Education

Comprehensive Community Medicine for Undergraduate Students

Bratati Banerjee

As per the latest CBME curriculum (NMC 2024 guidelines)

Online at MedOne

Thieme

Contents

Key Highlights

NMC competencies covered in this chapter are as follows:

CM 1.1	Define and describe the concept of public health
CM 1.2	Define health; describe the concept of holistic health, including concept of spiritual health and relativeness and determinants of health
CM 1.4	Describe and discuss the natural history of disease
CM 1.7	Enumerate and describe health indicators
CM 7.4	Define, calculate, and interpret morbidity and mortality indicators based on a given set of data

Mapped with the latest NMC competency framework, ensuring alignment with curriculum requirements and outcome-based learning.

Content is systematically organized into well-structured subtopics for easy navigation and stepwise conceptual understanding.

Subtopics covered in this chapter are as follows:

- History of medicine
- Concept of health
- Concept of well-being
- Concept of disease
- Indicators of health, disease, well-being, and health service

Methods of Health Communication

Individual Methods of Health Communication

Counseling

A structured heading hierarchy ensures logical flow of content and facilitates quick referencing and retention.

Clear, well-labelled illustrations enhance conceptual clarity and aid visual learning of key concepts.

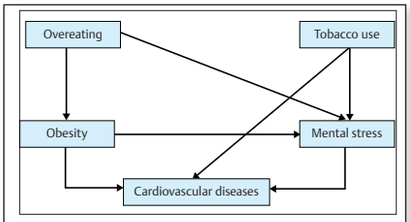

Fig. 1.5 Web of causation.

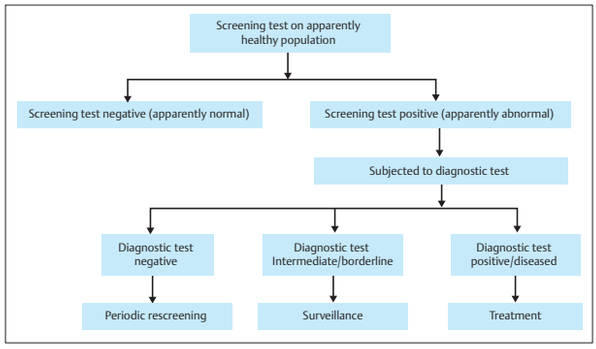

Fig. 6.1 Possible outcomes of screening and their management.

Simplified flowcharts provide quick revision of processes, pathways, and strategies.

Evaluation of Protein Quality

- *Digestibility coefficient (DC)*

$$DC = \frac{\text{Nitrogen absorbed}}{\text{Nitrogen intake}} \times 100 = \frac{\text{Nitrogen ingested} - \text{Nitrogen excreted}}{\text{Nitrogen intake}} \times 100$$

Important epidemiological formulas and indicators are presented in a clear, easy-to-recall format for quick application.

Well-structured tables present key data and comparisons in a concise format, enabling quick understanding and efficient revision.

Table 3.6 Classification of water according to level of hardness

Type of water	Level of hardness	
	mEq/L	mg/L
Soft	<1	<50
Moderately hard	1–3	51–150
Hard	3–6	150–300
Very hard	>6	>300

Check Your Progress

Scenario-Based Questions

1. A 2-year-old child with severe acute malnutrition was brought to the pediatric outpatient department of a district hospital. Prepare a plan of medical management for this patient on the following aspects.
   A. Initial assessment.
   B. Dietary management.
   C. Micronutrient supplementation.

Short Questions

1. What is the definition of health given by the World Health Organization?
2. What are the commonly considered factors that act as determinants of health?
3. Briefly outline the phases in the natural history of disease.
4. State the differences between rate, ratio, and proportion.
5. Briefly explain the current International Classification of Diseases.

Integrated review features including scenario based and short answer questions help reinforce concepts, promote analytical thinking, and support exam-oriented preparation.

All digital advances in public health including AI have been covered extensively

37 Digital Technology in Public Health

40 Important Definitions in Community Medicine

- **Opportunistic infection:** A condition in which a micro-organism takes the opportunity of the immune-deficient state of the host to cause a disease that would otherwise usually not have occurred in a normal individual.
- **Iatrogenic infection:** An infection transmitted during examination, investigation, or treatment procedures, for example, catheter-induced urinary tract infection.
- **Nosocomial infection:** An infection acquired by a patient from a health care facility that was neither present nor incubating when the patient came to the facility.

A separate section has been dedicated to cover all definitions of the book at one place for ease of students.

8 Epidemiological Research Studies

NMC competencies covered in this chapter are as follows:

CM 6.1 Formulate a research question for a study
CM 7.5 Enumerate, define, describe, and discuss epidemiological study designs
CM 7.10 Able to demonstrate development of research proposal
CM 7.11 Able to demonstrate the skills for critically appraising the research articles or research data

Subtopics covered in this chapter are as follows:

- Epidemiological research methods
- Quantitative research
- Descriptive epidemiology
- Analytical epidemiology
- Experimental epidemiology
- Quasi-experimental study
- Review studies
- Qualitative research

Research is a process of uncovering new information or providing new interpretation to existing facts. It is a systematic way of searching for the truth or reality. Research is generally considered to be of two broad levels as follows:

- ***Fundamental or pure research*:** It involves advancing the basic scientific knowledge, usually in a laboratory setting.
- ***Applied research*:** It studies the ways of practical application of the findings of pure or basic research.

Some other newly developed types of research are as follows:

- ***Translational research*:** It implies translation of fundamental research findings conducted in a laboratory to potential treatments for disease conditions.
- ***Implementation research*:** It identifies the implementation bottlenecks, with the purpose of improving health care and its delivery.
- ***Operations research*:** It is problem-oriented and deals with identifying the problem and working out mathematical solutions for making management decisions. This has been described in Chapter 31: Principles of Management.

Epidemiological Research Methods

Epidemiological research tries to answer the common questions encountered in the context of any disease or health and health-related problem. These questions are what, how many, how much, when, where, in whom, why, and how. In addition to these, there are some vital questions that need to be answered through research. These are how to control the problem, how to prevent its further occurrence in the future, what adversities can happen in implementing these measures, and how to minimize those adversities.

Epidemiological research is a kind of applied research. It deals with identifying the frequency of disease occurrence, the factors associated with diseases and other health-related problems, and the methods of preventing and controlling these problems. Research can be done on individuals or population, in a community setting or facility setting, which may be health care facility, educational institution, occupational organization or other facilities. Research can be conducted by collecting data by the researcher or by studying existing data from records, which

might have been routinely collected, and can be done as forward-looking study or by looking back on and analyzing what has already happened. These are the usual approaches to quantitative research. Research can also be conducted as qualitative research that helps in developing concepts for understanding social phenomena in natural settings, which takes into consideration the views, opinions, and experiences of the participants of research.

Background

More than 2,000 years ago, Hippocrates, the Greek physician also known as the "Father of Medicine," observed that diseases occurred due to the influence of environmental factors such as water, air, and diet. This marked the beginning of epidemiological research, thereby finding connection between exposure and outcome. In the nineteenth century, John Snow, an English epidemiologist, studied the occurrence of cholera in London and observed that it was related to consuming water supplied in the community. Snow highlighted the clustering of cases and their association with water supply from a particular company. This initiated the use of descriptive and analytical epidemiological research methods in understanding diseases in depth. James Lind conducted the first intervention study that established the role of citrus fruits in prevention of scurvy.

Epidemiological Research Methods

Epidemiological research may be conducted in various ways using different study designs, depending on the research question to be answered. However, all methods have been broadly grouped into two categories, as quantitative and qualitative study methods. The difference between these two categories is outlined in **Table 8.1**.

The various types of commonly undertaken epidemiological studies, with unit of research study, are as follows.

Quantitative Research

- Observational:
  - Descriptive:
    - Cross-sectional:
      - Prevalence survey—population.

Table 8.1 Differences between quantitative and qualitative research

Description	Quantitative research	Qualitative research
Goal	Obtain generalizable findings by describing and testing hypothesis	Generate theories and frameworks by understanding phenomena in depth
Based on	Numerical sciences	Behavioral and social sciences
Approach	Objective	Subjective
Type	Conclusive	Exploratory
Purpose	Identify and measure problem, find solution	Find reasons for the problem
Assessment	Result-oriented	Process-oriented
Research question	Specific and measurable	Complex and exploratory
Line of inquiry	Narrow close-ended questions—"what", "when," and "how many"	Broad open-ended questions— "why" and "how"
Sample size	Large, statistically calculated	Small, guided by data saturation
Sampling	Probability	Nonprobability
Data	Numerical and measurable	Verbal narrative, spoken, or written data
Data collected	Information on measurable variables	Experiences and feelings of participants
Assess	How much, how many, how frequently	The way people think, feel, behave
Method	Structured techniques, e.g., interview, observation, clinical investigations, etc.	Nonstructured techniques, e.g., in-depth interview, case study, focus group discussions, etc.
Tool	Structured, semi-structured, validated	Unstructured, topic guides
Data analysis	Univariate, bivariate, multivariate analysis	Thematic, content, and framework analysis
Data presented	Tables and graphs	Textual format, word clouds, tree maps
Analysis	Logical or statistical observations	Descriptive
Reasoning	Deductive	Inductive
Criteria for quality	Internal and external validity	Credibility, transferability, dependability, confirmability

    - Longitudinal:
        - Incidence survey—population.
        - Disease surveillance—population.
  - Analytical:
    - Cross-sectional:
        - Prevalence study—population.
        - Ecological study—population.
    - Longitudinal:
        - Incidence study—population.
        - Case–control study—individual.
        - Cohort study—individual.
        - Hybrid designs—individual.
        - Ecological time series analysis—population.
- Experimental:
  - Clinical:
    - Randomized controlled trial (RCT)—patients.
    - Cluster RCT—community.
  - Preventive:
    - Field trial—healthy individuals.
    - Community trial—community.
    - Quasi-experimental or nonrandomized trial—individuals or community.
    - Cluster RCT—community.

