the study design include:

- serious potential for bias associated with the losses to follow-up that are likely to occur when participants must be followed for month, year or even decades,
- validity of the result can be affected by loss of follow-up,
- is inefficient for the evaluation of rare diseases, unless large sample size included,
- if retrospective, requires the availability of existing records,

With a review of the cohort studies let us next focus on case-control observational studies.

4.3.2 Case-Control Study

The demonstration of the link between tobacco smoking and lung cancer, was first suggested by the case-control studies undertaken by *Sir Richaed Doll* and others after him, What is a case-control study and what is its importance in epidemiological research? We already know that a *case-control study* is a type of *observational analytic study* in which researchers use subjects who already have a disease or other condition and look back to see if there are characteristics of these subjects that differ from those who don't have the disease (as you may have also noticed in the example presented in Table 4.1 above). In other words, case-control studies use subjects who already have a disease or other condition and look back to see if there are characteristics of these subjects that differ from those who don't have the disease. Here you may have noticed that such as design moves in a reverse direction from the effect to ascertain the possible causes. Such a design, therefore, is also referred to as a *retrospective study* design by some investigators.

The starting point of most case control studies is the identification by researchers of an *outcome* or *effect* (e.g. lung cancer, heart disease, anaemia etc.) and a number of potential causative factors. If investigating lung cancer, for example, the factors selected might include smoking history, asbestos exposure etc.

The next important point in case-control studies is the identification of cases. A group of *cases* is selected, containing individuals who exhibit the outcome under investigation. Using medical records or interviews, researchers record the variables identified as risk factors, plus other non-risk variables which can then be used to select matching controls. Typically these would be demographic variables such as age, sex, race, income bracket, geographic area of residence etc.

A number of *control subjects* (or *controls*) are then chosen who do not exhibit the outcome or effect under investigation – there may be one or more per *case subject.* These controls should match the cases as closely as possible with respect to the non-risk variables (such as age, sex, race, income etc.); this allows the proposed non-risk variables to be ignored in the analysis. Sometimes more than one control group can be used. The case and control groups are then compared on the proposed causal factors, and statistical analysis is used to estimate the strength of association between each factor and the studied outcome.

This approach/design was developed as a response to needs that accompanied the shift from acute to chronic diseases as major public health problem. The case-control design offered a solution to the difficulties of studying diseases with very long latency period, since researchers could identify affected and unaffected individuals and then look backward in time to assess the important exposure, occurrence and/or event in one's early life rather than having to wait a number of years for the disease to develop. Thus case-control studies are particularly efficient and are of utmost importance to epidemiological research and their advantages include:

- well suited to the study of rare diseases or those with long latency period (such as cancer, heart diseases etc.) as already mentioned above,
- relatively quick to conduct as compared with other analytic design,
- the long period of follow-up is avoided, thus relatively inexpensive,
- require comparatively few subjects, existing records can be used,
- knowledge about the early stages of the development about a particular disease or outcome of interest is provided, and

- allows study and evaluation of multiple potential causes (factors) of disease as well as, the interrelationship among these factors.

Though case-control studies are generally considered more efficient, the major potential problems/limitations of this study design include:

- susceptible to bias from the differential selection of either the case or control,
- differential reporting or recording of exposure information between study groups, since it relies on records or recall for information on past exposure,
- is inefficient for the evaluation or rare exposure, unless the attributable risk percent is high,
- cannot directly compute incidence rates of disease in exposed and unexposed individuals, unless study is population based,
- validation of information is difficult to assess or sometimes impossible,
- selection of an appropriate comparison group may be difficult, and
- cannot establish sequence of events.

Despite these limitations, the case-control studies, you would notice, have become the most common analytic epidemiological study designs because of their advantage in being able to evaluate disease that occur many years following relevant exposure in a timely and cost-effective manner.

Finally, we move on to the third analytic observational study design, namely cross-sectional studies. You may recall studying about the cross-sectional studies/surveys under the descriptive studies as well. Here, the next sub-section we shall study about the analytic cross-sectional studies.

