FOURTH EDITION

Mahajan & Gupta

TEXTBOOK OF PREVENTIVE AND SOCIAL MEDICINE

Revised by

RABINDRA NATH ROY
INDRANIL SAHA

JAYPEE

Contents

4 General Epidemiology

Diseases have afflicted mankind since days of yore. Alterations in growth, disturbances of metabolism, degenerative changes with advancing years, accidents, poisons, tumors, cancers, and invasions of body by microorganisms, all seem to have occurred with varying extent and distribution with the changing environment in which man has lived.[1]

Epidemiology has been recognized as "the multi-disciplinary study of the distribution (person, place, time) and determinants (cause) of health-related states or events in specified populations and the applications of this study to control of those health problems".[2-4] Epidemiology has evolved over a few centuries. It has borrowed from sociology, demography, statistics, as well as other fields of study and it is still considered as neonate or budding science.[5] It was not until the 19th century that the fabric of epidemiology was finally woven into a distinct discipline with its own philosophy, concepts and methods.[6] Epidemiological principles and knowledge of distribution of disease may be utilized to describe the natural history of disease as well causal factors.[7] Thus, it is useful to know how the duration of a disease and the probability of the various possible outcomes (recovery, complication, death) vary by age, gender, and so on. Such knowledge is useful not only for prognostic purposes but also in advancing hypotheses as to what specific factors may be more directly involved in determining the course of disease in an individual.[8]

Epidemiology is the basic science of public health that deals with health and disease in population. It has been defined various way by different epidemiologist.

Definition

'Study of the distribution and determinants of health related states or event in a specified population and application of this study to the control of health problems' (Jhon M Last, 1988). Last's in his definition emphasized that epidemiological study is not only concerned with the disease but also with 'health related events'. The term 'epi' means among and 'demos' means people; any study undertaken among population to find the magnitude of health problems and their distribution, causes with a aim to suggest remedial measures for those problems are called epidemiology. The word 'study' denotes scientific inquiry on some problem or event. The epidemiological investigation to health problem involves following *two basic approaches*.

1. ***Asking questions***: Availability of data is prerequisite for any systematic investigation on health problem in population; key information can be approached through a series of questions:
   - What is the health problem, condition, what are its manifestation and characteristics?
   - Who are affected, with reference to with age, sex, social class, etc.?
   - Where does the problem occur, in relation to geographical distribution, residence, place of exposure, etc.?
   - When does it happen in terms of day, months, seasons, etc.?
   - Why does it occur, in terms of the contributing or causative factors?
   - So what can be done? What intervention may have been implemented? Have there been any improvement following any action?
2. ***Making comparisons***: The next basic approach is to make comparison and draw inference. Such comparison may be made between different population at a given time, between subgroup of population, or between various periods of observation. By making comparisons, the investigator attempt to find out the difference related to study variables among study and comparison group, which help to draw inference on contributing factor or etiology of a disease. To ensure the 'comparability' between the groups (i.e. study group and control group), both the groups should be as similar as possible to all factors that may relate to the disease except to the variable under the investigation. In other word we can say that 'the like can be compared with like'.

Types of Epidemiological study

Epidemiological studies can be broadly classified as observational and experimental study with further subdivision, however, these studies cannot be regarded

as watertight compartment; they complement one another.

Observational study

- Descriptive—it includes case report, case series, correlation/ecological study, cross-sectional/prevalence studies.
- Analytical—can be of following type
  - Group based—the unit of study is population as group, e.g. ecological study
  - Individual based
    i. Cross-sectional
    ii. Retrospective—this can be case control study
    iii. Prospective—this is cohort study, also called follow-up study

Experimental Study (Interventional study)

Experimental study (Interventional study)—include clinical trial, field trial, etc.

OBSERVATIONAL STUDY

In this study nature is allowed to take its own course, investigator only measure do not intervene or manipulate any variable. This includes descriptive and analytical epidemiology. The descriptive epidemiology is concerned with measuring frequency and study of distribution of health related problem in population whereas analytical epidemiology attempts to analyze the cause or determinants of disease (how the disease caused?) by testing the hypothesis that has been setout in the study.

EXPERIMENTAL STUDY (INTERVENTIONAL STUDY)

Unlike the observational study, the researcher in an experimental epidemiological study, control or manipulate one or more factors in the study to obtain information how the factors influence the variables in the study and draw inference. This manipulation may be deliberate application or withdrawal of suspect causal factor or changing one variable in the experimental group while making no change in the control group and comparing the outcome in both groups. The experiments are designed to test the cause effect hypothesis.

EVALUATIVE STUDY

Evaluative study are those that appraise the value of health care; they are setout to measure the effectiveness of different health services. They are of two main types: review and trials.

Types of Evaluative study

- Program review
- Trials
  - Clinical trials
  - Program trial
  - Trial of screening and diagnostic tests

Study design

Structuring of research design can be divided into observational and experimental type. Observational types of studies generally employ the method of sample surveys, where a sample of the population is observed for various characteristics, whereas surveys where the observations on cause and effect differ by way of a period of time (such as case-control studies and cohort studies) are considered to be analytical in nature, and inference of associations can be made.

Study design may be classified into different type according to time of measurement of specific factor (cause) and its effects (disease).

Cross-sectional

A survey or study that examines people in a defined population at one point of time. Cross-sectional study may be descriptive, analytical, or both. Descriptive cross-sectional survey usually provides prevalence data but repeated survey can be used to give an estimate of incidence. In cross-sectional surveys the information on cause and effect is simultaneously gathered and the time sequence cannot be determined (e.g. study of relationship between body built and hypertension). Hypothesis may be generated from this type studies. This approach is useful during investigation of epidemic. This cannot distinguish whether exposure preceded the development of a disease or presence of a disease affect the individual's level of exposure. From epidemiological study it has been noted that individuals with cancer have significantly lower level of B-carotene in blood but from the study it is not possible to comment, which is the cause and effect. It is to be noted that in cross-sectional design observations are made *at one point of time only*.