Qualitative Research

- Interview—individuals:
  - Key informant interview.
  - In-depth interview.
- Focus group discussion—group.
- Observation—facilities, individuals, groups:
  - Structure.
  - Procedure.
  - Behavior.
- Case study—individuals.
- Ethnography—community.

Quantitative Research

Quantitative epidemiological research for a new health or health-related problem passes through various stages for understanding its natural history, assessing its frequency of occurrence, identifying its determinants, working out measures for prevention or control of the problem, and assessing the efficacy, effectiveness, and efficiency of these measures. Research for all these purposes can be conducted by various methods, depending on the research question aimed to be answered. All research methods follow a pattern of hierarchy as depicted in **Fig. 8.1**.

The purpose of quantitative research is to identify the problem, measure the magnitude of the problem, identify risk factors associated with the problem, measure how strongly these risk factors are associated with the problem, and test efficacy of preventive and therapeutic measures designed to combat the identified problem. These studies deal with the subjects selected from a population and the findings of the research are projected to the population from which the study subjects have been selected. The various study designs available to carry out these functions are graphically represented in **Fig. 8.2**, which helps to understand the study design involved.

Ecological Study

Since ecological studies are included in both descriptive and analytical designs, which can be conducted as cross-sectional or longitudinal study, this is explained first. Here, instead of analysis of data from individual subjects, data of large groups of population, i.e., aggregate data, are collected and analyzed, which can be done in either way as follows:

- Large groups of population are selected, which are either followed up for a long duration covering the latent period of the disease under study to observe the incidence or assessed at one point of time to observe its prevalence.
- National decadal data regarding the disease are studied to observe any rise or fall in its magnitude.

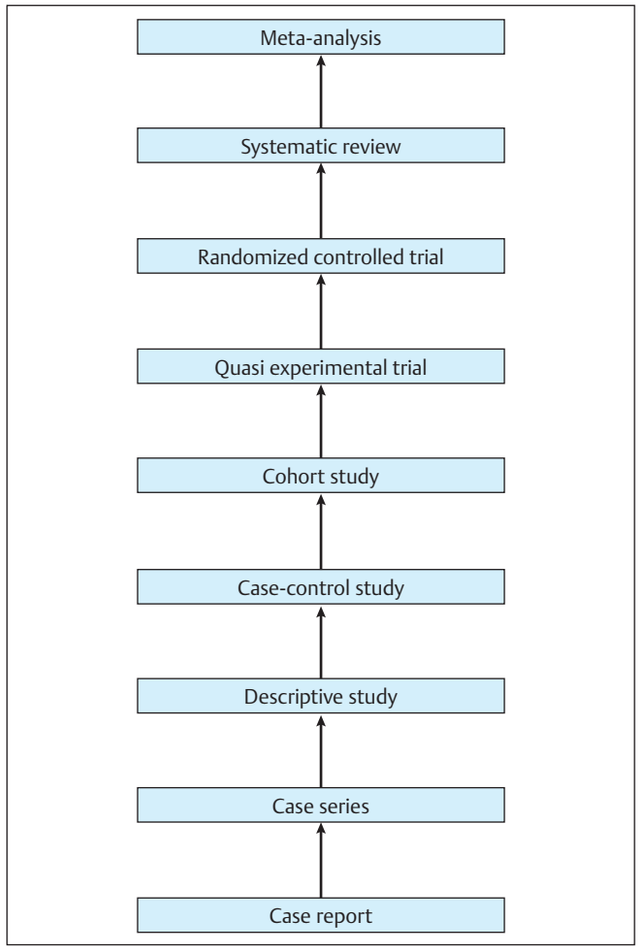

Fig. 8.1 Hierarchy of epidemiological research studies.

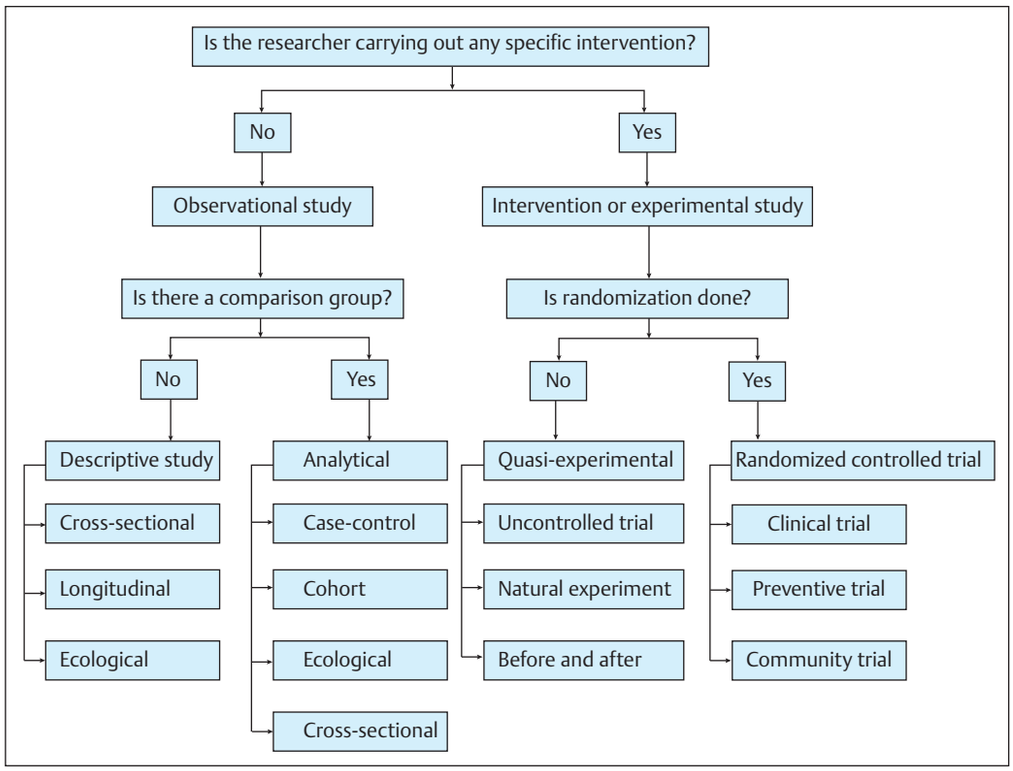

Fig. 8.2 Study designs for quantitative research

Ecological studies may be of three main types as follows:
- Cross-sectional ecological studies.
- Time-trend ecological studies.
- Solely descriptive ecological studies.

The first two designs are generally considered to be analytical, as these are comparative studies to see the differences in occurrence of health problems between two populations or two time periods. These are also known as correlational study where aggregate data are collected and analyzed to find association and/or correlation between exposure and outcome, e.g., association between tobacco use and cardiovascular diseases (CVDs), which can be done as cross-sectional or longitudinal study, in either way as follows:
- Two large groups of population selected with and without tobacco use, which are either followed up for a long duration covering the latent period of CVDs, to observe the incidence of CVDs, or assessed at one point of time to observe the prevalence of CVDs in both the groups.
- National decadal data regarding tobacco consumption and magnitude of CVDs can be compared to observe whether there is corresponding rise of CVDs with rise of tobacco consumption, or 1-year data of many countries can be observed to see whether countries with higher tobacco consumption have higher prevalence of CVDs.

The third type, as the name suggests, is of descriptive design as it aims to see the pattern of occurrence through description of the disease. This differs from the cross-sectional or longitudinal descriptive studies, as here data are not collected from individuals but from records of the population in general, e.g., to see decadal trend of noncommunicable diseases (NCDs) by observing the prevalence of NCDs or mortality due to NCDs, or occurrence of NCDs in different decades.

The main features of these designs have been depicted in **Table 8.2**.

This is a cost-effective and convenient design. However, there is a high risk of error as exposure and outcome may not be present together at the individual level, i.e., all smokers might not have CVD and those suffering from CVD might not be smokers. This is known as ecological fallacy or ecological bias, which is the most important problem in this design. This is because the association observed at the population level may not be present at the individual level.

Descriptive Epidemiology

The first step to research, following case report and case series, is a descriptive study. This describes the occurrence of disease or health-related problems in the population and attempts to identify the factors associated with the disease by answering the following research questions:
- When does the disease mostly occur?
- Where does the disease commonly occur?
- Who are mainly affected by the disease?

The answers to the above research questions typically describe the time–place–person distribution of the disease,

Table 8.2 Ecological study designs

Description	Cross-sectional	Time-trend	Solely descriptive
Study design	Across communities	Within the same community	Across communities or within the same community
Time frame	Same time period	Over time	At a point in time or over time
Example	Arsenic level in water and skin cancer	Level of air pollution in different seasons and occurrence of respiratory symptoms	Pattern of a disease in a large population

which help to identify the factors that are commonly associated with the disease, without establishing any cause–effect relationship. The aim of studies is to assess the magnitude of the disease and its various associated factors, thus helping to generate hypothesis.