4.3.3 Analytic Cross-Sectional Studies

You may have realized by now that there are two types of cross-sectional studies, one mostly descriptive study about which we have already discussed in sub-section 3.3.3 in Unit 3 earlier and the second the analytic cross-sectional disease.

Unlike in the descriptive cross-sectional study, which is thought of as providing a "snapshot" of the frequency and characteristics of a disease in a population at a particular point in time, in *analytic cross-sectional research design*, the status of an individual with respect to the presence or absence of both exposure and disease is assessed (simultaneously) at the same point in time. Analytic cross-sectional studies, thus, differ from solely descriptive cross-sectional studies in that they compare the proportion of exposed persons who are diseased with the proportion of non-exposed persons who are diseased.

Analytical cross-sectional studies are, therefore, those in which data on the prevalence of exposure and disease are obtained for the purpose of comparing disease differences between exposed and non-exposed. Comparison of differences is the analytical component of these studies.

This type of cross-sectional design is particularly well suited to acute conditions with short latent period such as typhoid and measles or to those chronic diseases that are not fatal (for example, congenital malformations). Analytic cross-sectional study is generally considered a rapid and an inexpensive way to provide clues for further and more valid investigations. Note, it seldom provides explicit answers.

Since exposure and disease status are measured at the same point in time, it may not always be possible to distinguish whether the exposure proceeded or followed the disease. Hence, cross-sectional surveys have certain *limitations* in terms of the following:

- information on all factors is collected simultaneously so it can be difficult to establish a putative "cause' antedated the "effect',
- susceptible to selection bias (e.g. selective survival),
- susceptible to misclassification (e.g. recall), and
- not good for rare diseases and rare exposures.

With a review of the cross-sectional studies, we end our study of analytic observational studies here. Before we move on to the experimental studies, the second category of analytic study design, let us recall what we have learnt so far by answering the check your progress exercise 1.

Check Your Progress Exercise 1

1) What is an analytic study? List the various analytic studies you may carry out while undertaking epidemiological research.

.....

.....

.....

.....

2) What is the objective of conducting a cohort study? Look up any reference book/journal or the internet and give one example of a cohort study undertaken in epidemiological research.

.....

.....

.....

.....

.....

3) Explain the process involved in case-control study. List any two advantages and two limitations of case-control studies.

.....

.....

.....

.....

.....

4) Differentiate between descriptive and analytic cross-sectional studies.

.....

.....

.....

.....

.....

Now with a comprehensive understanding about analytic observational epidemiological research designs, we next move on to the study of experimental analytic study designs.

4.4 EXPERIMENTAL/INTERVENTION STUDIES

Earlier in section 4.2, we learnt that analytical studies can be *observational* and/or *experimental/interventional.* Experiment *is an investigation of the effect of a deliberate intervention or a stimulus so as to change the course of event.* In an experimental design, the investigator keeps a control on the allocation of the experimental units to different types of intervention, as contrast to observational study design where the researcher is merely a passive observer who simply observes the natural course of event(s), noting who is exposed/non-exposed and who has/has not developed the outcome of interest. Thus you would have realized that in experimental studies, there is active assignment of participants to a particular treatment/exposure, the investigator(s) themselves allocate the exposure and then follow-up the cases for the subsequent development of the disease/event. A cause-effect type of relationship can be thus easily inferred.

An experiment, as you may already know, can be carried out in a laboratory, clinic or in the field. When the subjects for experiments in the field, lab or clinic are human beings, such experiments are generally termed as '*trials*'. The term "*clinical trial*" emphasizes the controlled aspect of the intervention, at the expense of the generalizability of the results. The term "*community trial*" emphasizes that the trial is carried out in a realistic setting and results may therefore be more generalizable (at the expense of having control over what subjects actually do). A community trial can involve an individual-level intervention (e.g., breast cancer screening), a community-level intervention (e.g., anaemia control), or interventions with elements of both levels (e.g., mass media promotion of physical exercise).