Longitudinal study design: In contrast to previous designed described there is another type called longitudinal study design. In this design the observations or data refer to for more than one point of time. The difference between cross-sectional surveys and longitudinal studies can be expressed similar to difference between snapshot and motion picture.

Retrospective (Backward Looking Study)

Here the investigator start with effect and goes back to find the cause.

Prospective (Forward Looking)

In this study the investigator start with causative factor and goes forward to the effect. The term prospective not necessarily mean that the study is carried out in

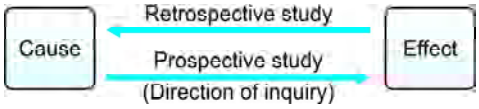

Fig. 4.1: Schematic presentation of prospective and retrospective study design

future, it can be carried out based on findings on record in past. Study population are divided into two groups, exposed to the factor of interest and not exposed to that factor, and then followed up to see and compare the development of disease in these two groups.

'Retrospective' and 'prospective' are distinguished by temporal relationship between initiation of study and the occurrence of disease outcome being studied. If the *outcome* and *exposure* both have already occurred at the initiation of investigation called retrospective (it can be case-control study or retrospective cohort study) and if the outcome of interest has not occurred at the initiation of investigation called prospective (also called cohort study) **(Fig. 4.1)**.

Choice of Study Design

A particular research question may be addressed using different epidemiological approach; the choice depend upon nature of the disease, type of exposure and availability of resources, as well as result from previous studies and gap in knowledge. The descriptive studies are primarily carried out for measuring frequency and describing pattern of disease or health related problem and for formulation of etiological hypothesis. On the other hand both case-control and cohort study can be used to test a hypothesis. For rare disease a case control study design is useful and for common diseases cohort study is suitable (large no of subjects available and need follow-up to get sufficient number of case).

Descriptive Epidemiology

The distinctive feature of this approach is that its primary concern is with description rather than with the testing of hypotheses or proving causality. This study is concerned with disease *distribution* and *frequency* in human population in relation to time, place and persons and identifies the characteristics with which the disease in the question is related. In this study the investigator tries to get the answer of questions about a disease or health related events. What is the problem and its frequency? Who are affected (person distribution)? When the disease occurs (time distribution)? Where (place distribution)? Descriptive studies are useful to formulate hypothesis.

Distribution is concerned with finding the frequency and pattern of disease or health related events in a population. Rate (number of events divided by size of the population) may be used to measure frequency, which allows valid comparisons across different populations. Pattern refers to the occurrence of health-related events by time, place, and personal characteristics. Sometimes we can study association between variables, which help in formulation of hypothesis.

MEASURING FREQUENCIES

The two main measures of frequency of disease, health problems and utilization of health services are incidence and prevalence. Incidence and prevalence may be expressed in absolute number or rate.

TIME TREND

These explain time distribution of occurrence of disease or health related events.

Secular Trend (Long-term)

Variation that occur over period of years, e.g. incidence of diphtheria showing decrease trend and diabetes, CHD, cancer showing rising trend since last few decade.

Periodic Trend (Cyclical Fluctuation)

Periodic fluctuation in occurrence of diseases is known as periodic trend, e.g. upsurge in influenza activity every 2 to 3 years result from antigenic drift of virus. Cause of periodic variation: (a) Variation in herd immunity, (b) Antigenic variation in agent.

Seasonal Trend

Annual variation in the disease incidence that is related in part to a season is called seasonal trend, e.g. community acquired infections and nosocomial infections show increased incidence in winter months because people inhale closed unfiltered air with droplet nuclei.

Acute (Epidemic) Trend

Short-term fluctuation is seen with epidemic outbreak. Epidemic is portrayed by epidemic curve, which is a graphical presentation of number of cases plotted against time.

PLACE DISTRIBUTION

World is not uniform in its characteristics, it varies in culture, standard of living, genetic makeup, etc.

Relative importance of these factors in etiology of a disease can be studied due to difference in place distribution, e.g. migration study can distinguish genetic and environmental factor in disease aetiology. To analyze by place, we usually organize data into a table, a map, or both. Variation may be classified under various levels.

International Variation

Stomach cancer is highest and breast malignancy is lowest in Japan, oropharyngeal cancer is high in India in comparison to other part of world.

National variation

Disease variation is also noted within the country.

Rural-Urban variation

Ch. bronchitis, mental illness, accidents, CHD are more common in urban area.

Local Variation

Geographical variation can best studied with aid of 'Spot map/ Shade map' which at a glance can show high or low frequency of a case. Clustering of cases may suggest common risk factors shared by all. Spot map used in 'John Snow cholera epidemic investigation' showed a common water pump in the Broad Street was source of infection thus helped to hypothesized that 'cholera is an water born disease'.

Person Distribution

Disease or a health related event is described by personal characteristics like demographic factors (e.g. age, race, sex, marital status), socioeconomic status, behaviors, environmental exposures, etc.

TYPES OF DESCRIPTIVE STUDY

- *Case series:* This kind of study is based on reports of a series of cases with no specifically allocated control group.
- Community diagnosis or needs assessment.
- Epidemiological description of disease occurrence.
- Descriptive cross-sectional studies or community surveys ('prevalence' study)
- *Ecological descriptive studies*: When the unit of observation is an aggregate (e.g. family, clan or school) or an ecological unit (a village, town or country) the study becomes an ecological descriptive study.

CASE REPORTS AND CASE SERIES

Case report is the descriptive study of the individual in terms of a careful, detailed report of a single patient. Case series means characteristics of a number of patients with a given disease. In other words, case series are the collection of individual case reports, which may occur within a fairly short period of time. These studies lead to formulation of new hypothesis.

