UNIT 3 DESIGN STRATEGIES IN RESEARCH: DESCRIPTIVE STUDIES

Structure

3.1 INTRODUCTION

This unit along with Unit 4 and 5 will provide a brief overview of the various design strategies used in epidemiological research. In this unit we will focus on descriptive studies used in epidemiology. A broad understanding about the epidemiological design strategies with major focus on descriptive studies their characteristics, limitations and strengths will be the main focus in this unit.

Objectives

After studying this unit, you will be able to:

- classify the design strategies in epidemiological research,
- define and discuss the basic characteristics of descriptive studies in epidemiological research,
- describe the different types of descriptive studies which can be conducted at the population and individual level, and
- enumerate the strengths and weaknesses of the descriptive studies.

3.2 DESIGN STRATEGIES IN EPIDEMIOLOGICAL RESEARCH

We already know that epidemiology deals with the distribution and determinants of disease frequency in human population. If you have an interest in research and have reviewed journal articles, research papers, you would have encountered various study designs used in epidemiological research. To illustrate, you may have read about the case studies or case-control studies or for that matter the follow-up/cohort and/or intervention/controlled trials etc. These are some of the designs employed in epidemiological research. It is important to note that the research strategies/design can be broadly categorized according to whether investigations focus on *describing the distribution of disease* or *elucidating its determinants*. Based on these two broad categories therefore design strategies in epidemiology are classified into two categories namely – *descriptive* and *analytic*. Refer to Figure 3.1 which classifies the epidemiologic design strategies. *Descriptive studies*, as the name implies, deals with describing the general characteristics of the disease distribution, particularly in relation to person, place or time. *Analytic studies*, on the other hand, focus on the determinants of a disease by testing the hypothesis (formulated from the descriptive studies), with the ultimate objective of assessing whether a particular exposure causes or prevents the disease. As you would have noticed each of these design strategies

include different type of studies. To illustrate the descriptive studies include the correlational studies, case studies, cross-sectional surveys etc. Analytic study design includes the observational studies, cohort studies, experimental/intervention studies. We will learn about these design strategies in this and the subsequent unit.

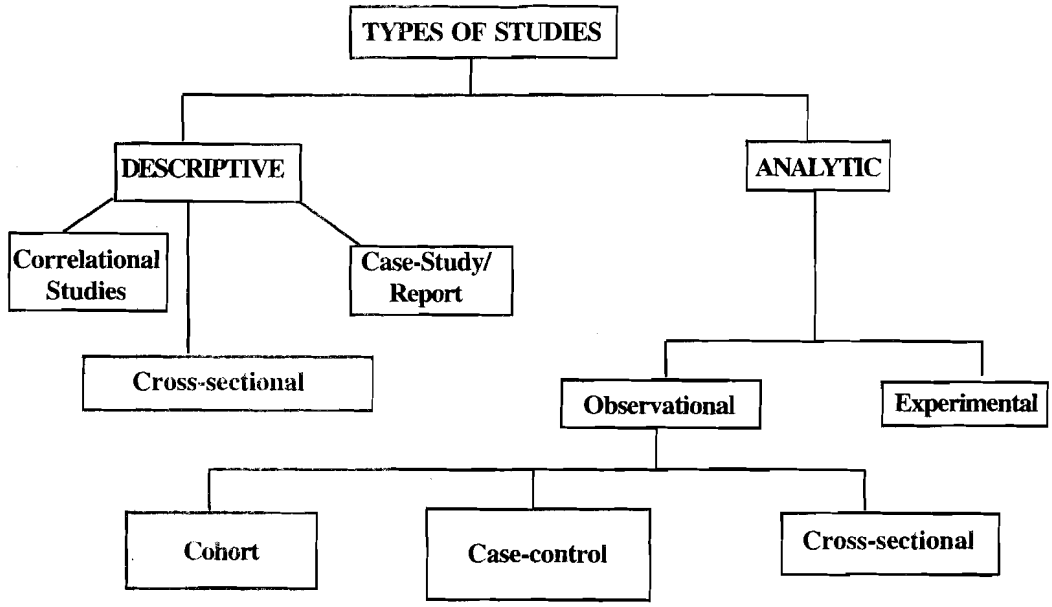

Figure 3.1: Overview of epidemiological design strategies

Whether the study is "descriptive" or "analytic", it is important to clearly identify the objectives of the study (preferably identifying the specific parameters to be measured) and the rationale (i.e., the case for conducting the research) as you will realize after going through this unit and the information presented in the next unit related to analytic studies. So let us begin our review with descriptive studies.