The intervention/trials can be *therapeutic* or *preventive/rehabilitative. Therapeutic trials* are conducted among patients with a particular disease to determine the ability of an agent (drug, diet, supplement etc.) or a procedure to diminish symptoms, prevent recurrence, or decrease risk of death from that disease. Look at Figure 4.4 which illustrates a therapeutic trial involving tuberculosis patients.

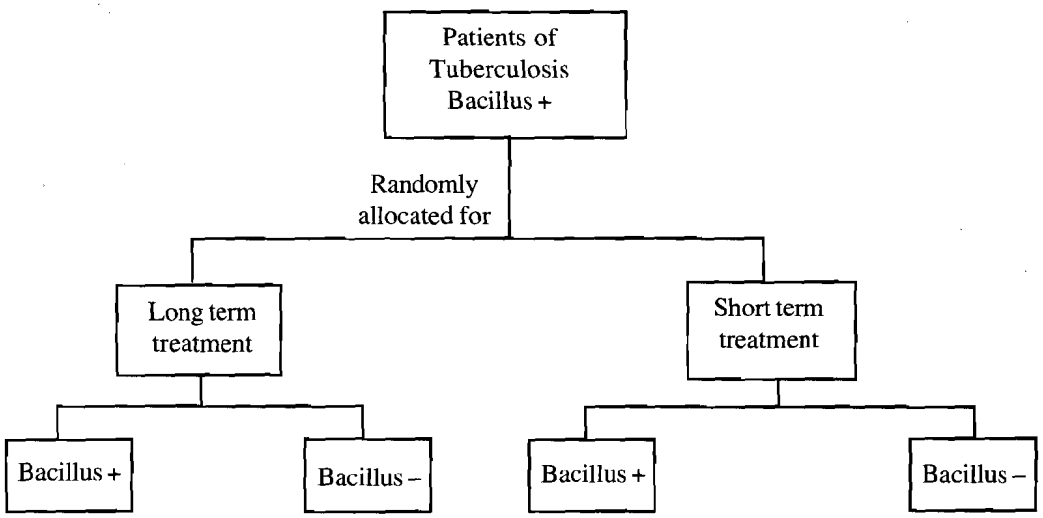

Figure 4.4: A therapeutic trial

Preventive trial, on the other hand, involves the evaluation of whether an agent/ substance or procedure reduces the risk of developing disease among those free from that condition at enrolment. Figure 4.5 illustrates the preventive trial.

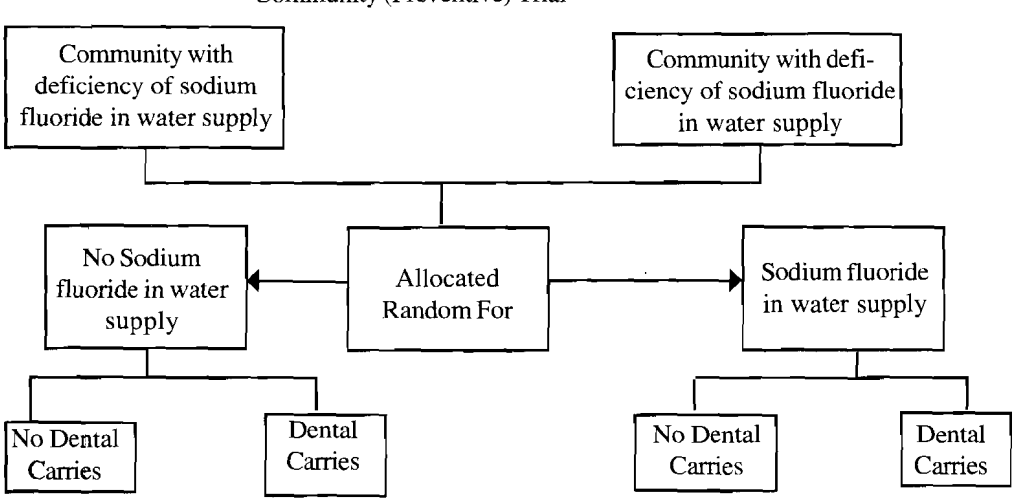

Figure 4.5: The preventive trial

As already highlighted above, clinical trials can be conducted at an individual or community level. While *therapeutic trials* are virtually always conducted among *individuals,* primary *preventive trials* can be carried out among either individuals or *entire populations* as in the case of Newburg-Kingston dental carries study illustrated in Figure 4.5 above. One community (Newburg) in this trial was allocated at random to receive sodium fluoride added to water supply, while the other (Kingston) continued receiving water without supplementation. This clinical trial indicated important and statistically significant reduction in the development of decayed, missing or filled teeth in the community receiving fluoridation.