Advantages

- New diseases are recognized, for example: Acquired Immunodeficiency Syndrome (AIDS), 5 cases of *Pneumocystis carinii* pneumonia
- Formulation of new hypothesis concerning possible risk factors.

Disadvantages

- Case reports are based on experience of only one person
- Cannot be used to test the presence of valid association
- Presence of risk factors may be purely coincidental and hence unreliable.

The planning phase of a descriptive cross-sectional study:

The following steps should be followed in conducting a descriptive epidemiological survey:

- Formulation of study objectives
- Planning of methods
  - Study population
  - Variables
  - Methods of data collection
- Methods of recording and processing data
- Comparing with known indices.

Objectives of a Descriptive Study

In formulating the objectives, the researcher expresses what he wishes the study to yield. The researcher may investigate the characteristics of population or obtain information on health status and health service (e.g. to find out incidence, prevalence, case fatality rate, distribution of some events, etc); and sometimes may find the association between variables, which help in formulation of hypothesis (e.g. the incidence of disease may be measured by time, place, person). Objectives should be expressed in specific and measurable term. The more specific the objectives the more easy it is to generate reliable and valid data.

The Study Population

Definition, sampling and sizing: This is the individual unit of study (persons, families, medical records, specimens, etc). It should be clearly and explicitly defined in terms of age, sex, occupation and other relevant criteria. The procedures for finding and inclusion of subjects (e.g. volunteers, hospital populations, people in the community) in the study should be clearly mentioned. The whole of the population in a geographical area or a representative part of it (sample) may be taken as study population.

Variables: Selection, Operational Definition and Measurement

The characteristics that are measured referred to as variables, which may be measured numerically (e.g. weight, height) or in terms of category (e.g. sex, presence or absence of a disease). Each of the variables used in the study should be clearly and explicitly defined. There are two kind of definition—conceptual

and operational. The conceptual definition defines the variables as we conceive it where as the *operational definition* ('working definition') define the characteristic as we actually measure it. An example of operational definition of obesity: A weight, in under clothes without shoes which exceeds by 10% or more of standard weight for age, sex and height in a specified population.

The disease under the study should be defined by a set of standard criteria called *'Case Definition'*. By using a standard case definition we ensure that every case is diagnosed in the same way, regardless of when or where it occurred, or who identified it. A case definition consists of clinical criteria and sometimes specified by limitations on time, place, and person. The clinical criteria usually include confirmatory laboratory tests, if available, or combinations of symptoms (subjective complaints), signs (objective physical findings), and other findings. For example, 'Clinical measles' may be defined as follows:

- Any person with fever and maculopapular rash (i.e. non- vesicular or without fluid), with cough or coryza (running nose) or conjunctivitis (red eyes).
- The variables under the study should be measured and *described by time, place and person* (described in earlier section).

Methods of collection, recording, processing and analysis of data should be planned before starting the study.

Comparing with Known Indices

By making comparisons, the investigator attempt to find out the difference related to study variables, which help to draw inference on contributing factor or etiology of a disease.

Determinants means to search for causes and other factors that influence the occurrence of health-related events.

Analytical (Explanatory Study): This study aims to explain a situation, i.e. to study the determinative processes of a disease or event. This tries to analyze the relationship between health status and other variable. Why does the disease occur in these people? Why certain people fail to make use of health services? Can decease incidence of a disease be attributed to preventive measures? Analytical study may be group based or individual based. In the group based study the researcher attempt to compare the data related to a variable in a group of population. For example, correlation study, a type of analytical study uses data from group of population as unit to compare the disease frequency among different group, e.g. per capita consumption of meat and rate of colonic cancer among population of different countries showed positive correlation. This type of study cannot test the hypothesis as they refer to group of population rather than to individuals. Individual based analytical studies though undertake survey of groups but they utilize information about each individual in the group. Two most important study designs under this category are be **Case-control study or Cohort study (Table 4.1).**

Case control study: In this study an investigator starts with diseased subjects and look back to study the exposure to the suspected factor. The diseased subjects taken for the study are called cases and another group without disease called comparison group are taken to compare the rate of exposure to the suspected factor in these two categories **(Fig. 4.2)**.

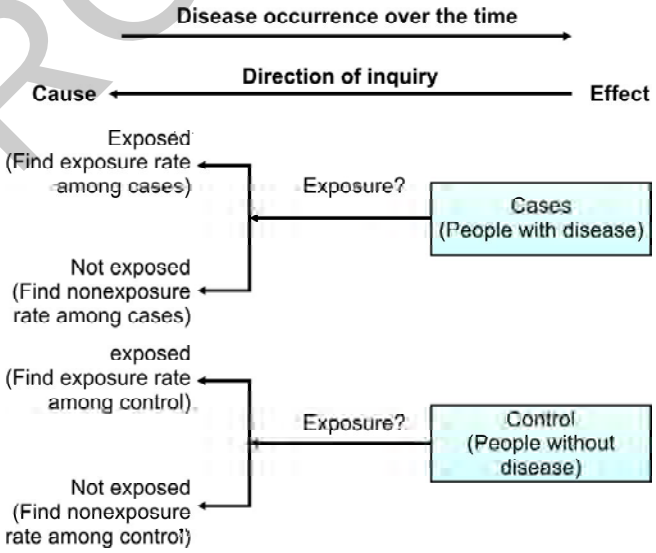

Fig. 4.2: Case control study design (Effect to cause study)

TABLE 4.1: Strength, weakness and main difference between case control cohort study

Case control	*Cohort*
• Proceed from effect to cause	• Proceed from cause to effect
• Start with diseased population	• Start with people exposed to the factor under study
• Case control provide information about one outcome only	• Useful for evaluating more than one outcome related to single exposure
• Allow to study the range of exposure	• Usually focus on one exposure only
• Suitable for study of a rare disease	• Impractical to consider cohort study for rare diseases
• For rare exposure study, case control may not suitable one	• Suitable for rare special exposure study
• Cannot estimate the incidence of a disease, so only can give estimate of relative risk (odd's ratio)	• Can provide accurate estimate of incidence of a disease—possible to find RR and attributable risk
• Time, cost, involvement is more	• Time, cost, involvement is more, more
• No problem of drop-out but record based information may be a problem	• Being a follow-up study there is more chance of drop-out

Features of a case control study:

- Both exposure and outcome have occurred before the beginning of the study.
- Proceeds from effect to cause.
- Use a control or comparison group to support or refute an inference.