3.3 DESCRIPTIVE STUDIES

Descriptive studies are concerned with the distribution of disease, including consideration of what population or sub-group do or do not develop the disease, in what geographical location, and how the frequency of occurrence varies over time. Descriptive studies, in fact, seek to delineate the magnitude of the problem in different population groups, say in terms of prevalence and incidence, or to establish normal or abnormal levels of measurements. In simple terms descriptive studies are concerned with describing the distribution of disease or the health condition by *person*, *place* or *time*.

Basic indices of person examined in descriptive studies include the *demographic and life-style variables* such as:

- Age and sex distribution
- Socio-economic status
- Family structure, including marital status and number of single-parent families
- Racial, ethnic and religious composition
- Consumption of various foods, supplements, medicines etc.

Characteristics of *place* refer to geographical distribution of a disease, as already mentioned earlier, in terms of variation among countries or within the country, such as between urban or slum or rural areas. With regards to *time*, descriptive studies may examine seasonal patterns, growth patterns or compare frequency data of today with that of 5, 10 or 100 years ago.

Descriptive studies may use information from very diverse sources such as the census data, vital statistic record, data from national surveillance programmes, and employment health examination records, clinical records from hospitals and private clinics, as well as, national figure on consumption of food/nutrients, medication or other products. Since this information is often routinely collected and easily available, descriptive studies are generally far less expensive and time-consuming than analytic studies as you would get to know later in the next unit. In fact, a stimulus to investigation may come from a surveillance activity or descriptive study.

From our description so far, you would have realized that generally, in a descriptive study, the emphasis is *on estimation rather than testing*. Some of the quantities we might want to estimate are:

1) the prevalence of a disease,
2) the natural history of a disease.
3) the resources required to treat the disease.
4) attitudes and perceptions about the disease condition and so on.

Therefore, data so obtained from descriptive epidemiological studies are useful as they provide *valuable information to health/nutrition care providers, policy makers and administrators planning for health/nutrition care utilization and for allocating resources efficiently and also planning for effective preventive and therapeutic education programmes*. Further, the descriptive studies present the first important clues about the possible determinants of a disease condition. They are valuable to epidemiologists in describing the disease pattern i.e. reveal patterns associated with a specific disease without an emphasis on pre-specified hypotheses. Hence, these types of studies are sometimes referred to as *hypothesis generating studies* (to contrast them with hypothesis testing studies). The three important uses of descriptive studies therefore include *trend analysis*, *health/nutrition-care planning*, and *hypothesis generation* (and/or formulating research questions).

Though descriptive studies are quite useful, a frequent error in reports of descriptive studies is overstepping the data: studies without a comparison group allow no inferences to be drawn about associations, causal or otherwise.

There are a number of descriptive study design options, including *correlation studies of populations*, as well as, *case study/report, cross-sectional surveys among individuals* as you may have already noticed earlier in Figure 3.1 which illustrates the various types of descriptive studies. Each of these descriptive study designs provides information on various characteristics of person, place or time, and each has unique strengths and limitations. Let us get to know about these different descriptive study designs next. We begin our study with correlation studies.

3.3.1 Correlational Studies

Correlational studies as already mentioned earlier uses data from *entire population(s)* to compare disease frequency between different groups/segments during the same period of time or in the same population at different point of time. Correlational studies explore the statistical connection between disease in different population groups and estimated exposures in groups rather than individuals. For example, they may correlate anaemia rates by country with estimates of low birth weight, or low intake of iron rich foods or parasitic infestation such as hookworm infestation etc.