Next, we shall consider a few issues linked to the design and conduct of clinical trials.

4.4.1 Issues in the Design and Conduct of Clinical Trials

To ensure valid and reliable results, there are a few issues which we may like to consider in the design and conduct of clinical trials. These include selection of the study population, allocation of the treatment regimens, maintenance and assessment of compliance, and achieving high and uniform rates of ascertainment of outcome.

The first important issue is the *selection of a study population* among whom the intervention study is to be conducted. Consider the population hierarchy illustrated in Figure 4.6, which highlights a number of interrelated populations, from which the intervention study group would be derived. At the top of the hierarchy, is the reference population i.e. the general group to whom the study findings will be universally applicable. This reference population may include all human beings or it may be restricted by geography, age, sex or any other characteristic that may have been thought to modify the magnitude of the effect seen in the trial. The experimental population is the actual group in which you would carry out the trial. It is essential to determine whether the proposed experimental population is sufficiently large to achieve the necessary sample size for the trial. It is also essential to choose an experimental population that will experience a sufficient number of outcomes of interest to permit meaningful comparison between the various treatments/interventions within a reasonable period of time and lastly choose an experimental population from which we can likely obtain complete and accurate follow-up information for the duration of the trial. Once the experimental population has been defined, subjects are invited to participate in the study after being fully informed as to the purpose, benefits, possible risks and the research procedure. Those willing to participate must be then screened for eligibility based on a predetermined criterion (i.e. factors such as previous history of any end point under study, a definite need for the study treatment, any contradictions etc.). Those who are eventually willing and determined to be eligible

to enroll in the trial compose the actual *study population*, which you would notice are often a relatively small sub-group of the experimental population. Subjects from this study population are then allocated into the various treatment groups as shown in the population hierarchy in Figure 4.6. The effects of a treatment, procedure or programme can be compared with those of one or more of a variety of groups, such as another dose of the same supplement/drug, another strategy or an approach or a placebo.