Methods of Conducting Descriptive Study

There are two main ways of conducting a descriptive study: cross-sectional and longitudinal.

Cross-sectional Study

This is the simplest and easiest way of conducting descriptive research. This is a one-time contact of the study group at a particular point of time. However, the period of study may vary depending on the size of the population to be studied, and may even take months to complete. This should not create confusion with longitudinal studies, which take longer time due to multiple follow-ups. Diseases for which cross-sectional type of study is suitable are chronic diseases of long duration, e.g., hypertension, diabetes, tuberculosis. This is because most of the cases of acute conditions of short duration, for which recovery is quick, or highly fatal diseases for which death rate is high may not be encountered during single contact.

Longitudinal Study

This requires following up the study group over a period of time to identify new cases or observe the progress by multiple contacts with every subject in the group. Longitudinal study provides valuable information that may not be obtained through cross-sectional study. However, this study design is difficult to conduct and is costly. Diseases for which this type of study is applicable are acute diseases of short duration, e.g., diarrhea, acute respiratory infections. Disease surveillance is an ongoing process and hence is always longitudinal. The differences between cross-sectional and longitudinal studies are shown in **Table 8.3**.

Steps of Conducting Descriptive Study

Define and Describe the Population under Study

The population under study should be defined at the outset by the geographical boundary from where the population will be selected. It should be described by number and structure, in terms of relevant sociodemographic variables such as age, sex, and occupation. The population should not be grossly different from other population of the region or nearby areas. It should preferably be a stable one to reduce attrition or loss to follow-up, particularly in longitudinal studies of long duration. Hence, migrant population, visitors, tenants, etc., should not be included in the study, unless the research question pertains to specific criteria of such population, e.g., migrant studies.

The eligibility criteria should be clearly specified in terms of inclusion criteria, i.e., who can be included in the study, and exclusion criteria, i.e., who should be excluded out of those included in the study in the initial phase. The entire population or a representative sample can be taken for study. The population can be selected from the community, hospital, or any other organization such as workplace and educational institution.

The population under study should be large enough for the findings to be meaningful. The defined population or its sample considered for study forms the denominator for calculating rates. The results obtained for the sample is projected to the population from which the sample has been selected, provided the sample is representative of the population and large enough for results to be statistically valid. All of these are described in detail in Chapter 10: Biostatistics.

Define the Disease under Study

Having defined the population, the disease under study should be defined. For this purpose, an operational definition should be drawn up for accurately classifying diseased and nondiseased individuals and also for maintaining

Table 8.3 Differences between cross-sectional and longitudinal studies

Description	Cross-sectional study	Longitudinal study
Other name	Prevalence study	Incidence study
Contact with study subjects	One time	Follow-up
Time required	Less	More
Cost incurred	Less	More
Duration	Shorter	Longer
Number of observations	One	More than one
Result obtained	Similar to a snapshot	Similar to a movie
Measure of morbidity	Prevalence	Incidence
Type of diseases study is suitable for	Chronic diseases	Acute diseases
Useful for studying	Magnitude of problem	• Magnitude of problem • Natural history of disease • Identification of risk factors

uniformity in identifying cases by different data collectors or over different time periods. The operational definition should include history, symptoms and signs, examination, and investigation, which will be required to be fulfilled to designate an individual as diseased. Absence of the requisite criteria will classify an individual as nondiseased. The steps of examination and investigation should be standardized and communicated to all data collectors for uniformity in data. The definition and the steps outlined at the initiation of research should be strictly adhered to throughout the research period.

Collect and Analyze Data

Data are then collected, analyzed, and interpreted in the following ways:

- ***Describe the disease*:** The disease is described in terms of time, place, and person to outline the occurrence in various groups of population. This helps to identify which factors pertaining to these three characteristics are more commonly associated with the disease. However, as explained earlier, these factors do not establish any causal association or cause–effect relationship. These simply describe the disease occurrence in various population groups when classified under these three characteristics. The various factors describing time–place–person distribution have been explained in the next section.
- ***Measure the disease*:** The magnitude of the disease is next measured for morbidity, disability, mortality, and any other parameter that may be of interest to the research question. Measurement should be done in all relevant subgroups of the population, in terms of the various factors related to time, place, and person. Subgroup analysis helps in identifying the factors, the presence of which in a group is more likely to have higher burden of the disease.
- ***Compare with different groups*:** Comparison is made with different population, with subgroups of population, with groups of population that have increased risk of the disease under study, and also with data of known population such as national population.

Formulate Hypothesis

From the results obtained, hypothesis is formulated regarding association of the disease with particular time, place, or person. The hypothesis should not be vague, but should specify all relevant details, as follows:

- Population to which the hypothesis applies.
- Disease for which the hypothesis has been formulated.
- The associated factor that is a likely cause of the disease.
- Dose–response relationship, indicating the amount of the factor needed to cause the disease.
- Time–response relationship, indicating the duration of exposure to the factor needed to cause the disease.

This hypothesis is further tested by means of analytical study, through which all the above-mentioned details are studied.

For example, the hypothesis should be presented as *Lung cancer is present in 10% of smokers who have smoked 30 or more cigarettes per day for 20 or more years.*

Time–Place–Person Distribution

Examples of associated factors in terms of time, place, and person may be as follows, which have been explained further in detail:

- ***Time*:** Time of day, month, year, decade; seasons; time trend of the disease; periodicity, etc.

- ***Place*:** Rural–urban; towns and cities; slums and other areas; different climatic zones such as tropical and temperate; different terrains such as plains, hilly, riverine, desert, and forest; different geographical regions such as countries, states, and districts.
- ***Person*:** Demographic, socioeconomic, ethnic, cultural, behavioral, etc.

Time Distribution

Time distribution should be described according to time period that is relevant for the disease under consideration, e.g., for hours in case of food poisoning, days or weeks in outbreaks such as plague, months in case of seasonal diseases such as dengue fever, years in case of NCDs, and decades in re-emerging diseases. This helps to assess presence of any time fluctuation in occurrence of the disease or periodicity in increase or decrease of the disease, which helps in identifying factors that may be involved in disease causation or prevention. It also helps to understand whether the change was due to variation in etiological agent; change in diagnosis, reporting, management; or alteration in composition of population. It helps to provide guidelines for prevention and control of the disease condition or epidemic.

Commonly, three types of fluctuation of diseases with respect to time have been described. These are short-term, medium-term, and long-term fluctuations.

Short-term Fluctuation

This is most important in case of increase in occurrence of disease, which is also known as epidemic or outbreak. The increase is described in comparison with the expected frequency with which it usually occurs, which is also known as the endemic level of the disease among the specified population in the specified time period. Increase in disease occurrence to be labeled an epidemic is considered when increase in frequency is more than two standard errors of the mean frequency of a previous defined time period, which is usually considered as the past 3 to 5 years in case of rapidly occurring epidemics. The concept of standard error has been explained later in Chapter 10: Biostatistics.

The pattern of the epidemic can be depicted by drawing an epidemic curve, which is a graph of time distribution, showing time of occurrence of cases in the x-axis and number of cases in the y-axis, shown in **Fig. 8.3a–c**. It is a type of histogram that gives the following information:

- Magnitude of the epidemic over time.
- Type of epidemic.
- Pattern of spread.
- Course of epidemic, i.e., rising/stable/declining/stopped.
- Most likely time of exposure.
- Link the epidemic to a certain source.

Three important types of epidemics have been described, as follows.

Common-Source Epidemic

This originates from a common source, which may be single-exposure or multiple-exposure epidemics.

- ***Single-exposure epidemic*:** This is also known as point-source epidemic, as exposure to the disease agent for all affected population occurs from a common source and for a single time. Thus, exposure is brief and almost simultaneous and all the cases occur within a single incubation period. The epidemic occurs explosively with a narrow time interval. The epidemic curve rises and falls sharply, showing a single steep peak coinciding with the median incubation period and no secondary waves (**Fig. 8.3a**). These epidemics usually occur due to an infectious or toxic agent, such as food poisoning. Classical examples of common-source, single-exposure epidemics are the Bhopal gas tragedy in India due to leakage of toxic fumes from an industrial plant and Minamata disease in Japan due to consumption of fish contaminated with methyl mercury.
- ***Continuous or multiple-exposure epidemics*:** These occur due to repeated exposure of people to the disease agent from a common source. Here, there is no explosive sudden rise of cases, as exposure may be at different times. Hence, the cases occur beyond one incubation period. The epidemic curve is more flat, rises gradually often showing a plateau, and may show secondary waves. The epidemic finally declines due to infection of all susceptible individuals or removal of the source of infection (**Fig. 8.3b**). Repeated exposure can happen from contaminated sources such as a contaminated hand pump. Classical examples of common-source, multiple-exposure epidemics are Legionnaires' disease from contaminated air conditioner and Typhoid Mary, who was an asymptomatic healthy carrier of typhoid and a cook by profession, thus transmitting the infection to many others.