Correlational studies are used to look for relationships between variables. The correlation coefficient, denoted by r, is the descriptive measure of association in correlation studies. This coefficient (value can vary between +1 and –1) quantifies the extent to which there is a linear relationship between exposure and disease. We will learn about correlation as a data analysis method later in Unit 13 in this course.

Interestingly, while correlational studies can suggest that there is a relationship between variables, finding a correlation does not prove that one variable causes a change in another variable. In other words, correlation does not equal causation. For example, a correlational study might suggest that there is a relationship between dietary components and coronary heart disease, but it cannot prove that a change in one variable *causes* a change in the second variable. Other factors might play a role, including life-style habits, stress and a myriad of other variables.

Correlational studies are therefore useful for the formulation of hypothesis. For example, correlation data can raise the hypothesis that consumption of n-3 fatty acids can decrease the risk of coronary heart disease. Testing this hypothesis would require the design and conduct of analytic studies among individuals, which would account for the effect of other risk factors. We will learn more about analytic studies later in this unit. Here continuing our discussion regarding the strengths of correlational studies, the chief strength of correlational studies is that they constitute the first step in investigating a possible exposure-disease relationship, which contributes to their frequent use in epidemiological research. Further they can be done quickly and inexpensively, often using the already available data.

The main limitation of correlation studies is the inability to link exposure with disease in particular individual. Since correlational studies refer to whole population rather than individuals, it is not possible to link an exposure to occurrence of disease in the same person. A second major limitation of correlational studies is the lack of ability to control for the effect of potential confounding factors. Finally, correlational data represent average exposure/consumption levels rather than actual individual levels/ values.

Next, let us learn about case reports and case studies.

3.3.2 Case Study/Report

Case study and/or report, is a form of qualitative descriptive research that is used to look at *individuals*, a small group of participants, or a group as a whole. It refers to the collection and presentation of detailed information about a particular participant, event, institution or a small group, frequently including the accounts of subjects themselves. While using case study, researchers do not focus on the discovery of a universal, generalized truth, nor do they typically look for cause-effect relationships; instead, emphasis is placed on exploration, observation and description.

We are already aware that in research, the scientist study variables Case studies typically examine the interplay of all variables in order to provide as complete an understanding of an event or situation as possible. This type of comprehensive understanding is arrived at through a process, which involves an in-depth description of the entity (person/situation/event) being studied/evaluated, the circumstances under which it is used, the characteristics of the people/process involved in it, and the nature of the community/area in which it is located.

By using case studies we, therefore, seek a holistic understanding of the event or situation in question using inductive logic-reasoning from specific to more general terms. Case studies are often referred to interchangeably with ethnography, field study, and participant observation. Unlike more statistically-based studies which search for quantifiable data, the goal of a case study is to offer new variables and questions for further research. Statistical-based studies (which are quantitative) are concerned with the distribution of a particular trait, or a small number of traits, in a population, whereas the case study (which is qualitative) is concerned with the whole variety of traits/variables to be found in a particular instance/situation. There are different types of data collected in case studies. These may include: documents, medical records, interviews, direct observation, and participant observation and artifacts.

Individual case studies/reports can be expanded to case series, which describe characteristics of a number of individuals, patients, instances etc. Case series can be therefore referred to as a collection of individual case reports/studies, which may occur within a fairly short period of time. This design has historical importance, as it was often used as an early means to identify the beginning or presence of an epidemic. This is best illustrated by the recognition and formulation of hypothesis concerning possible risk factors of acquired immunodeficiency syndrome (AIDS) by early epidemiologists based on five cases of pneumonia reported among young, previously healthy, homosexual men in the US in the early 1980's. As a result of these case series, the Centre for Disease Control (USA) immediately initiated a surveillance programme to quantify the magnitude of this problem (i.e. AIDS) and develop the diagnostic criteria for what appeared to be a new disease.

