EPIDEMIOLOGY
Principles and Practice

Susmita Bhattacharya



JAYPEE

Contents

1. **Introduction to Epidemiology** *1*
2. **Epidemiological Studies** *7*
3. **Measurement of Disease Frequency and Association** ... *52*
4. **Causal Association/Cause Effect Relationship** *64*
5. **Epidemiological Concept of Disease and Natural History of Disease** *70*
6. **Epidemiological Investigations** *88*
7. **Screening for Disease** *95*
8. **Mortality Statistics** *106*
9. **Morbidity Statistics** *112*
10. **How to Design an Epidemiological Study** *116*

Annexure *121*

Bibliography *123*

Index *125*

2

Epidemiological Studies

In previous chapter we have learned that epidemiology is study of distribution and determinants of disease or health related states or events. To know this we have to enquire and investigate. The sequence of enquiry in epidemiology is as follows:

- Who are affected? — **Person**
- How many are affected? — **Frequency**
- When were they affected? — **Time**
- Where did the affected persons lived — **Place**
- How did it happen? — **Cause**
- What can be done to prevent and control the event in present situation and in future? — **Application**

These are the six questions, answers to which are searched in epidemiological studies.

The basic categories of epidemiological studies are two:

1. Studies primarily focusing on description of diseases – *Descriptive epidemiology.*
2. Studies that look into the determinants or cause of the disease.
   A. Analytical epidemiology
   B. Experimental epidemiology.

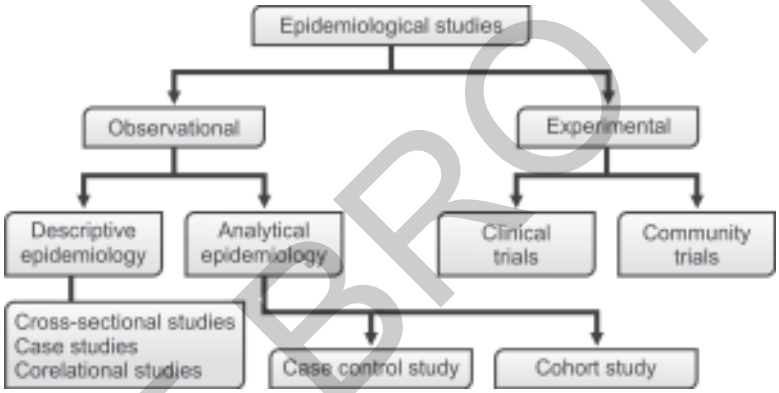
Of these three types of epidemiology, descriptive epidemiology and analytical epidemiology uses the observation as a tool. That means, the investigator observers the event and studies it in relation to time, place and person.

JAYPEE BROTHERS

In analytical epidemiology the events are only observed, the chain of events are not manipulated or altered by any drug, vaccine or procedure. The examples of analytical studies are:

1. Case control study
2. Cohort study.

In experimental epidemiology the investigator has controls over the events. Investigator manipulates it, for example clinical trials.



DESCRIPTIVE EPIDEMIOLOGY

In descriptive epidemiology the disease or health related event is studied in relation to place and time. These are not individual oriented studies. We describe patterns of disease or health related events occurring in community or group of persons. Let us understand how descriptive studies are useful for describing the disease in relation to person, place and time.

PERSONS

Descriptive data on persons address the question "Who are affected by the disease? While describing the disease we are interested in knowing, who are the persons affected. What are their ages and gender. Are they married or not? What is the

JAYPEE BROTHERS

level of education, occupation and income? These are some of the variables, which help in understanding the disease distribution. These characteristics are shown in Table 2.1.

Age and gender are two basic characteristics, which are routinely studied in any epidemiological study.

Age

We are well aware that certain diseases, such as measles, whooping cough and chickenpox are seen in children and are known as childhood diseases. Similarly, hypertension and coronary heart disease is rarely seen in young adults. Shift in the age group in a time period could be due to many factors. For example, because of immunization of children with vaccines like measles or mumps age shift has been observed in some countries. This is because of waning of immunity after certain years after immunization.

Similarly there is an increase in proportion of young men affected by coronary heart disease. In this case the explanation is not simple. It could be due to stress, obesity, high fat diet, and lack of exercise or variability of technology for diagnosis. The observed changes are not sufficient to put an explanation to the possible reason or cause. For this we need further studies of analytical nature, which looks into cause or determinants of the disease. Age is also important variable in studies on mortality. Maternal age is an important factor in preterm birth and low birth weight.