Steps of case control study:

- Statement of the hypothesis
- Selection of cases and control
- Matching between cases and controls
- Measurement of exposure
- Analysis and interpretation.

STATEMENT OF THE HYPOTHESIS

This should be based on hypothesis which has been formulated from previous descriptive study or from previous experience.

Selection of Cases and Control

Defining the cases: The cases should be defined beforehand to avoid bias in the study. Diagnostic criteria and eligibility criteria should be established for cases.

Sources of cases: The cases can be taken from hospital or community.

Selection of control: The controls should be similar to cases as much as possible in respect of different variables except for presence or absence of the disease under study. Controls are not needed in the study in which hypotheses are not tested. Selection of controls depends on the nature of study. In a retrospective survey the association between a postulated cause and disease, the study group is compared with control group. Control should be selected from same source population from which the cases have been taken. Control should be representative of source with respect to exposure. Time during which a subject is eligible to become a control should be same in which an individual become a case.

Source of control: The controls can be taken from hospital, relatives, neighbours or general population.

Control from hospital or clinic: In a case control study, the controls can be selected from the patients with other disease (other than the disease under study) from same *hospital or clinic*. In this type of control selection, the study and control population are similar at least to some extent as they are from same parent population (catchments), and are subject to same selective factor. However, the selected control group being ill may not be representative of person without the disease under study.

Population control: If the cases are representative sample of a defined population and controls are sampled directly from that population, a random sampling of control may be a suitable one. If a population register exist or can be compiled this may be desirable method of control selection.

Size of control: Ideally case control ratio should be 1:1, but when there is doubt regarding the matching of all variables, several controls may be taken to increase comparability (e.g. 1:4). Any specific deficiency in matching can be compensated by inclusion of another group and thus will increase the power of test.

Neighborhood control: Where source population cannot be enumerated, instead of going through random sampling control match, the investigator may take control people who reside in the same neighborhood.

Matching between Cases and Controls

The controls should be similar to cases as much as possible in respect of different variables and matching can ensure this. Matching can be done various ways:

Individual matching: Each control may be so selected that he or she should be similar to the study subject in respect of different variables. This type one to one control can be taken from spouse, sibling, friends, neighbor, fellow worker, etc. This one to one close matching may not be possible if we wish to control more than two or three variables simultaneously.

Group or stratified matching: If control is taken as group and matching is done with study group for different variable like age, sex, occupation, etc. called group matching.

A combination of above may be used. Whatever method is used for selection of controls, clear cut rules should be laid down to ensure objectivity. For example, in individual matching, the degree of similarity must be expressed clearly for each characteristic.

Measurement of exposure: Measurement criteria must be defined clearly and same criteria should be used for measuring variables among the cases and controls.

Analysis and interpretation: A variety of statistical test are available some commonly used test are described below:

Frequency distribution of all variables: It is advisable to start the analysis by examining the frequency distribution of variables.

Summary of frequency distribution: Summary statistics of frequency distribution such as mean percentage, rate, of relevant variables can be calculated.

Association between variables: Analysis is done by finding and comparing the rates of exposure to a suspected factor among cases and controls. Simple methods of cross tabulation with a pair of variable may reveal association. Basic analytic framework for a case control study is 2 × 2 **(Table 4.2)**. For example if the intension is to test the hypothesis that 'cigarette smoking

causes lung cancer', the investigator begin with lung cancer cases (a+c) and matched control (b+d) to find out if there is any difference in exposure to a risk factor. This difference can be found out by statistical test of significance.

Calculation of Odds ratio: Since the typical case control donot provide incidence rate of disease in exposed and nonexposed group we donot get true measurement of relative risk from this study; instead we use the odds ratio as a measure of estimation of disease risk associated with exposures. The odds ratio is sometimes called the **cross-product ratio**, because the numerator is the product of cell *a* and cell *d*, while the denominator is the product of cell b and cell c **(Table 4.2)**. The odds ratio is calculated as:

$$Odds\ ratio = ad/bc$$

TABLE 4.2: A case control study of smoking and lung cancer with hypothetical data

Exposure	*Cases (lung cancer)*	*Control (without lung cancer)*
Smokers	a (95)	b (70)
Nonsmokers	c (5)	d (30)
Total	a+c (100)	b+d (100)

Exposure rate:

- Cases = a/(a+c) = 95/100=95.0%
- Control = b/(b+d) =70/100=70.0%

Test of significance for the difference noted in exposure rate: p value < 0.001

Estimation of risk by Odds ratio: ad/bc =2850/350=8.14.

COHORT STUDY (FOLLOW-UP STUDY)

It is an observational analytical study in which individuals are identified on the basis of presence or absence of exposure to a suspected risk factor for a disease and followed over time to determine the occurrence of subsequent outcome. This is also called 'cause to effect study' as the outcome of interest has not occurred at the initiation of investigation **(Fig. 4.3)**. It has the advantage of establishing the temporal relationship between exposure and health outcome, and thus they measure the risk directly. The Framingham study is a well-known cohort study which has followed over 5,000 residents of Framingham, Massachusetts, since the early 1950's to establish the rates and risk factors for heart disease.[9]

Features of a Cohort Study

- Exposure has started or yet to start but outcome has not yet occurred. However, in retrospective cohort study both have occurred, but investigation is designed to proceed from cause to effect.
- Proceeds from cause to effect.
- Use a nonexposed group to support or refute an inference.