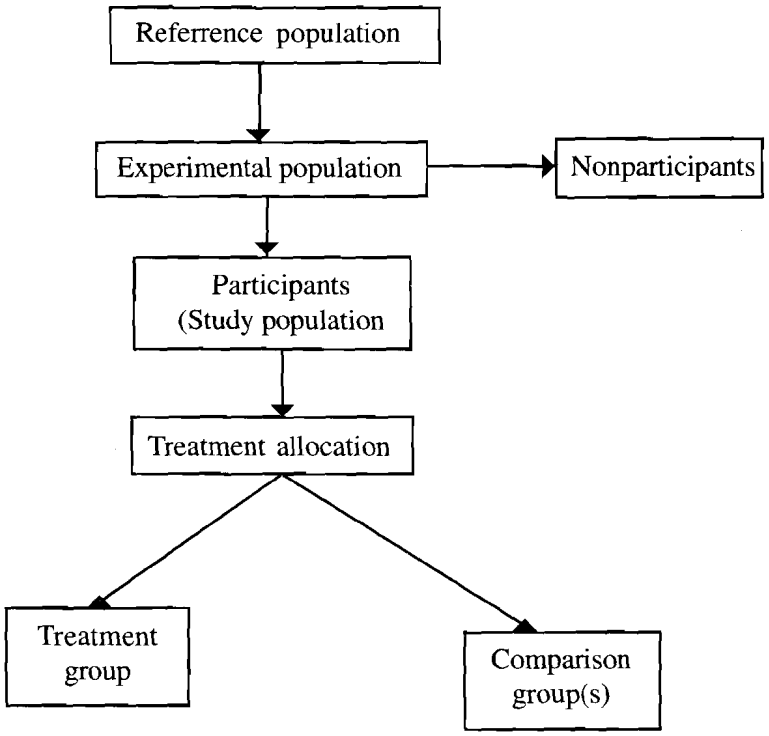

Figure 4.6: Population hierarchy for an experimental/intervention study

Allocation of study regimens i.e. allocation of participants into the various treatment groups is another issue to be considered in the conduct of intervention studies. To maximize the probability that the groups receiving these differing interventions will be comparable, allocation of subjects to these different interventions should be random. Random assignment means that each individual (from the study population) has the same chance of receiving each of the possible intervention and that the probability that a given subject will receive a particular allocation is independent of the probability that any other subject will receive the same intervention/treatment. *Randomization* is therefore a cardinal principle of experimentation, wherein the potential for bias in allocation to study groups is removed, and investigators can be confident that observed differences are not due to the selection of particular patient. Study groups will tend to be comparable with respect to all variables (geography, age, sex, race etc.) except for the intervention being studied. Two methods most commonly used for random assignment are the use of *a table of random numbers* or by *draw of lots*. Use of a computer-generated randomization list is also practiced. We will learn more about this later in Unit 6.

Thus, we have seen that randomization tend to remove the possible bias of the investigator in allocation and observing subjects with different treatments. Randomization is the only possible way to achieve control for any influence of unknown variable. When the sample size is sufficiently large, both known and unknown confounding factors are distributed equally among treatment groups, which provides a degree of assurance about the comparability of the study groups that is simply not possible in any observational study design. Thus, there is an inherent confidence in the results of a well-designed and conducted randomized trial. In fact, you would realize that randomization is a necessary ingredient for validity of the statistical methods that are used for data analysis.

An intervention study also requires the active participation and cooperation of the study subjects. Consequently, the *maintenance and assessment of compliance* is yet another issue to be considered in the design and conduct of all clinical trials and intervention studies. After agreeing to participate, subjects may deviate from the protocol for a variety of reasons, including side effects, migration, forgetting to take the treatment, or a long intervention period or simply withdrawing their consent after randomization. This deviation from the protocol may influence the study results. Therefore, possible startegies should be adopted to try to enhance the compliance among the participants in the trial. As already mentioned earlier in this sub-section, selection of population of individuals who are both interested and reliable can enhance compliance. Other way of attempting to increase compliance would include frequent contact with subjects by home or clinic visit, telephone or mail or by giving incentives (detailed information not ordinarily available from their usual sources of health care).

Monitoring compliance is important since noncompliance will affect the results and decrease the statistical power of a trial to detect any true effect of the study treatment. Thus the assessment of any intervention study/clinical trial must take into account the extent to which there was adherence to the intervention protocol by the study groups. The higher the degree of compliance with the offered intervention, the greater the extent to which observed difference between those allocated to differing treatment reflect real differences in the effect of the treatment themselves.

Finally another crucial issue to be considered in the design and conduct of an intervention study is the *ascertainment of the outcome(s) of interest.* Ascertainment of the outcome may require a *complete follow-up of study participants* over the duration of the trial. If the proportion of outcomes that are not ascertained is large or differs among the study groups, the result could be an under- or overestimate or even, by chance, reflect the true effect. To avoid this situation, where it is not possible to know the direction of the bias, it is crucial to keep the numbers of individuals lost to follow-up to an absolutely minimum.

Further, there can be the *potential for observation bias in the ascertainment of outcome* in an intervention study in that knowledge of an participant treatment status, might consciously or not, influence the observation, identification or reporting of relevant issues in other term biased inferences as illustrated in Figure 4.7. If the researcher knows that a particular subject is a case or is a control, this may well affect the way the questions are asked, investigations done or interpretations made. There is a tendency in the subjects too to respond differently depending upon whether they are in treatment group or in placebo group. To control this bias, one precaution that can be taken is in the designing of the placebo. The placebo can be so designed so as to look like the treatment so that there are no apparent differences for the subjects or for any research staff to detect.