Propagated Epidemics

These epidemics thrive by transmission of disease agent from infected host to a susceptible host. Transmission may be human-to-human, animal-to-human, or vector-to-human. Transmission continues till the pool of susceptible host is exhausted or herd immunity is achieved. The epidemic curve shows multiple peaks, which are progressively taller, with each being one incubation period apart and leveling off gradually (**Fig. 8.3c**).

Slow Epidemic

This type of epidemic, as the name suggests, is slow in onset as well as slow in progression, thus taking a long period to develop fully. Examples are human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS), which took decades to establish as a serious public health problem, and rise of NCDs, which are also known as modern epidemic.

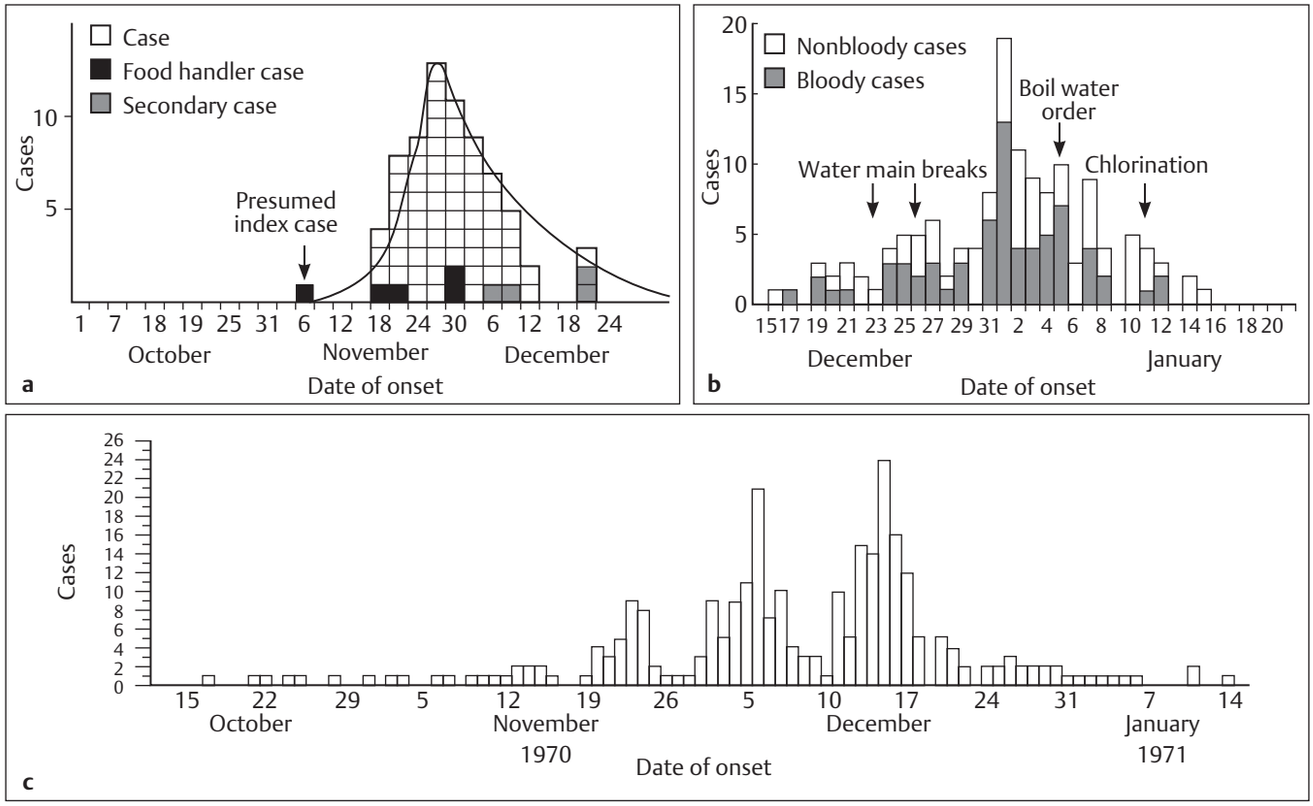

Fig. 8.3 Epidemic curve. **(a)** Hepatitis A Cases by Date of Onset, November–December, 1978 *Source: Centers for Disease Control and Prevention. Unpublished data; 1979.)* **(b)** Diarrheal Illness in City Residents by Date of Onset and Character of Stool, December 1989–January 1990 *(Source: Centers for Disease Control and Prevention. Unpublished data; 1990)* **(c)** Measles Cases by Date of Onset, October 15, 1970—January 16, 1971 (Source: Centers for Disease Control and Prevention. Unpublished data; 1979.)

Medium-term Fluctuation

This is also known as periodic fluctuation, as these occur at a gap of a time period that is determined by various factors. Two main types of periodic fluctuation have been described.

Seasonal Trend

This is due to influence of the climatic condition in different seasons, due to variation in temperature, rainfall, humidity, etc. This is commonly seen in most communicable diseases, e.g., childhood viral infections such as measles and chicken pox are most common in winter and early spring; mosquito borne diseases are more common in postmonsoon season; bacterial diarrheal diseases are more common in summers due to breeding of flies. Some NCDs such as chronic obstructive pulmonary disease show seasonal trend in exacerbation of symptoms, due to air pollution.

Cyclic Trend

This is increased occurrence of diseases over short periods of time that commonly occurs due to fall of herd immunity by rising susceptible population, which may be due to new birth cohort or immigrant population who do not have immunity to the disease. Another reason for cyclic trend that is commonly seen in influenza is mutation of the virus resulting in antigenic drift or antigenic shift, which has been discussed under influenza in Chapter 11: Respiratory Infections. Cyclic trend may also be seen in very short periods, e.g., road traffic accidents occurring more during end of the week due to increased alcohol consumption.

Long-term Fluctuation

This is also known as secular trend where the fluctuation is seen over a long period of time, which may continue over several years or decades. This pattern is commonly seen in increasing incidence of NCDs, where CVDs, diabetes, cancer, etc., have shown to be consistently rising due to alteration in lifestyle. There has also been decreasing pattern of certain diseases, e.g., vitamin A deficiency disorders, due to active implementation of prophylaxis program.

Place Distribution

Study of place or geographical distribution helps to perceive difference in disease occurrence between countries and within countries, thus identifying environmental

influences that may increase or decrease disease causation. The environmental influence may be due to physical environment, such as terrain and climate, or due to social environment such as culture and practices. Geographical variations are described as follows.

International Variation

This describes the distribution of diseases in all countries across the world. This gives an interesting picture of disease occurrence in different countries, which can subsequently be attributed to the genetic predisposition of the population, climatic conditions of the region, socioeconomic condition of the country, or cultural practices of the community—e.g., diabetes is higher in Southeast Asia due to genetic predisposition of the population, vector-borne diseases are higher in tropical regions due to climatic conditions, communicable diseases and malnutrition are higher in sub-Saharan Africa due to low socioeconomic condition of the countries, stomach cancer is higher in Japan due to cultural practice of consuming smoked food.

National Variation

Within the country also there are variations in disease occurrence, which can be attributed to factors that often vary across states, such as climatic and geographical condition of the state, cultural practices of the community, socioeconomic condition of the population, and availability, accessibility, and utilization of health services, etc.—e.g., iodine deficiency disorders are more common in hilly regions; kangri cancer is common in Kashmir due to cultural practice of using kangri for heating; communicable diseases are common in population of low socioeconomic classes. This can be depicted by the area map shown in **Fig. 8.4**.

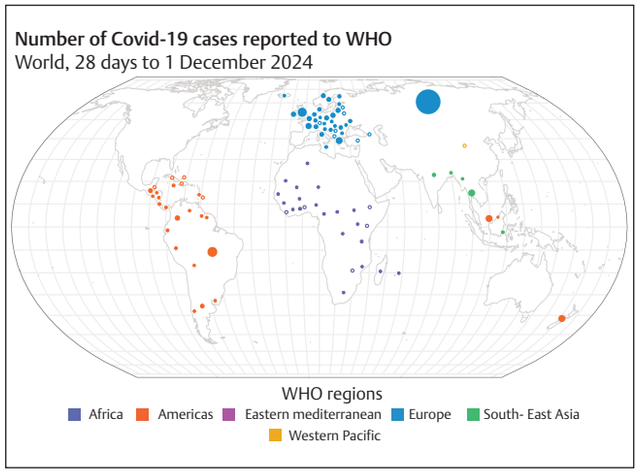

Fig. 8.4 Area map for COVID-19 in the world. (World Health Organization 2023 data.who.int, WHO Coronavirus (COVID-19))

Local Distribution

There are local variations also due to multiple factors. A classic example of local variation was identified and depicted by John Snow in his epidemiological investigation of cholera in London, where water supply was seen to be the cause of cholera in localized pockets of the city. Other than common sources of infection, other conditions also may show local variation, e.g., malnutrition is more in areas where the population belongs to low socioeconomic status; diseases spread by fecal–oral route are more common in slum areas where people live in insanitary conditions; communicable diseases are more in rural areas, while NCDs, particularly chronic lung diseases, are more in urban areas due to lifestyle and air pollution. Local distribution can be depicted by means of spot maps.

Person Distribution

The third component considered under descriptive epidemiology is the person, whose various characteristics are studied to understand the influence of these factors on the disease, so as to formulate a hypothesis. However, as with the other components, i.e., time and place, these characteristics do not establish any etiology or causal association. A descriptive study merely highlights the degree to which the factors studied are associated with the disease concerned, which help to identify high-risk groups. These may be demographic, sociocultural, socioeconomic, behavioral, etc. Common factors under consideration are described here.