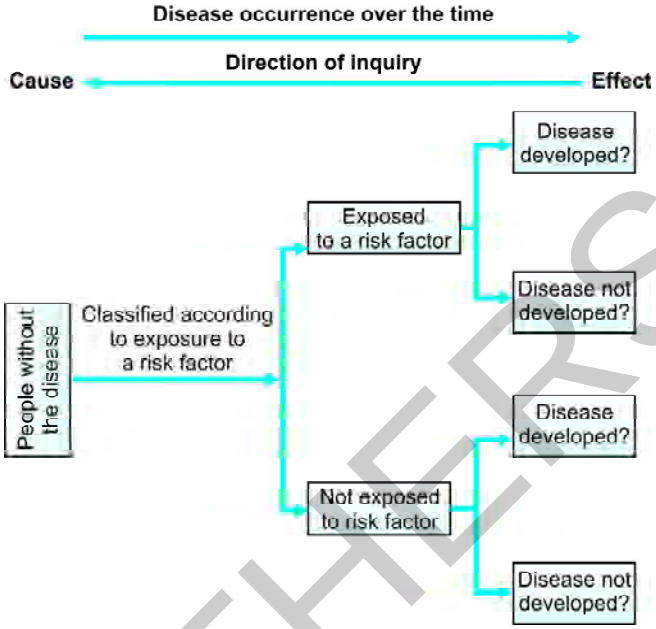

Fig. 4.3: Cohort study design (Cause to effect study)

Cohort: A well-defined group of people who share some common characteristic or experience called cohort. A group of people born during a particular year is called birth cohort, a cohort of smokers has the experience of smoking in common. There are two cohorts in cohort study, one of them is described as exposed cohort (exposed to the putative cause or condition) and other is unexposed or reference cohort (not exposed to the putative cause or condition). There may be more than two cohorts when exposure is classified according to level or type of exposure.

Indication: When exposures are uncommon but incidence of disease among the exposed group is comparatively high then cohort study may be suitable one (e.g. radiation exposure).

Type of Cohort Study

Depending upon the temporal relationship between the initiation of the study and the occurrence of the outcome (e.g. disease) the study can be classified under following heading.

- *Prospective cohort study:* The study subjects are classified on the basis of presence or absence of exposure and followed up to find the development of the outcome of interest. In this type, exposure may or may not have occurred but outcome must not have occurred at the beginning of the study.

- *Retrospective (historical) cohort study:* The subjects are also classified on the basis of presence or absence of exposure but in this type both the exposure and the outcome of interest have already occurred at the banging of the study. A historical cohort study depends upon the availability of good data or records that allow reconstruction of the exposure of cohorts to a suspected risk factor and follow-up of their outcome (e.g. mortality or morbidity) over time. The study can be carried out quickly and with limited resources.
- Combined cohort study having both retrospective and prospective design.
- Inserting case control with cohort study (nested case control).

Methodology (Steps) of Cohort Study

- Selection of study subject
- Selection of comparison group
- Obtaining information on exposure
- Follow-up
- Analysis and interpretation.

SELECTION OF STUDY SUBJECT

Initially the members of cohort must be free from the disease under study. The study subjects may be drawn from:

GENERAL POPULATION

The study subjects are chosen from general population (not a special exposure group). Subsequently they are divided into two groups described as exposed cohort and unexposed or reference cohort. Both the group should be representative of corresponding segment of general population. Relatively common exposure such as smoking, coffee drinking a large number of exposed subjects could be identified from general population. Famous 'Framingham heart study' selected study cohort from the resident of Massachusetts and followed them for 30 years.

SPECIAL GROUP

For rare exposure, such as related to a particular occupation, or environmental condition in a specific geographical location, it is more efficient to choose cohort from special group, e.g. doctors, nurses, occupational group, special exposure group, etc. It gives sufficient number of exposure population within reasonable time. For example, to study the relationship between industrial solvent and carcinoma, a retrospective cohort can be selected from particular occupation group.

Selection of Comparison Group

The comparison group (unexposed or reference cohort) should be as similar as possible to exposed cohort with respect to all factors that may relate to the disease except to the variable under the investigation. Comparison group is required to compare the difference in the rate of disease occurrence among two groups. There are various ways of selecting the comparison group. The controls can be taken from hospital, relatives, neighbors or general population.

Internal comparison: A single general cohort is entered in the study then its members are classified into different exposure groups on the basis of information obtained before the development of disease. The cohort study undertaken by Doll and Hill (1950) classified British physicians into smokers and nonsmokers group which acted as an internal comparison.

External comparison: In a cohort study of special exposure group it may not be possible to identify a portion of cohort that can be assumed to be nonexposed to the suspected risk factor for comparison. In such situation an external comparison group can be taken from general population or any other special exposure cohort, which is similar with study cohort. A cohort of radiologist can be compared with cohort of ophthalmologist to investigate the effect of radiation on development of malignancy.

Comparison with general population rate: Disease experience of study cohort can be compared with that of general population, e.g. lung cancer mortality of uranium mineworkers can be compared with that of general population.

Obtaining information on exposure: The goal is to obtain complete, comparable and unbiased information. Exposure information should be collected in such a manner that the study group can be classified according to degree of exposure. Information about the exposure may be obtained from number of sources.

- From cohort members by interview or mailed questionnaire.
- Review of available records.
- Medical examination or special test.
- Other sources.