Gender

There are observed differences in males and females in relation to both mortality and morbidity. There are many factors, which contribute toward difference in incidence and prevalence of diseases and other health related conditions in men and women. It is documented that males have higher prevalence of heart disease. We know reason is biological. However, many a time reasons for gender related observed differences in prevalence of disease could be due to sociocultural or behavioral factors for example higher prevalence of anemia or malnutrition in

JAYPEE BROTHERS

females are more due to poor nutrition of females which is largely related to lower status in many communities across the the world.

Table 2.1: Characteristics in relation to persons in descriptive studies

- *Age:* Age specific mortality and morbidity—childhood diseases.
- *Gender:* Gender specific diseases, e.g. coronary heart disease, anemia, sex linked genetic disorders.
- *Ethnicity:* Certain traits are seen in some ethnic groups—Sickle cell disease
- *Marital status:* Mortality higher in unmarried
- *Occupation:* Determines exposure to harmful factors, e.g. silicosis, asbestosis.

Place

Descriptive study answers the question where do the affected persons live? In epidemiological study the major issue is description of disease in relation to geographical area. The variation in geographical distribution of disease gives us clue to the etiology and natural history of disease.

Geographic data can be presented as spot map and this is routinely used in studies of epidemics and outbreaks. Another factor in relation to place is mobility of individuals or movement of population. Population movement has led to changes in distribution of disease all over world. In India movement of population occurs to great extent. People move to cities or urban area in search of job. Movement of population occurs during harvesting, during fairs and festivals. It may lead to clustering of cases in particular locality during particular period. The geographical variation in prevalence of cases could be due to factors related to lifestyle, immunity, access to health services, exposure to disease and spread of disease amongst the group.

John Snow in his classic study of cholera described the disease occurrence and possible source of infection by using spot map.

JAYPEE BROTHERS

DIAGRAM-SPOT MAP (FIG. 2.1)

The geographic variation in disease or health related conditions could be:

1. *International variation:* For example, leprosy, tuberculosis and malaria is endemic in India and developed countries do not have problems of these diseases.
2. *Intracountry variation:* Distribution of disease within the country is not uniform. Many diseases are located or endemic in some states and their level of endemicity varies.
3. *Variations within the state:* Even in state the disease prevalence may vary, for example malnutrition is more prevalent in tribal area. Accidents are more in cities and metros.

Time

Descriptive data in relation to time addresses the question "when does the disease occur? If we note that the prevalence of disease or health related condition increased or decreased over a particular time period, we compare it to previous reference period. This gives an important clue to the cause or etiology. For example decrease in fatal accidents after legislation on seat belt in cars. Another example is decrease in prevalence of tuberculosis after directly observed short course therapy.



Fig. 2.1: Spot map of a village showing cases of disease

JAYPEE BROTHERS

The time pattern of the disease is also related to seasonality of the disease. Seasonal fluctuations are well documented in diseases like diarrhea and acute respiratory disease. The cases of diarrhea are more during rainy season and reason is contamination of water.

In chronic diseases, study of trends are useful. Here disease is studied over years or decades.

The secular change occurs gradually. The secular trends of a disease could be due to one of the following factors:

1. Changes in age distribution of population – If life expectancy increases, old people live longer and there could be increase in diseases seen in old age.
2. Changes in survival pattern due to advancement of medical science, people survive, for example tuberculosis and cholera, killed large number of persons fifty years ago.
3. Changes in diagnostic techniques.
4. Changes in actual incidence, due to environmental or lifestyle. For example, increase in air pollution is likely to increase incidence of respiratory diseases and lung cancer. Increased consumption of dietary fat can lead to increase in coronary heart disease.

TYPE OF DESCRIPTIVE STUDIES

There are three main types of descriptive studies:

1. Co-relational studies
2. Case series
3. Cross-sectional studies.

Co-relational Studies

In co-relational studies the characteristics which are chosen for study are used to describe the disease in relation to some factor of interest. Here in these studies the entire population is used. For example, to describe the mortality of persons due to coronary heart diseases we can co-relate to a factor of interest – like annual sale of cigarettes in that population and this can contribute to formulate the hypothesis. This can be proved or disproved in subsequent analytical epidemiological study.

JAYPEE BROTHERS

Merits

1. Easy, quick and inexpensive.
2. Can be used as first step to investigate possible exposure – disease relationship.
3. It can be done by using available data from Government bodies or private health agencies.
4. It is used regularly by Government and policy makers.