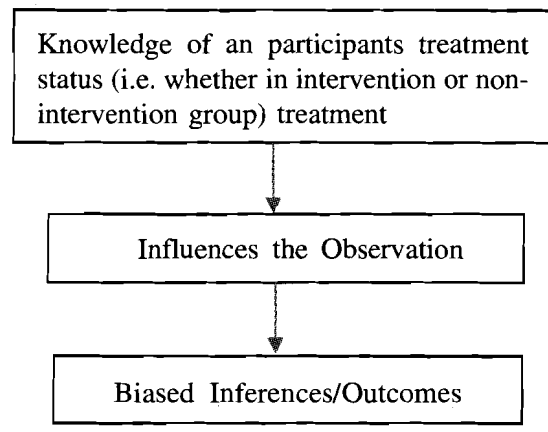

Figure 4.7: Potential for observation bias in the ascertainment of outcome in an intervention study

The second tool used in clinical trials to remove these sources of bias is the process called *blinding*. Researchers frequently use the terms "single," "double," and "triple" blind to communicate the blinding status of persons involved in the trials. What do these terminologies mean? Let us find out.

When either the subject or the observer is blind about the treatment received by particular subject, then this is called *single blinding*.

When both the subject and the observer are blind about the treatment received by particular subject, then it is referred to as *double blind*, as can be seen in Table 4.2.

Sometimes the data analyst becomes interested in certain findings and can gear the analysis and interpretations accordingly. To avoid this, the analyst is also kept blind about the subjects/codes. The codes are broken only after the data analysis is complete. This type of trial is thus called the *triple blind trial* (refer to Table 4.2)

Table 4.2: Various types of blinding used in clinical trials

	Types of Blinding		
	Single	**Double**	**Triple**
Subject	+	+	+
Observer		+	+
Data Analyst			+

(+) Blind with respect to subject's allocation

( ) May be aware of subject's allocation

Single blind trials you would realize are simpler to execute than double blind studies. Clinical trials or interventional studies are certainly more difficult to design and conduct than observational epidemiological studies, due to their unique problems of feasibility, cost and the most important consideration related to the *ethics of conducting a trial* on the human subjects some of which could be sick.

The important ethical considerations to be considered in intervention studies/clinical trials include:

- have the subjects been informed about the potential benefits and possible side-effects and their consent obtained?
- is it ethical to use a placebo on some subjects who are suffering from a disease or are sick, and
- is the treatment regimen under test reasonably safe?

Having gone through the discussion above, it must be evident to you that intervention studies are precarious. Trials that are sufficiently large, randomized and carefully designed, conducted and analyzed can provide the strongest and most direct epidemiologic evidence on which to make a judgment about the existence of a cause-effect relationship.

We will end our discussion on this topic by summarizing the strengths and limitations of intervention/experiment studies

Strengths

- Most like an experiment
- Provides strongest evidence for causality in relation to temporality and control for unknown "confounders"
- Fulfills the basic assumption of statistical hypothesis tests

Limitations

- Expensive, time consuming, sometimes ethically questionable.
- Subjects are often a highly selected group (selected for willingness to comply with treatment regimen, level of health, etc.) and may not be representative of all people who might be put on the treatment (i.e., generalizability may suffer).

Now answer the questions given in check your progress exercise 2 and assess your understanding about the experimental studies.

Check Your Progress Exercise 2

1) Explain the following briefly:

   a) Clinical trial: .....

   .....

   b) Community trial: .....

   .....

   c) Double blind trial: .....

   .....

2) Differentiate between therapeutic and preventive trials, giving examples.

   .....

   .....

   .....

   .....

3) Discuss the importance of intervention studies in analytic epidemiological research.

   .....

   .....

   .....

   .....

4) List the issues we would keep in mind in the design and conduct of intervention studies.

   .....

   .....

   .....

   .....

4.5 LET US SUM UP

This unit focused on analytic epidemiological research designs. We learnt that analytic studies focus on the determinants of a disease by testing the hypothesis formulated from descriptive studies, with the ultimate goal of judging whether a particular exposure causes or prevents diseases. There are various types of analytic study design options

namely observational and experimental/intervention studies, which we as a researcher can adopt.