The biggest strength of using case study/case series in epidemiological research is that it deals with creativity, innovation, and context. It is useful in the *recognition of new diseases* and the *formulation of hypothesis* concerning the possible risk factors linked to the disease. While case studies and case series are very useful for hypothesis formulation, they cannot be used to test for the presence of a valid statistical association. The fundamental limitation of case studies is that they are difficult to generalize because of inherent subjectivity and because they are based on qualitative subjective data, generalizable only to a particular context. Although case series can be sufficiently large to permit quantification, the interpretability of such information is severely limited by the lack of an appropriate comparison group.

From case studies, we move on to the third type of descriptive study used in epidemiological research i.e. cross-sectional surveys which are discussed in the next sub-section.

3.3.3 Cross-Sectional Study/Survey

We would like to begin our study of cross-sectional studies by informing our learners that cross-sectional studies can be either descriptive or analytical. We will learn about the analytic cross-sectional study in the next unit, here we will focus on the descriptive cross-sectional studies.

Most of the descriptive studies, you would realize, are cross-sectional in nature and are generally called *surveys*. What do we mean by the term cross-sectional survey? Cross-sectional survey is a type of descriptive epidemiological study design *in individuals* that involve observation of some subset of a population of items all at the same time. Cross-sectional studies can therefore be thought of as providing a "snapshot" of the frequency and characteristics of a disease in a population at a particular point in time in other words a 'snapshot' of the health status of the population at a specified time.

Descriptive cross-sectional surveys, therefore as you may have realized by now are, in general, useful for assessing disease status at a point in time. They simply characterize the prevalence of disease in a specified population. Prevalence can be assessed at either a point in time (point prevalence) or over a defined period of time (period prevalence). We will learn about point and period prevalence later in Unit 11. The basic *strength* of cross-sectional studies lies in the fact that:

- they can study an entire population or a representative sample,
- they provide an estimate of prevalence of all factors measured, and
- have greater generalizability.

Note descriptive cross-sectional studies are generally useful for planning or administering preventive or health care services, surveillance programmes, and surveys and polls. Further, it is important to note that cross-sectional studies are useful only

at finding some kind of association rather than for testing the hypothesis. This aspect, as you may have realized by now, is true for all descriptive studies.

Analytic cross-sectional studies, in contrast compare the proportion of exposed persons who are diseased with the proportion of non-exposed persons who are diseased.

One examples of cross-sectional surveys in India is the AIDS Prevention Control Project (APAC), which has been implementing intensive HIV/AIDS prevention activities in the Indian state of Tamil Nadu since 1995. The project launched the first wave of Behavioral Surveillance Surveys (BSS) in 1996 and followed up with successive surveys (in 1997, 1998 and 1999) to observe trends in high-risk behaviour among selected population groups. In Tamil Nadu these subpopulations included persons at high risk, such as female sex workers (FSWs), "bridge" groups, such as truck drivers and their helpers (THs), who have significant sexual contact with both high- and low-risk groups, and moderate to low-risk groups with varying socio-demographic characteristics, such as factory workers and students.

The Indian National Family and Health Survey (NFHS) is yet another widely known example of cross-sectional demographic and health surveys that have been carried out in several developing countries with the primary purpose of collecting information on health, fertility and other family issues from ever married women of fertility age. The first wave (NFHS-I) was completed between April 1992 and August 1993 with a sample of ever married women of age between 13 and 49. The second wave (NFHS-II) was completed between November 1998 and December 1999, sampling ever married women of age 15-49. Each survey contained reports from approximately 90,000 women, sampled from all Indian states using a stratified and clustered survey design. We will learn about clustered, stratified and other sampling techniques later in Unit 6). The largest component of the surveys is an individual questionnaire administered to each ever married woman of fertility age in the sample. The questionnaire also includes information on health, contraception and fertility preferences, as well as, a complete birth history and very detailed information on the health status of younger children. Such data are of great value to public nutritionists and policy makers and administrators in assessing the health status and health care needs of the Indian population.

With a review of the cross-sectional surveys we end our study of descriptive epidemiological studies. We hope you would have got a good insight into descriptive studies their strengths and limitations. Surely you would be in a better position now to consider using descriptive studies as part of your research work. Next, we shall move on to study the analytic studies in the next unit. Before we move on to analytic studies let us recall what we have learnt so far by answering the check your progress exercise 1.