Follow-up: At the beginning of the study, method should be developed to obtain data for assessing the outcome. The entire study participant should be followed up from point of exposure.

Analysis: The basic analysis of data from a cohort study involves the calculation of incidence rate of a specified outcome among both the group and estimation of risk. The rates can be compared among various groups with different degree or grade of exposure. The common measurement of analysis are following:

Measures (Refer Basic Measurement in Epidemiology in Chapter 3)

- Relative risk (true measurement of risk)
- Attributable risk (measurement of potential impact)
- Population attributable risk (measure of impact in population).

Relative risk (RR): It estimates the magnitude of association between exposure and disease. It indicates the likelihood of developing the disease in the exposed group relative to the unexposed group. It is a ratio of the incidence of the disease among the exposed persons to that in the unexposed persons:

RR = Incidence of disease among exposed persons/ Incidence of disease among nonexposed persons.

To calculate RR, a 2 by 2 table has to be made. The pattern of the table is exactly the same as that given in reference to the Odds Ratio. Exposure to the risk is shown in rows (horizontal) and presence or absence of disease is shown in columns (vertical).

Example: A cohort study is conducted to investigate the effect of smoking habits on lung cancer. It is found that among 200 smokers, 140 developed lung cancer while among 200 non-smokers 70 developed lung cancer. The following results were found:

Exposure Smoking Habit	*Lung Cancer Present*	*Lung Cancer Absent*	*Total*
Smokers (Exposed)	140 (a)	60 (b)	n1=200
Nonsmokers (Nonexposed)	70 (c)	130 (d)	n2=200
Total	210	190	400

From the table, it can be seen that the incidence of lung cancer over the total time period of the study among the exposed people is 140/200 = 0.70 or 70 percent. The incidence among the non-exposed persons is 70/200 or 0.35 or 35 percent. The relative risk is therefore 0.7/0.35 = 2.0

This is interpreted, as "people who smoked cigarettes were 2 times more likely to develop lung cancer than the nonsmokers".

Evaluative Study

In **evaluative study** some form of value judgements may be required. Attempt should be made to reduce the subjective element in judgment by using explicit criteria in assessment. Basic questions that are addressed in evaluative study provide the framework for setting study objectives. Following are the basic questions of an evaluative study:

- *Requisiteness (appropriateness) of care:* To what extent is the care needed?
  - Degree of need as judged by professionals
  - Need can be assessed from relative importance of problem, extent and severity of problem, perceived need (expressed by public), expressed demand (utilization of services)
- *Quality:* Quality of care need to judge on following aspect:
  - Structure evaluation (about facilities and settings)
  - Process evaluation (regarding performance of activities)
- *An appraisal of the performance of services indicate:*
  - What kind of services and how much?
  - Coverage of services, utilization of services, degree of compliance, community participation, etc.
  - Outcome evaluation (regarding the effect).

Appraisal of outcome requires clear-cut criteria of effectiveness. Effectiveness is the extent of achievement of pre-established target or goal attained as result of activities. If pre-established target or goal cannot be used as criteria, the investigator will need to formulate suitable criteria.

- *Efficiency (Economic efficiency):* Unit cost analysis, cost effective ration, cost benefit ratio
- *Satisfaction:* Client satisfaction and job satisfactions
  - Require attitudinal survey
  - Not necessarily means high quality services
  - Satisfaction ensure compliance.

USES OF EPIDEMIOLOGY

Refer Chapter 3 Page 22.

Some Example of Well-known Epidemiological Study

John Snow's classic study on cholera epidemic: John Snow is called the father of field epidemiology. Conducted his classic study on cholera in 1854 in the Golden Square of London. Snow believed that water was a source of infection for cholera (hypothesized). He began his investigation by marking the location of water pumps source for human consumption in the locality and then looked for a relationship between the distribution of cholera case households and the location of pumps. He noticed large number of cases in Broad Street area and then used this information to map the distribution of cases on what epidemiologists call a spot map; he observed the clustering of cases around a particular water sources (Broad Street pump). He also gathered information on water consumption in other area and noticed few number of cholera cases where the resident obtained water from alternate source. Consumption of water from the Broad Street pump was the one common factor among the cholera patients and concluded that the Broad Street pump was the most likely source of infection. Snow removed the handle of the Broad Street pump and aborted the outbreak.

Search for Cause and Risk Factors

Retinopathy of prematurity (ROP): In 1942, Terry first described the presence of grayish white opaque membrane behind the lens in premature babies known as 'retrolental fibroplasia' or (ROP), the cause of which

was not known at that time. Epidemiology of disease revealed peculiar clustering of ROP in a neonatal unit, where paradoxically the premature infants' survival rate had been improving. Based on preliminary observation, a well-controlled multicentric trial concluded that uncontrolled oxygen as toxic to premature retina. Following that liberal oxygen use was discontinued in premature infants.[10]

Diethyl stilbesterol (DES) and vaginal carcinoma: Over a period of four years a physician diagnosed clear cell Ca of vagina in seven young girls aged 15 to 22 years in a hospital of Boston in the year 1977. This disease had never been reported before in this age group. The apparent clustering of cases, led these worker to design a case control study and result proved that the use of DES during pregnancy was associated with occurrence of vaginal carcinoma in their offspring.[10]

Assessing the impact of legislative policy or law:[10] Epidemiological research can help to assess the influence of legislative policy or law on health of public at large. The effect of government law could be positive, in such situation epidemiological study can provide scientific basis to support the law. On other hand if the law is inflective or harmful, epidemiology can provide scientific basis to revert the policy or law in question. Example of some research that proved to be beneficial is as follows:

- Labor law to protect worker from occupational hazards
- Mandatory seat belt policy
- Antismoking policy.