Age

This is the first and foremost host factor to be described, as many diseases have been found to be occurring in particular age groups. Examples are measles in children, malaria in adults, and hypertension in elderly population. Some diseases occur equally among population of all age groups, e.g., acute respiratory infections, gastrointestinal problems. Some diseases show a bimodal distribution, i.e., two separate peaks in the age–incidence curve, e.g., breast cancer, in which incidence falls to some extent around menopause and shows a second peak in late postmenopausal period of about 65 years of age. However, bimodality may also be due to error in sampling, by either sample size being too small or sampling technique not being proper, thus resulting in a heterogeneous sample.

Sex

This is another important characteristic that is described at the outset. Many diseases show a distinct predilection for a certain gender, e.g., diabetes and obesity are more common in women, while lung cancer and coronary artery diseases are more common in men. Sex difference may occur due

to genetic or environmental factors. Genetic factor may be due to sex-linked genetic inheritance. Environmental and behavioral factors may be due to difference in lifestyle between males and females, e.g., malaria is more common in males as they tend to have more of outdoor activity and their bodies are less covered than women in general.

Ethnicity

Population subgroups of different racial origin also differ greatly from one another in case of certain diseases, which may be due to genetic or environmental factors. One example is that South Asian population, including Indians, have more predilection to NCDs, particularly diabetes. The reason for this has not yet been explained fully, although the amount of fat present in the body is considered to be responsible for it because they tend to have a higher proportion of fat than Caucasians with the same body mass index, which is attributed to the increased risk.

Cultural Factors

There are many traditional cultural practices that are followed by religious or social communities since time immemorial, which are handed down from generation to generation. Many of these are intimately related to health, in positive or negative way, which have been explained in Chapter 2: Social and Behavioural Sciences.

Social Class

Families are classified into different socioeconomic classes, based on certain criteria, which include education, occupation, income, material possessions, social status, etc. All these factors basically determine the standard of living and other determinants that are intimately related to health. There are several scales used for such classification, which have been described in detail in Chapter 2: Social and Behavioural Sciences. Health and disease have been established to be associated with the class that a person belongs to and thereby certain diseases have been seen to occur more frequently to a particular socioeconomic class. For example, communicable diseases are more frequent in lower classes, while NCDs are more common in upper classes. However, these are not watertight compartments and, hence, there may be variations. Also, social classification varies from country to country and hence its effects should be judged accordingly.

Education

Although education bears an indirect relationship, it is established to be a strong determinant of health, mainly through increased awareness that leads to following a healthier lifestyle, better health-seeking behavior, and increased treatment compliance. Female literacy has an even higher effect, as an educated woman takes better care of all family members and especially the children, thus improving the health status of the entire family. Kerala is a very good example of this, as the state has the highest level of female literacy among all states of the country, along with most of the health indicators being the highest.

Occupation

Occupation has an important role in the health status and disease in an individual. Hence, some diseases are more frequent in people engaged in certain occupations, while some diseases occur exclusively in certain occupational groups. This may be due to altered lifestyle in way of sleep, meals, substance abuse, physical posture, ergonomics, etc., or due to occupational exposure to certain harmful practices or substances. Examples include computer vision syndrome, which occurs more frequently in professionals in information technology, and pneumoconiosis, which are diseases specific to workers involved in the respective occupations, etc.

Marital Status

This has also been found to be associated with health and disease, as studies have shown lower morbidity and mortality in married people compared with unmarried or separated ones, when controlled for other sociodemographic factors. This may be due to better lifestyle, more responsible behavior, and being more protected and secure. Single, separated, divorced, and widowed individuals have higher risk of contracting HIV or sexually transmitted infection (STI) than currently married individuals.

Behavior

This is a very strong determinant of health and disease. The main behavioral factors that lead to causation of NCDs are unhealthy diet, lack of physical activity, use of tobacco, and consumption of alcohol. High-risk behavior in the form of sexual activity and injectable drug use is responsible for transmission of HIV and other STIs. Risk behavior and negligence lead to road traffic and other accidents.

Migration

Migration within and outside the country also has a very important role. Migration within the country may be from rural to urban areas in search of occupation, where the migrants are compelled to live in temporary shelters with low levels of housing and sanitation, with resultant ill health. Although migration by the affluent class is not affected in this way, there may be other effects in the form of adoption of the lifestyle and other practices of the place where migrated to. Migration also contributes to transmission of diseases across regions.

Many other host factors may be detailed to describe the person distribution, as relevant to specific health problems.

Uses of Descriptive Epidemiology

Descriptive epidemiology has the following main uses:

- ***Assessing magnitude of the problem*:** This is described in terms of incidence, prevalence, attack rate, death rate, etc., from longitudinal and cross-sectional descriptive epidemiological studies.
- ***Formulating a hypothesis*:** Association between a factor and disease helps in formulating hypothesis regarding etiology of the disease as derived from increased occurrence in different population groups described in terms of time, place, and person.
- ***Health services*:** The data obtained regarding magnitude of the problem are used in planning and evaluation of health services.
- ***Research*:** The hypothesis formulated is further tested through analytical and experimental epidemiological research to establish causal association, which helps in designing prevention and control strategies.

Examples of Descriptive Study

Some classical studies that have followed the descriptive design are as listed:

- ***Time distribution*:** Legionnaires' disease—This was identified in 1976, when an outbreak of severe pneumonia with sudden deaths occurred in Philadelphia. Within a week, more than 130 people were hospitalized, of which 25 had died. On investigation, all the patients were found to be participants of the American Legion Convention, which was held at a hotel. The disease was identified to have occurred by transmission of a new bacterium through the air conditioner of the hotel. This organism was named as *Legionella pneumophila*, after the legion.
- ***Place distribution*:** Cholera epidemic in London—In 1854, there was an epidemic of cholera in London. John Snow, an epidemiologist, investigated this epidemic and, using spot maps for observing the local distribution of occurrence of cases, identified a common water pump in Broad Street as the source of infection.
- ***Person distribution*:** Typhoid Mary—In the early 1900s, multiple cases of typhoid fever occurred in Ireland. On investigation, it was reported that all cases occurred in affluent families where a cook named Mary Mallon worked. On testing the blood, urine, and stool samples of Mary, she was found to be an asymptomatic intestinal carrier of typhoid.

Analytical Epidemiology

This is the next step in determining etiology of a disease through epidemiological research. In addition to establishing association, these studies also determine strength of the association and estimate the risk of acquiring the disease in presence of exposure to the causal factor. Like descriptive studies, analytical studies can also be of cross-sectional or longitudinal design, although the latter is more common. While descriptive cross-sectional study formulates hypothesis by finding association between diseases and other factors, analytical cross-sectional study, on the other hand, tests the hypothesis formulated through descriptive studies by testing the association of these factors with the disease in question. The basic difference between descriptive and analytical cross-sectional studies should be clearly understood. The difference lies in the research question with which the study begins. This is explained in **Table 8.4**.

Two main types of longitudinal analytical studies are conducted. These are case–control study and cohort study, which are described here in detail. The differences between case–control and cohort study have been detailed in **Table 8.5**.

Case–Control Study

These are also called retrospective or backward-looking studies, as at time of initiating the study the exposure and outcome have both already occurred and the investigator looks back in time to identify presence of factors that might pose a risk for occurrence of the disease under study. Hence, the study is named so, as it goes backward from effect to cause. This is the first step to testing a hypothesis formulated through descriptive studies.

These are comparison studies where comparison is done between two groups, individuals of one group having the disease under study that is known as the case group, which is compared with the other group that comprises individuals who do not have the disease and is known as the control group. The presence or absence of factors suggested in the hypothesis, and if present the degree to which these are present, is studied and reported through analysis as estimation of risk.

Case–control studies have several advantages over cohort study. However, as a study design, it is a weaker study than cohort study due to its many disadvantages. These are stated in **Table 8.5**.

Steps of Conducting a Case–Control Study

Step 1: Defining the Disease and Exposure

- ***Disease*:** The disease under study should be defined with an operational definition drawn up as explained for descriptive study.
- ***Exposure*:** Exposure criteria also need to be specifically defined before initiating the study, along with measuring the degree of exposure.

Measurement of disease and exposure should ideally be done in the same manner in both case and control groups.