Check Your Progress Exercise 1

1) Illustrate through a flow chart the classification of design strategies in epidemiology.

.....

.....

.....

.....

.....

2) Differentiate between descriptive and analytic design strategies, giving examples.

.....

.....

.....

.....

.....

3) Give one example each of descriptive study you would use at the population level and individual level. Explain the study briefly.

.....

.....

.....

.....

.....

4) List the strengths and limitations of descriptive cross-sectional studies.

.....

.....

.....

.....

.....

3.4 LET US SUM UP

The unit introduced the various study designs employed in epidemiological research. It elaborated on the fact that research strategies/design can be broadly categorized according to whether investigations focus on describing the distribution of disease or elucidating its determinants. Based on these two broad categories therefore design strategies are classified into two categories namely – descriptive and analytic.

Descriptive studies were the main focus of study of this unit. Descriptive studies, as the name implies, deals with describing the general characteristics of the disease distribution, particularly in relation to person, place or time. They are valuable to epidemiologists in describing the disease pattern i.e. reveal patterns associated with a specific disease without an emphasis on pre-specified hypotheses. Hence these types of studies are sometimes referred to as hypothesis generating studies. The important uses of descriptive studies include trend analysis, health/nutrition-care planning, and hypothesis generation (and/or formulating research questions).

There are a number of descriptive study design options, including correlation studies of populations, as well as, case study/report, cross-sectional surveys among individuals.

Correlational studies explore the statistical connection between disease in different population groups and estimated exposures in groups rather than individuals. Correlational studies are therefore useful for the formulation of hypothesis. Further they can be done quickly and inexpensively, often using the already available data.

Case study and/or report, is a form of qualitative descriptive research that is used to look at individuals, a small group of participants, or a group as a whole. It refers to the collection and presentation of detailed information about a particular participant, event, institution or a small group, frequently including the accounts of subjects themselves. By using case studies we therefore seek a holistic understanding of the event or situation in question using inductive logic-reasoning from specific to more general terms. Individual case studies/reports can be expanded to case series, which describe characteristics of a number of individuals, patients, instances etc. The biggest strength of using case study/case series in epidemiological research is therefore that it deals with creativity, innovation, and context. It is useful in the recognition of new diseases and the formulation of hypothesis concerning the possible risk factors linked to the disease. Finally the unit expanded on the cross-sectional studies. We learnt that most of the descriptive studies are cross-sectional in nature and are generally called surveys. Descriptive cross-sectional analysis surveys, therefore as you may have realized by now are, in general, useful for assessing disease status at a point in time. They simply characterize the prevalence of disease in a specified population.

3.5 GLOSSARY

Co-founding factors/ Variables : refers to two variables that are cofounded when their effect on a response variable connot be distinguished from each other.

Prevalence : is a statistical concept referring to the number of cases of a disease that are present in a particular population at a given time.

3.6 ANSWERS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress Exercise 1

1) Refer to Figure 3.1 and prepare the classification flow chart on your own.

2) Descriptive studies, as the name implies, deals with describing the general characteristics of the disease distribution, particularly in relation to person, place or time. Analytic studies, on the other hand, focus on the determinants of a disease by testing the hypothesis (formulated from the descriptive studies), with the ultimate objective of assessing whether a particular exposure causes or prevents the disease.

   Give examples of the two types of studies on your own based on your understanding of the topic.

3) At the individual level, case study and/or report, which are a form of qualitative descriptive research, can be used to look at individuals, a small group of participants, or a group as a whole. At the population level, correlation studies can be used which uses data from entire population(s) to compare disease frequency between different groups/segments during the same period of time or in the same population at different point of time.

4) The basic strength of descriptive cross-sectional analysis lies in the fact that, in general, they can study an entire population or a representative sample, they provide an estimate of prevalence of all factors measured, and have greater generalizability. Descriptive cross-sectional studies are generally useful for planning or administering preventive or health care services, surveillance programmes, and surveys and polls. The main limitation of cross-sectional studies is that it is useful only at finding some kind of association rather than for testing the hypothesis.