Experimental Studies

An experimental study is a most definitive tool for evaluation of clinical research. It is a gold standard for evaluating effectiveness as well as side effects of therapeutic, preventive and other measures in clinical medicine as well as in public health.

An experimental study is defined as a study comparing the effect and value of intervention(s) against a control in a group of subjects. The basic difference between observational and experimental study is the intervention (manipulation).[11] It may be mentioned that Phase-I and Phase-II trials during development of a new drug are often conducted without a control group but Phase III trials are actual clinical trials having control groups.

Phase 1 Trial: These usually constitute the first step towards clinical experimentation and research into new or improved drugs, etc. Animal experiments are also part of Phase 1 trial.

Advantages: Cheap and less time consuming

Purposes: Study of:

- The adverse effect of drugs
- Benefits of the drug
- Absorption, excretion, metabolism of drugs.

Phase 2 Trials (Quasi-experimental design): These generally constitute the second step during drug research. They may use a quasi-experimental study design that may or may not have a control group. They are more expensive than Phase 1 trials but are still less time consuming and yield better results.

Limitations: Cannot control for observer/assessors bias and bias due to sampling variation.

Purpose: To study:

- Benefits of drugs
- True effect of drugs.

Phase 3 Trials (True experimental design): These are also referred to as Randomized Controlled Trials (RCT).

- There is a clear control group similar to the experimental group
- The terms of follow-up and all other conditions are kept similar for the two groups
- Blinding, preferably double-blinding, is observed to minimize bias
- The subjects are randomly allocated to the treatment or the control group as per predetermined randomisation procedure.

Example: A new analgesic "A" is to be tried for postoperative pain. Its efficacy is to be compared with a standard analgesic "B" already in use. It has been decided to try both of them in patients who have undergone a particular type of surgery with similar results:

- Patients are randomly assigned to two groups;
- Either analgesic A or B is given to subjects in a particular group, the drug being contained in similar color coded packs;
- Two groups are followed postoperatively for 3 to 4 days to look at the pain scores;
- After the results have been compiled, the code is broken to form comparative groups.
- The results are then compared to look at the true difference between the two groups receiving analgesic A or B.

TYPES OF EXPERIMENTAL STUDIES

These may be as follows:

- Preventive or prophylactic trials
- Therapeutic or clinical trials
- Community (field) trials.

Preventive or Prophylactic Trials

Here intervention takes place before the disease has occurred, e.g. study of vaccines or risk factors (stress, smokers, etc). Example: vaccinating one group against hepatitis B and leaving the other unvaccinated to study the efficacy of Hepatitis B vaccine. The most famous and one of the earliest vaccine trials was the one carried out by Louis Pasteur to a nine-year-old boy Joseph Meister on July 6th 1885.

Therapeutic or Clinical Trials

A clinical trial is an experiment with patients as subjects. Hence, the unit of study is a patient. The goal is to evaluate one or more new treatments for a disease or condition. The major ethical dilemma in such a trial is to decide about using placebo, which is a preparation containing no medicine or no medicine related to the complaint and administered to cause the patient to believe he/she is receiving treatment. It may sometimes be difficult to decide while planning a drug trial as to whether the control group should be given placebo or the standard medicine against which the drug in question is to be tested.[11]

Example

- Treatment of carcinoma breast comparing surgery, radiology and drug treatment.
- Studying a new drug for hypertension management and giving the drug to one group and placebo to another group.

Types of Therapeutic or Clinical Trials

Randomized Control Studies

These are comparative studies with an intervention group and control group. Subjects are assigned to intervention or control group as per proper predetermined procedure of randomization. Randomization is a process by which all subjects are equally likely to be assigned to either group.

Advantages

- Strong ability to prove causation
- Minimize or remove the potential bias in the allocation of subjects to intervention group or to the control group
- Randomization tends to produce comparable groups
- Randomization would guarantee the validity of statistical tests of significance.

Disadvantages: Ethical concerns like, depriving a subject from receiving a new therapy or intervention, which is believed to be beneficial regardless of validity of the evidence for that claim, i.e. randomized control trial deprives about one-half the subjects from receiving the new and presumed better intervention.

Nonrandomized Concurrent Control Studies

Here the subjects are assigned to either the intervention or the control group without randomization.

Advantages: Relatively easy to conduct by selecting the group of people to receive the intervention and selecting the control group by means of matching key characteristics.

Disadvantages: Intervention groups and control groups are not strictly comparable because of selection bias. This is a serious drawback and hence all efforts should be made to have random allocation system.

Historical Control Studies

Here a group of subjects on a new therapy or intervention is compared with a previous group of subject on standard or control therapy. In other words, a new intervention is used in a series of subjects and results are compared to the outcome in a previous series of comparable subjects. Such studies are, by their very nature, nonrandomized. The argument for using a historical control design is that all new subjects can receive the new intervention where it is felt that no subjects should be deprived of the possibility of receiving a new therapy or intervention.

Advantages

- Subjects may be more willing to participate in a study if they can be assured of receiving a particular therapy or intervention
- Less time consuming because all new subjects will be on new intervention and compared to a historical group
- Relatively cost-effective.

Disadvantages

- Potential for bias
- Results may be misleading because of:
  - Different structure and characteristics of the two groups
  - Shift in diagnostic techniques and criteria for the disease under study can cause major changes in the recorded frequency of the disease, thereby questioning the validity of the study
- Data for control group may not be accurate and complete.

Crossover Design

Here each subject participates in the study twice, once as a member of the intervention group and once as a member of the control group. It allows each subject to serve as his/her own control. In other words, each subject will receive, at different times, both treatments A or B. The order in which A or B are given to each subject is randomized.