Table 8.4 Difference between descriptive and analytical cross-sectional studies

Description	Descriptive cross-sectional study	Analytical cross-sectional study
Level in hierarchy	Lower	Higher
Aim of study	Generate hypothesis	Test hypothesis
Objective	Assess magnitude of the disease and its various associated factors	Test association of these factors with the disease under study
Research questions	• When does the disease mostly occur? • Where does the disease commonly occur? • Who are mainly affected by the disease?	• Does the disease occur significantly more in a particular defined time? • Does the disease occur significantly more in a particular defined place? • Does the disease occur significantly more in a particular defined group of population?
Hypothesis	Formulated	Tested
Association	Identified	Established
Causation	Not known	Not known

Table 8.5 Differences between case–control and cohort study

Description	Case–control study	Cohort study
Other name	Retrospective study	Prospective study
Level in hierarchy	Lower	Higher
Direction	Backward	Forward
Starts with	Disease	Exposure
Proceeds from	Effect to cause	Cause to effect
Exposure	Already occurred	Already occurred
Disease	Already occurred	Not yet occurred
Groups	Diseased and nondiseased	Exposed and nonexposed
Suitable for studying	Rare disease	Rare exposure
Factor studied	Any associated exposure found	Exposure selected for study
Outcome studied	Disease selected for study	Any associated disease found
Analysis	Exposure rate in diseased	Disease rate in exposed
Causal association	Cannot be established	Can be established
Incidence	Cannot be calculated	Can be calculated
Estimation of risk	Odds ratio	Relative risk and attributable risk
Time	Less	More
Cost	Less	More
Number of subjects required	Less	More
Number of contacts with subjects	One	Multiple
Recall bias	Present	Minimal
Ethical issues	Minimal	Present

Step 2: Source of Cases and Controls

Cases and controls may be selected from hospital or from the general population. Controls may be selected from the same hospital as cases, but suffering from another disease. It is convenient to select cases from hospital. However, these subjects may not be representative of the population in terms of certain factors, most important of which is treatment or health-seeking behavior, and hence the results obtained may not be valid if projected to the entire reference population. However, in case of a rare disease, subjects may be selected from hospital as it might be difficult to identify cases in the community. Cases in the population may be identified from disease registry or hospital records. Controls may be relatives of the cases, e.g., sibling or spouse, from the same neighborhood, i.e., locality, educational institution, workplace, or from the general population.

Step 3: Selection of Cases and Controls

- ***Diagnostic criteria*:** The disease should first be defined by outlining a standard operational definition of the disease, along with the diagnostic criteria, the presence or absence of which will be used to identify individuals to be included in case or control group, respectively. Ideally, all individuals should be subjected to the same procedures for diagnosis, including laboratory diagnosis if required, to avoid missing subclinical or asymptomatic cases and erroneously including them in the control group. In addition to this, the eligibility criteria should also be defined, which should be met for inclusion in either group. One example of diagnostic criteria is shown here.
  - ***Diagnostic criteria for hypertension*:** An individual showing systolic blood pressure of more than or equal to 140 mm Hg or diastolic blood pressure of more than or equal to 90 mm Hg or both, as measured on left arm with the subject in sitting position after 5 minutes of rest, or having a documented record of taking antihypertensive drugs.
- ***Eligibility criteria*:** In addition to the diagnostic criteria, some other criteria may be included to determine eligibility to participate in the study. These criteria may include age group, gender, socioeconomic stratum, geographical location, other health conditions, etc. This is essential to control for other factors that might have an influence on the exposure or outcome under research. Eligibility criteria have two components: inclusion and exclusion criteria. Inclusion criteria state the factors that an individual belonging to the reference population should meet, so as to be eligible to be included in the study. Exclusion criteria state the factors that, if present, disqualify the subjects to participate in the study, even if the subjects meet the inclusion criteria. An example of eligibility criteria is shown below.
  - ***Inclusion criteria*:** Adult population aged more than or equal to 40 years, of either gender, with (case group) or without (control group) hypertension as defined under diagnostic criteria, in absence of any diagnosed secondary cause of hypertension, namely, renal or hormonal causes.
  - ***Exclusion criteria*:** Subjects who are too ill to respond or are mentally deranged.
- ***Matching*:** The controls should be matched with the cases in terms of the factors that might have effect on the outcome other than the suspected etiological factor under study. Matching indicates the process of selecting controls that are similar to the cases with respect to certain relevant variables to avoid distortion of results or confounding. The concept of confounding has been explained in Chapter 6: Basic Concepts in Epidemiology.
  The most common factors considered for matching are age and gender, education, occupation, etc. However, overzealous matching may result in reducing the association. The process of selecting cases and controls based on the operational definition and the eligibility criteria ensures matching to a great extent. Matching may be done in groups or pairs. In group matching, stratification of the entire reference population is done on the basis of the criteria to be matched and cases and controls are selected from each stratum. In pair matching, once a case is selected, a matched control is selected following the diagnostic and eligibility criteria.
  - ***Group matching*:** Reference population is stratified as age group 40 to 59 years, 60 to 79 years, and 80 years and above and then same number of cases and controls are selected from each stratum.
  - ***Pair matching*:** For a 55-year-old male case who is a sedentary worker, a 55-year-old male sedentary worker control is selected. Similarly, all controls are selected on matching with every case, based on the factors decided for matching.

Step 4: Data Collection

Data regarding exposure as per operational definition of the study, are collected from both the case and control groups.

Step 5: Analysis

The final step is analysis, by which the hypothesis is tested and either accepted or rejected. The steps of analysis are as follows:

1. ***Construction of table***: A 2 × 2 contingency table is first constructed for analysis, as shown in **Table 8.6**.
2. ***Calculating exposure rate***: This determines the rate of exposure, i.e., to smoking in the example given here, which is calculated for both the diseased and nondiseased groups. According to the hypothesis that the occurrence of hypertension is associated with smoking,

Table 8.6 A 2 × 2 contingency table

Study group	Case (hypertensive)	Control (normotensive)	Total
Smoker	a	b	$a + b$
Nonsmoker	c	d	$c + d$
Total	$a + c$	$b + d$	$a + b + c + d$

the exposure rate of smoking should be more in the case group than the exposure rate in the control group, i.e., $(E_{Case}) > (E_{Control})$.

- Exposure rate in cases $(E_{Case}) = a/a + c$.
- Exposure rate in controls $(E_{Control}) = b/b + d$.

3. ***Estimation of risk***: Ideally, the estimation of risk should be calculated from incidence of hypertension in both groups, i.e., smoker and nonsmoker. According to the hypothesis, the incidence of hypertension should be more in smokers than the incidence of hypertension in nonsmokers. However, for calculating the incidence of disease in a population, a denominator is required, which is the total population of smokers and nonsmokers in whom hypertension is present. Since a case–control study is conducted on cases of the disease and control group without the disease, there is no denominator available. Hence, risk estimation cannot be done using incidence.

   Therefore, risk estimation is done in a different way in case–control studies. It is a measure of odds of exposure to the factor in case group versus odds of exposure in the control group. This is known as odds ratio (OR) or cross product ratio, the calculation of which is shown as follows. Concept of odds is explained in Chapter 10: Biostatistics.

   - Odds of exposure in case group = a/c.
   - Odds of exposure in control group = b/d.
   - $$\text{OR} = \frac{\text{Odds of exposure in case group}}{\text{Odds of exposure in control group}} = \frac{(a/c)/(b/d)}{= a/c \times d/b} = ad/bc$$

   Ideally, risk should have been estimated by relative risk (RR), which is equal to incidence in exposed divided by incidence in unexposed, as follows:

   - $$\text{RR} = \frac{\text{Incidence in exposed}}{\text{Incidence in unexposed}} = \frac{(a/a + b)}{(c/c + d)}$$

   In case of rare diseases, the values of "a" and "c" will be very small and hence can be ignored in the denominator. In such condition $a + b$ will be almost equal to b, and $c + d$ will be almost equal to d. The formula of RR can therefore be rewritten as follows:

   - $\text{RR} = (a/a + b)/(c/c + d) \approx (a/b)/(c/d) = ad/bc = \text{OR}$

4. ***Testing of hypothesis***: This is done by applying statistical test of significance. Exposure rate and OR may appear to be more in cases than in controls. However, this difference may also occur by chance. To rule out the chance factor, the differences detected should be statistically significant, which is done by applying a statistical test of significance. There are many such statistical tests of significance, the choice of which depends on the type of variables to be considered. Statistical tests of significance have been described in detail in Chapter 10: Biostatistics.

Advantages and Disadvantages of Case–Control Study

Case–control study has several advantages for which this type of study is performed. However, there are some important disadvantages also.

Advantages

- Suitable for rare diseases, as study starts with cases of the disease.
- Study of several etiological factors is possible.
- No risk to subjects, as both exposure and outcome have already occurred.
- No ethical issues, as subjects are no longer required to be exposed to suspected etiological factor.
- No problem of loss to follow-up or attrition, as subjects are observed at one time when data are taken regarding the past experience.
- Requires less number of subjects, as there is no attrition.
- Requires less time, as no follow-up is required.
- Inexpensive, as less number of subjects and less time is required.
- Easy to conduct.
- Risk estimation can be done.

Disadvantages

- Recall or memory bias can occur, as the estimation of exposure is based on history given by subjects, which may be faulty, or from records, which may be incomplete.
- Incidence cannot be measured.
- RR cannot be measured, and risk can only be estimated approximately from OR.

- Causal association cannot be established, as exposure and outcome have both occurred before initiating the study.

Examples of Case–Control Study

Some classical studies that have followed the case–control design are as follows:

- ***Adenocarcinoma of the vagina and diethylstilbestrol*:** In the late 1960s, adenocarcinoma was diagnosed in eight young women who were born in hospitals in Boston. Case–control study was done taking four matched controls of female births during the same time at the same hospital for each case, and it was observed that in seven of the eight cases, diethylstilbestrol was given to the mother during the antenatal period. Subsequent research confirmed this association.
- ***Thromboembolic disease and oral contraceptives*:** Case–control study was done for venous thrombosis or pulmonary embolism in women without any known medical condition. Cases were the women admitted in hospitals with venous thrombosis or pulmonary embolism and controls were women admitted in same hospital with other conditions and were matched for age, marital status, and parity. Use of oral contraceptives was much higher in the case group than control group.