Limitations

1. It is not possible to link exposure to the person who gets the disease.
2 . The confounding factors cannot be controlled or measured in these studies.
3. The co-relational studies usually tell about average exposure but not the individual exposure which may be very important in disease causation.

Case Reports and Case Series

In this type of study the individual cases or groups of patients are studied. In clinical medicine these cases gives an important clue to the possible hypothesis. The early epidemiological study of AIDS is a classic example. Between October 1980 and May 1981, 5 cases of *Pneumocystis carinii* pneumonia were reported in young homosexual men in Los Angeles. These were unusual cases, as the pneumonia due to *Pneumocystis carinii* occurred in elderly cancer patients who were immune- suppressed due to chemotherapy. Similarly, in early 1981 Kaposi's sarcoma was diagnosed in large number of homosexual young men. This was different, as the said malignancy was exclusively seen in older patients and both men and women were equally affected. These descriptive data lead to analytical studies.

Limitations

1. It is restricted to a personal experience and there is no planned observation.
2. There is no comparison group available.

JAYPEE BROTHERS

Cross-sectional Surveys

These are also known as prevalence surveys. These are descriptive studies where in a defined population, the exposure to the possible disease causing agent and disease is studied simultaneously. However, it is not restricted to disease but we can also study the relation between environmental factor and health related behavior.

In cross-sectional survey there can be a specific time window. In some situations we conduct survey over period of a year or few months. The cross-sectional surveys can be at a fixed point of time for example, at the time of entry in school or at the time of employment. The cross-sectional surveys give information about the frequency and characteristics of disease in population which is studied.

Merits of Cross-sectional Studies

1. It is easy and less time consuming.
2. The information obtained is useful for assessing health status and health needs of the community.
3. National health surveys, which are cross-sectional surveys, are used for planning health program and assessing the impact of ongoing programs.
4. Useful for assessing prevalence of disease or other health related out come in occupational health.
5. Useful for formulating the hypothesis.

Limitations

1. It is not useful to study incidence.
2. It does not give information about cause and effect relationships.
3. It cannot test the hypothesis because the risk factors or characteristic may change over a period of time.
4. Not useful for rare cases or exposures. However, the cross-sectional surveys can be used as an analytical study to test the hypothesis. This is possible in cases of characteristics are unalterable for example, association between blood group and certain diseases.

JAYPEE BROTHERS

HOW TO PLAN A DESCRIPTIVE STUDY

Cross-sectional studies or surveys are most popular form of descriptive study.

The steps are:

1. *Define the objectives of the study:* What are we looking for? What do we want to know is to be clearly defined.
2. *Define the population:* The next step is to decide where you are going to undertake the study.

   Which is the population group you are going to study? The study population could be from a defined geographical area like slum, municipal ward, or village. The study area could be an institution like school, college, factory, office, etc. or it could be special group like women attending family planning center, children from school or blood donors from blood bank. Selection will depend on disease condition you intend to study.

   Descriptive study is a study of population. In cross-sectional study the reference is the population group in which disease or condition is to be studied. The size of population to be selected will depend on the nature of disease. For rare diseases you will have to take large population and for common diseases you may work with smaller population.

   Define the disease or condition: A clear and precise definition of disease or condition is a must. It is necessary for comparison with other groups and also for validity of the study. It is advisable to use standard definitions like definitions by World Health Organization or Health Ministry or standard textbooks. If standard definition is not available the investigators should develop a working definition. In absence of clear definition the study becomes invalid.
4. Study the disease in relation to time place and person. The data is to be collected using correct techniques and methods. Data is analyzed and disease is described in-relation to person, place and time.

JAYPEE BROTHERS

5. Compare with known information from other scientific studies.
6. Put a possible hypothesis to explain the observed pattern of disease and its relationship with other variables.

ANALYTICAL EPIDEMIOLOGY

Case Control Studies

These are observational studies; the investigator does not interfere with the natural history of the disease. In case control studies comparisons are made between a group of persons that have the disease and a group that do not. The persons with disease are labeled as "cases" and persons without the disease are called "controls". The groups are then compared in relation to risk factors or exposure to a particular substance or presence or absence of a characteristic of interest (Fig. 2.2).

The most commonly cited example of case control study is that of association between cigarette smoking and lung cancer. In case control study we go backward from effect to cause to study the association. Persons with lung cancer are enrolled as cases. People without lung cancer are enrolled as control group. Then we enquire from effect to hypothetical cause that is smoking.