Advantages: It allows assessment of whether a subject does better on A or B. Since each subject is used twice, once on A and once on B, the possibility of individual differences between subjects affecting the comparison of two groups are minimized. In other words, variability is reduced because the measured effect of the intervention is the difference in the individual subject's response to intervention and control. This reduction in variability means that the sample size needed is smaller.

Disadvantages: Effects of intervention during the first period may be carried over into second period. Hence, this design should be used only when there is ample evidence that the therapy has no carry-over effects.

Withdrawal Studies

Such study design is used when subjects on a particular treatment for chronic diseases are taken off therapy or have the dosage reduced. The objective is to assess response to discontinuation or reduction of the drug or its dose. The study is conducted using necessary randomization as safeguard against bias.

Advantages: This design may be validly used to assess the efficacy of an intervention that has never conclusively been shown to be beneficial.

Disadvantages

- The study sample is a highly selected one. Only those subjects are likely to have been on a drug for several years who, in the opinion of the physician, are benefiting from the intervention. Any one who has major adverse effects from drug would have been taken off and not been eligible for the withdrawal studies.
- Subjects and disease status may change over time so the results may be misleading.

Factorial Design

Here the attempt is to evaluate two interventions compared to the control in a single experiment. This design, with appropriate sample size, can be very informative when there is little chance of interaction.

Factorial Design

	Intervention A	*Control*
Intervention B	a	b
Control	c	d

a = A + B
b = B + Control
c = A + Control
d = Control + Control

Advantages

- With little increase in sample size two experiments can be conducted in one go
- Used for determining the interaction of two drugs.

Disadvantages

- There may be interaction between two groups, meaning thereby that the effect of intervention A may differ depending upon the presence or absence of interaction B, or vice versa. It is more likely to occur when the two drugs are expected to have related mechanism of action.

Group Allocation Design

Here a group of individuals, clinic(s) or community is randomized to a particular intervention or control. Such design is ideal when there is difficulty in approaching the individuals about the idea of randomization. Giving all subjects a specific intervention may be quite acceptable. In this design the basic sampling units are groups, not individual subjects.

Studies of Equivalency

In some instances, an effective intervention has already been established and is considered the standard. New interventions under consideration may be less expensive, have fewer side effects, or have less impact on an individual's general quality of life, and thus may be preferred. Studies of this type are called studies of equivalency or trials with positive controls. The objective is to test whether a new intervention is as good as an established one. The control or standard treatment must have been shown to be effective, that is, truly better from placebo or no therapy. It cannot be statistically shown that two therapies are identical, as an infinite sample size would be required. Hence, if intervention falls sufficiently close to the standard, as defined by reasonable boundaries, the two are claimed to be the same. The investigator must specify what he/she means by equivalence. It means specifying some value "d", such that two interventions with difference less than "d" might be considered equally effective or equivalent. Specification of "d" may be difficult but, without it, no study can be designed.

COMMUNITY (FIELD) TRIALS

They involve intervention on a community-wide basis. Here, the unit of study is a community. These are trials which are conducted on communities instead of individuals. Appropriate randomization should be used as far as possible, though this may sometimes be difficult due to practical considerations. They require greater number of subjects than clinical trials and, hence, are more expensive.

Examples

- Fluoridation trials for prevention of tooth decay.
- Deworming trials for ascariasis.[12,13]

SUMMARY OF THE METHODOLOGY OF INTERVENTION TRIALS

- Formulation of hypothesis.
- Decide the methodology for studying the affect of the independent variable on the dependent variable.
- Develop strategies for measuring the outcome and controlling the independent variable.

Salient Features

- All the chapters have been thoroughly revised and updated
- Various domains that are important, both in theory, practical and viva of MBBS examination, have been highlighted with examples
- Postgraduate study materials have also been incorporated with references for further reading
- Various flow charts have been introduced for clarity of understanding
- National Health Programs have also been included and revised
- Students will be benefited for their preparation in postgraduate entrance examination.

Rabindra Nath Roy has completed his MBBS and MD (PSM) from Calcutta University, Kolkata, West Bengal, India. He has been working in West Bengal Medical Education Services since last twelve years and now posted as an Associate Professor, Department of Community Medicine, Burdwan Medical College and Hospital, Burdwan,West Bengal, India. He has been teaching undergraduate, graduate and postgraduate students, and is Faculty in postgraduate entrance institutions. He has participated as state level trainer for different projects/programs in the State of West Bengal.

He has been associated with different professional bodies and actively participated in academic activities of these bodies. He has published numbers of scientific articles in national and international journals. He is officiating as an Assistant Editor, Indian Journal of Public Health (IJPH), as well as Reviewer of national and international journals. He has successfully completed different research projects and guided dissertation at MD/MS/MDS/DNB level.

Indranil Saha (Gold Medalist) MBBS MD has been working as an Assistant Professor, Community Medicine, Burdwan Medical College and Hospital, Burdwan, West Bengal, India. He has been teaching undergraduate, graduate and postgraduate students, and is Faculty in postgraduate entrance institutions. He has also done Professional Diploma in Clinical Research. He is also a state level trainer of different projects in West Bengal, India, like Integrated Management of Neonatal and Childhood Illnesses (IMNCI) and Maternal and Perinatal Death Inquiry and Response (MAPEDIR). He is a Life Member of Indian Public Health Association (MIPHA), Indian Medical Association (IMA) and Indian Association of Preventive and Social Medicine. He is one of the member of Editorial Board—Indian Journal of Public Health (IJPH), as well as Reviewer of national and international journals and is credited with many national and international research paper publications. He is one of the author of 'Biostatistics: MCQ and Essentials' for MBBS, postgraduate students and students aspiring for postgraduate entrance examination.

Shelving Recommendations
PREVENTIVE & SOCIAL MEDICINE, BIOSTATISTICS AND COMMUNITY MEDICINE

ISBN 978-93-5090-187-8
9 789350 901878