Cohort Study

This is the next in hierarchy of epidemiological research. It is a longitudinal analytical study for testing of hypothesis, which is forward-looking and hence also known as prospective study. A cohort is defined as a group of people sharing a common characteristic or experience within a defined time period, e.g., babies born within a defined time period form a birth cohort. Similarly, there may be marriage cohort, academic group cohort, professional group cohort, exposure cohort for some specific exposure, etc.

The study is conducted with two groups of subjects: the study cohort and control cohort. The study cohort is the group that is exposed to the suspected cause, while the control cohort is the unexposed group. The two groups are followed up for the defined time period at the end of which the incidence of the disease in both the groups is measured and compared. Hence, the study goes forward from cause to effect.

Steps of Conducting a Cohort Study

Steps of conducting a cohort study are similar to that of case–control study. The difference lies in the eligibility criteria of study and control groups and in the manner analysis is done.

Step 1: Defining the Disease and Exposure

- ***Exposure*:** Exposure criteria need to be specifically defined before initiating the study. Also, measurement of the exposure should be done in the same manner in both study cohort and control cohort to ensure that exposure is not present in the control group.
- ***Disease*:** The disease under study should be defined with an operational definition drawn up as explained for descriptive study. Measurement of disease should ideally be done in the same manner in both case and control groups.

Step 2: Source of Study Cohort and Control Cohort

Subjects can be selected from the general population as per specified eligibility criteria. Exposure cohort may also include subjects from specified geographical location, special groups such as select groups based on occupation (health personnel, policemen, pilots), educational groups (college students, school students), volunteers, insured persons, etc. Subjects may also be exposure groups that include those known to have been exposed to a certain factor that is suspected to be an etiological cause of a certain outcome. Examples of exposure groups are radiologists exposed to radiation, factory workers exposed to certain chemicals, security guards exposed to prolonged hours of standing, traffic police exposed to air pollution, people living in areas with high arsenic level in water, etc.

Step 3: Selection of Study Subjects

Diagnostic and eligibility criteria should be specified, including both inclusion and exclusion criteria, as described in detail under case–control study. Data should be taken properly to confirm absence of exposure in the control group to avoid misclassification resulting in selection bias.

If exposure is present, data should include the type as well as the degree and duration of exposure to the factor under study. Comparison may be made as follows:

- ***Internal comparison*:** Here, no separate group is taken. Comparison is made between subjects with various degrees and duration of exposure, within the study cohort itself. This helps to establish dose–response and duration–response relationship for establishing causal association.
- ***External comparison*:** When the degree or duration of exposure is not available, or when the exposure is not a quantitative variable, a separate group may be taken as control cohort. This group should be matched with the study cohort with regard to all possible factors that might influence the outcome, other than the exposure under study. Matching has been explained under case–control study.

- ***Comparison with the general population rates*:** This is a simpler way of making comparison, i.e., with rates of various subgroups of the general population. However, for this purpose, authentic and reliable data must be available.

Step 4: Follow-up

- Both the groups are followed up for the defined period of time to detect occurrence of disease. Contact details of all subjects should be noted down properly, along with those of other family members, neighbors, and friends, so that they can also be contacted if required. However, in spite of best and utmost efforts, some amount of loss to follow-up or attrition may happen due to various reasons such as movement out of the area, loss of interest, and death. Attrition is more common in case of studies of long duration. This problem is addressed by increasing the sample size to cover the loss. An attrition rate of 10% is acceptable, beyond which the results are generally not considered to be valid.

Step 5: Data Collection for Ascertaining Outcome

For this purpose, same method of data collection and same diagnostic work-up should be undertaken for both the groups, including examination and investigation. Data regarding disease, as per operational definition of the study, are collected from both the case and control groups. Data may be collected in any of the following ways:

- ***Periodic examination*:** Here, each subject of both study and control cohort is examined periodically, at specified intervals depending on the requirement of the outcome in question. This is the best method of collection of data during the period of follow-up. However, this is possible only when the sample size is small and resources are adequate.
- ***Record review*:** Here, records of hospitals or physicians' clinics are reviewed to detect occurrence of disease of interest in the study population. This is possible only when there is a very accurate and well-developed record maintenance and record linkage system.
- ***Routine surveillance of morbidity and mortality statistics*:** This is not a very effective method as all the health care establishments that the study population is likely to attend may not be included in the surveillance system.
- ***Mailed questionnaire, telephone calls, periodic home visits*:** This is the easiest way and most commonly followed, particularly in case of a large sample. However, the subjects should be educated and highly motivated to understand the importance of the study and their role in it, and cooperate throughout the period of study.

Step 6: Analysis

The final step is analysis, by which the hypothesis is tested and either accepted or rejected. Following are the steps for analysis in cohort study:

- ***Construction of table*:** The study started with two groups as smoker and nonsmoker, who were followed up to detect development of hypertension. Similar to case–control study, here also a 2 × 2 contingency table is constructed. This is shown in **Table 8.7**.
- ***Calculation of disease incidence rate*:** This determines the rate of developing the disease, which is calculated for both the groups. According to the hypothesis that the occurrence of hypertension is associated with smoking, the incidence rate of hypertension should be more in the study cohort or exposed group, i.e., in smokers, than incidence rate in the control cohort or unexposed group, i.e., in nonsmokers, i.e., $(I_{Exp}) > (I_{Unexp})$.

  Incidence rate in study cohort $(I_{Exp}) = a/a + b$.

  Incidence rate in control cohort $(I_{Unexp}) = c/c + d$.
- ***Estimation of risk*:** The risk of developing disease in those who are exposed is estimated in many ways through cohort study, the three main ones being RR or risk ratio, attributable risk (AR) or risk difference, and population attributable risk (PAR).

  - $\text{RR} = \dfrac{\text{Incidence in exposed}}{\text{Incidence in unexposed}} = \dfrac{(a/a+b)}{(c/c+d)}$
  - $\text{AR} = \dfrac{\text{Incidence in exposed} - \text{incidence in unexposed}}{\text{Incidence in unexposed}}$
  - $\text{PAR} = \dfrac{\text{Incidence in population} - \text{incidence in unexposed}}{\text{Incidence in population}}$

  RR is expressed in number and may have a decimal, e.g., RR = 2 or RR = 2.5. It is interpreted as the risk of acquiring the disease by the exposed individual is two times or two and a half times higher than the risk in unexposed individual. RR < 1 indicates the exposure under consideration has a protective effect.

Table 8.7 A 2 × 2 contingency table

Study group	Developed hypertension	Did not develop hypertension	Total
Smoker	*a*	*b*	*a* + *b*
Nonsmoker	*c*	*d*	*c* + *d*
Total	*a* + *c*	*b* + *d*	*a* + *b* + *c* + *d*

AR and PAR are expressed as percentages. AR indicates the amount of risk in the exposed group that can be attributed to the exposure. This is so because certain degree of risk remains in all individuals, whether exposed or not. This is the baseline risk that is considered to be the risk of developing the disease by the unexposed group. AR indicates the possible reduction of risk that will be achieved if the exposure is eliminated. For example, AR of 80% means that 80% of total risk in the smoker can be attributed to smoking, and the rest 20% is due to some other factors. Hence, cessation of smoking will result in reduction of risk of developing hypertension by 80%.

PAR estimates the risk that can be attributable to the exposure in the total population, as the population comprises both exposed and unexposed individuals in varying proportions. PAR indicates the possible reduction of risk that will be achieved if the exposure is eliminated in the population. This helps in assessing effectiveness of public health programs targeted toward the entire population, whereas the elimination of exposure will happen only in the exposed group. For example, in a situation with PAR of 60%, following a tobacco cessation health education program the rate of smoking will be reduced in the population by 60%. This is due to cessation of smoking by the smokers or exposed group, but there will be no change in the unexposed group as they were not smoking prior to the program. PAR is of use in taking managerial decisions.

- ***Testing of hypothesis*:** Similar to that in case–control studies, this is done by applying statistical test of significance. Incidence rate and the various risks (RR, AR, PAR) may appear to be more in exposed group than in unexposed group. However, this difference may also occur by chance. To rule out the chance factor, the differences detected should be statistically significant, which is done by applying a statistical test of significance, which has been described in detail in Chapter 10: Biostatistics.

Advantages and Disadvantages of Cohort Study

This type of study is superior to case–control study and has several advantages for which this is performed. However, there are some important disadvantages also, as follows.

Advantages

- Suitable for rare exposures, as study starts with exposure to the suspected etiological factor.
- Study of several outcomes due to the exposure of interest is possible.
- Recall or memory bias is absent as subjects are followed up.
- Incidence can be measured.
- RR, AR, and PAR can be measured.
- Causal association can be established, as exposure is present but outcome has not occurred at the time of initiating the study.
- Dose–response and duration–response relationship can be established.

Disadvantages

- Problem of attrition, as subjects are observed over a long period of time.
- Requires large number of subjects to cover for possible attrition.
- Requires more time as it involves follow-up of subjects.
- Expensive as more number of subjects and more time are required.
- Difficult to conduct.
- Risk to subjects as exposure is present.
- Ethical concerns present, as subjects should not be continued to be exposed to known harmful etiological factor.
- Regular follow-up may alter subject's behavior and hence subject may not be eligible to remain in the exposed group.
- Subjects in control group may become exposed to the factor.
- Diagnostic criteria may change over time.