Cause ← Effect
(Smoking) (Lung cancer)

Hence case control study is essentially a retrospective study.

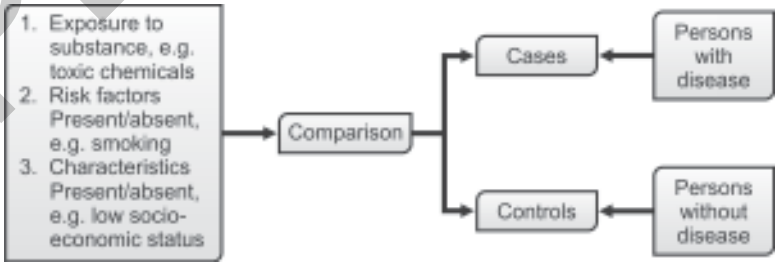


Fig. 2.2: Schematic presentation of case control study

JAYPEE BROTHERS

Basic Design of Case Control Study

In case control study our aim is to draw a conclusion that whether or not a valid statistical association is present between exposure and disease. For this following things are necessary:

1. Comparability of cases and controls with respect to all other factors except the factor under study.
2. Base line risk for developing the disease should be known, as all are not equally susceptible to develop a disease even with identical exposure level.
3. Accuracy of data collected and completeness of the information obtained from both groups.

THE STEPS IN CASE CONTROL STUDY ARE AS FOLLOWS

Selection of Cases

The prerequisite of selection is definition of cases. A precise definition is essential for accuracy and validity of case control study. If we choose a symptom or group of symptoms to define a "case", we should keep in mind that those may be due to different cause. For example, fever, cough, vomiting, etc. could be due to different diseases. Similarly, when we select a case based on disease of particular organ or anatomical site we should be precise. If we select patients of cancer uterus we should be precise, whether they will be patients of cancer corpus uteri or uterine cervix. If precise definition is not available we can classify the cases into definite case or probable case (Table 2.2).

Table 2.2: Sources of cases in case control study

• All "cases' diagnosed in a single hospital
• All "cases" diagnosed in one or more hospitals
• All "cases' diagnosed in all hospitals in defined geographical area
• All cases diagnosed in the community selected for study

Hospital-based Case Control Study

This is comparatively easy, but problem is that all cases may not visit the particular hospital or clinic; some may not even

JAYPEE BROTHERS

reach any health facility. However, hospital based studies has contributed valuable information, as they are feasible and practical.

Population-based Case Control Studies

In this type of studies, the cases are selected from a defined community. It eliminates the bias as mentioned in preceding paragraph. However, these are difficult, time consuming and hence less popular.

While selecting a patient as a case, it is to be remembered that "cases" could be "Incident" (new) cases or "Prevalent" (both old and new) cases. If prevalent cases are selected, they represent outcome of determinants of disease and duration of disease. Recently diagnosed cases are better for case control study because the data collection may be consistent. It is better if case control study is restricted to cases where complete and reliable information on exposure to risk factor and diseases related details are available. It is desirable that cases enrolled in case control study are representative of other cases of same disease in the community. This is fairly easy when cases are drawn from defined population. Thus, in a case control study of acute respiratory illness or childhood diseases in a nursery school or a day care center cases can be studied.

However, many a times this may not be possible. There are patients who do not go to hospital or they may not be aware about their illness at all. This happens in cases of diseases like hypertension or diabetes.

To find representative patients of disease under study as cases may be difficult. In this situation controls should be similar or comparable to "cases" with respect to age group, gender, socioeconomic status, etc.

Selection of Controls

The second step in case control study is selection of "controls". The question thus arises "why do we need control group? The control group provides us background for assessing exposure occurring in community or population. In case control study of

JAYPEE BROTHERS



Epidemiology
Principles and Practice

Salient Features

- Features epidemiology in a simple and readable way.
- Presented in a manner, to make understanding and application easy for both, Medical Students and Public Health Practitioners.
- Helpful for medical students appearing for various entrance examinations where knowledge of epidemiology is a key to success.
- Postgraduate students and researchers in the field of Medicine, Dentistry, Health Sciences, Public Health, etc. will find this book useful.

Available at all medical book stores
or buy directly from Jaypee Brothers through online shopping at www.jaypeebrothers.com

or call + 91-11-32558559

JAYPEE BROTHERS
Medical Publishers (P) Ltd.
www.jaypeebrothers.com

ISBN 978-81-8448-924-8