The term observational study is used for an analytic study which investigates the natural course of event/disease. The researcher in this study design simply observes the natural course of event(s) noting who is exposed/non-exposed and who has/has not developed the outcome of interest. Types of observational studies include the cohort studies, case-control studies and the cross-sectional studies. A brief review of each of these analytic observational studies was presented in the unit highlighting their strengths and limitations.

Opposed to observational studies are *interventions* or *experimental studies*, which was the second focus of this unit. Intervention studies we read necessarily entail an intervention, wherein the investigator(s) themselves allocate the exposure (in contrast to being simply an observer as in observation studies) and then follow-up the cases for the subsequent development of the disease/event. A cause-effect type of relationship can be thus easily inferred. When the subjects for experiments in the field, lab or clinic are human beings, such experiments are generally termed as '*trials*', which can be a clinical trial or community trial or a therapeutic or preventive clinical trial.

There are a number of issues to be considered in the design and conduct of intervention studies which include selection of the study population, allocation of the treatment regimens, maintenance and assessment of compliance, and achieving high and uniform rates of ascertainment of outcome.

4.6 GLOSSARY

Placebo : an inactive substance or preparation used as a control in an experiment or test to determine the effectiveness of a medicinal drug/supplement etc.

Seropositive : showing a positive reaction to a test on blood serum for a disease.

4.7 ANSWERS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress Exercise 1

1) Analytic study is a comparative study (uses comparison groups) intended to identify and quantify associations, test hypotheses (in contrast to hypothesis formulation), and identifies causes (determinants). The various types of analytical studies which can be carried out include the observational studies and the experimental studies.

2) The objective of a cohort study is to observe a large group of people over a period of time to test the hypothesis regarding the causation of the disease. Give an example of a cohort study based on your review of journals/reference book/ internet.

3) The process of case-control study starts with the identification by researchers of an outcome or effect. The next important point is the identification of cases, containing individuals who exhibit the outcome under investigation. A number of control subjects are then chosen who do not exhibit the outcome or effect under investigation. The case and control groups are then compared on the proposed causal factors, and statistical analysis is used to estimate the strength of association between each factor and the studied outcome.

Refer to sub-section 4.3.2 and present the strengths and limitations of case-control studies based on your own understanding.

4) Descriptive cross-sectional study is thought of as providing a "snapshot" of the frequency and characteristics of a disease in a population at a particular point in time. Analytical cross-sectional studies, on the other hand, are those in which data on the prevalence of exposure and disease are obtained for the purpose of comparing disease differences between exposed and non-exposed. Comparison of differences is the analytical component of these studies.

Check Your Progress Exercise 2

1) a) When human beings are used for experiments such designs are called clinical trials, and the term "*clinical trial*" emphasizes the controlled aspect of the intervention, at the expense of the generalizability of the results.

   b) The term "*community trial*" emphasizes that the trial is carried out in a realistic setting (i.e. in the field), which can involve an individual-level intervention, a community-level intervention, or interventions with elements of both levels and results may therefore be more generalizable.

   c) A double blind trial is a trial when both the subject and the observer are blind about the treatment received by particular subject.

2) Therapeutic trials are conducted among patients with a particular disease to determine the ability of an agent or a procedure to diminish symptoms, prevent recurrence, or decrease risk of death from that disease. Preventive trials, on the other hand, involves the evaluation of whether an agent/substance or procedure reduces the risk of developing disease among those free from that condition at enrolment.

3) Intervention studies are most like an experiment. They provide strongest evidence for causality in relation to temporality and control for unknown "confounders". In fact, they can provide the strongest and most direct epidemiologic evidence on which to make a judgment about the existence of a cause-effect relationship. They also fulfill the basic assumption of statistical hypothesis tests.

4) Selection of the study population, allocation of the treatment regimens, maintenance and assessment of compliance, achieving high and uniform rates of ascertainment of outcome and ethical considerations are a few issues which we should keep in mind keep the design and conduct of intervention studies.