Types of Cohort Study

- ***Prospective cohort study*:** Here, at the start of the study, the disease has not occurred in either of the two groups. This is the classical type that has been described in detail.
- ***Retrospective cohort study*:** Here, both exposure and outcome have occurred and the study goes back in time. However, it is different from case–control study as here the comparison groups are based on status of exposure, while in case–control study comparison groups are based on status of disease.
- ***Combination of prospective and retrospective cohort study or ambidirectional cohort study*:** This type of study starts somewhere in between and follows up the subjects for the rest of the period of study. For example, in a planned 5-year observation period, the study may look back on data of past 4 years and subsequently follow up for 1 year, thus completing the total 5-year period.

In case of any retrospective element in a cohort study, all data should be based on available records and not on history given by the subjects. Outcome data are easily available from medical records, but it is difficult to get exposure data in most cases. Hence, this kind of study is mainly suitable for occupational exposures where proper documentation is present.

Examples of Cohort Study

Some classical studies that have followed the cohort design are as listed here.

- ***Prospective cohort study for multiple risk factors*:** The Framingham Heart Study (1948) is a prospective cohort study initiated in 1948 to assess the relationship of multiple risk factors with development of coronary heart disease. The population of Framingham, United States, aged 30 to 59 years was examined every 2 years for a period of 20 years. The study has been able to establish association of coronary heart disease (CHD) and death due to it with multiple risk factors such as male gender, advancing age, cigarette smoking, obesity, high blood pressure, high blood sugar, and high serum cholesterol. The study is now continuing on its third generation of participants.
- ***Prospective cohort study for single risk factor*:** Doll and Hill's study on smoking and lung cancer (1951) is a prospective cohort study for a single risk factor, smoking, conducted on British doctors. Two cohorts, smoker and nonsmoker, matched for age, education, and social class, were formed and the participants were sent questionnaires on smoking habits. Subsequent questionnaires were sent at regular intervals to identify change in smoking habits. Follow-up of incidence of cancer and cause-specific mortality is still continuing. This study was able to establish causal association between cigarette smoking and risk of death from lung cancer, myocardial infarction, and chronic obstructive pulmonary disease.
- ***Retrospective cohort study*:** A retrospective cohort study was conducted in 1974 after cases of angiosarcoma of the liver were reported among workers of a factory producing polyvinyl chloride in Kentucky, United States. RR of angiosarcoma was higher in those workers exposed to vinyl chloride monomer.
- ***Combination of prospective and retrospective cohort study*:** Court-Brown and Doll's study on effects of radiation conducted in 1955 is a combination of prospective and retrospective cohort study done on individuals who had received large doses of radiation therapy for ankylosing spondylitis over the last 20 years, i.e., between 1934 and 1954. In 1955, after initiating the study with retrospective data, it was further followed up prospectively. The study showed association between exposure to radiation and subsequent death from leukemia or aplastic anemia.

Calculation of Estimated Risk in Analytical Studies

Estimation of risk can be made from both case–control and cohort study, as explained. Risk calculation for both designs is shown with a hypothetical example in **Table 8.8**.

$$\text{OR} = \frac{a/a}{bc} = \frac{8 \times 25}{15 \times 2} = 6.67$$

$$\text{RR} = \frac{a/a+b}{c/c+d} = \frac{0.35}{0.07} = \frac{184}{54} = 5$$

$$\text{AR \%} = \frac{(a/a+b) - (c/c+d) \times 100}{(a/a+b)}$$

$$= \frac{(0.35) - (0.07) \times 100}{(0.35)} = \frac{0.28 \times 100}{0.35} = 80\%$$

$$\text{PAR \%} = \frac{(a+c/a+b+c+d) - (c/c+d) \times 100}{(a+c/a+b+c+d)}$$

$$= \frac{(10/50) - (2/27) \times 100}{(10/50)} = \frac{0.2 - 0.07 \times 100}{0.2}$$

$$= 65\%$$

Thus, from the example, risk measures are as follows:

- OR = 6.67.
- RR = 5.
- AR = 80%.
- PAR = 65%.

Note that while OR and RR do not have any unit, AR and PAR are expressed in percentage.

Hybrid Studies

Some studies incorporate two or more study designs within them and are hence known as hybrid study designs. These are recently being widely used in epidemiological research. The commonly used quantitative hybrid designs are described here. Hybrid study may also be a combination of quantitative and qualitative methods.

Table 8.8 Calculation of risk

Study group	Developed hypertension	Did not develop hypertension	Total
Smoker	8 (a)	15 (b)	23 ($a + b$)
Nonsmoker	2 (c)	25 (d)	27 ($c + d$)
Total	10 ($a + c$)	40 ($b + d$)	50 ($a + b + c + d$)

Nested Case–Control Study

This is a hybrid design combining two analytical study designs into one, as a case–control study nested in a cohort study. In this design, a population free of the disease is identified and initial or baseline data are collected from the entire population by interview, examination, and/or investigation, which are stored without analysis. For the purpose of investigation to be conducted later, biological samples collected, such as blood, urine, stool, and tissue, are frozen and preserved. The enrolled population is subsequently followed up over time to identify occurrence of disease. As and when disease occurs in any individual, a matched control is selected and both the case and control are enrolled for the case–control part of the study. When the desired sample size of case–control pairs is obtained, the study is terminated and case–control analysis is done. The exposure data or risk factors are obtained from the baseline data collected at the start of the study. The frozen samples are thawed and tested for the parameters of interest in the study.

Advantages of this nested case–control study design are as follows:

- Reduced selection bias, as cases and controls are selected from the same population.
- Recall bias is absent as exposure/risk factor data are collected before disease occurs.
- Since exposure/risk factors were present before occurrence of disease, causal association can be established.
- It is economical and hence cost-effective to conduct, particularly in clinical studies involving biological sample testing, as analysis of biological samples is required to be done only for subjects selected for the case–control study and not for all members of the parent cohort.
- Incidence and hence risk estimation can be done as the denominator is available, which is the entire population initially enrolled in the study and from whom the disease subjects are identified.

However, one major disadvantage is that it can be conducted for only one disease at a time, as studying multiple diseases will make it complex because multiple control groups will need to be selected. This problem is overcome in the case–cohort study.

Case–Cohort Study

This design entails selecting a population free of the disease under study, from which a small sample is selected as a subcohort to serve as control. The entire population is followed up and cases are selected as and when they develop. Finally the cases are compared with the initially selected subcohort as control. This study design is similar to nested case–control study, the only difference lying in the process and time of selection of controls.

Advantages of this study design are similar to nested case–control design. In addition, there are some very important advantages of case–cohort design over nested case–control design, the main one being that the same control group can be used to compare with different diseases forming different case groups in the study. It is therefore flexible as it allows testing hypotheses that were not anticipated when the cohort was drawn at the inception of the study. Subcohort can be used to calculate person–time risk.

However, one great disadvantage of the case–cohort design is that it requires a more complicated statistical analysis, as the subjects in the subcohort may also become diseased and thus become cases during the period of follow-up.

Case-Crossover Study

This is a unique study design where the case serves as its own control, so that the study is self-matched. It is useful when exposure or risk factor is transient for a short period of time and the health problem is an acutely occurring one. Comparisons are made over different periods of time. Hence, each person is a case at the time of occurrence of the health problem, which is the case window, and a control at any other suitably considered period of time when the health problem was absent, which is the control window. The exposures or risk factors at these two periods of time are compared. This design is very helpful in studying accidents.

Advantages of case-crossover study design are as follows:

- It is efficient as it selects only cases, and the controls are self-matching.
- Multiple control windows can be selected and used for one case window.

The main disadvantage is that of information bias due to inaccurate recall of exposure during control window. Also, it requires careful selection of time period for the control window as the circumstances should be similar to circumstances associated with case window, which might always not be possible.

Experimental Epidemiology

In the earlier days, experimental study indicated research on laboratory-bred experimental animals such as rat, mouse, and guinea pig. Currently, experimental study indicates intervention study on humans as well. Classically, experimental study means an RCT, which is a true experiment. Experimental or intervention study is similar to cohort study, a difference being that in the former the exposure, which is an intervention or manipulation, which may be promotive, preventive, curative, or rehabilitative in nature, is under control of the researcher, whereas in cohort study, which is an observational study, exposure is usually a risk factor to the disease on which the researcher has no control and the researcher only observes the effects without taking any action. Another difference is that while

cohort study proves causal association between exposure and outcome, intervention study proves causal association between intervention and outcome.

Since intentionally exposing the study subjects has ethical issues, intervention or experimental studies are only conducted for testing prevention or treatment measures to test the effectiveness and efficiency of health services directed toward controlling the exposure. If at all done for studying causal association between exposure and outcome, it is done as study to see the effect of withdrawal of exposure or cessation of risk factors. Hence, these studies give further scientific proof of the role of etiological agents in causation of the disease. However, these studies involve serious ethical issues, require prolonged time, incur great cost, and are often not feasible to conduct.

Types of Experimental Studies

Experimental studies may be conducted on animals or humans.

Animal Experiments

These have been conducted since long time and are now used for the following purposes:

- Experimental reproduction of human diseases in animals that helps to study the natural history of the disease of interest, confirm the etiology of the disease, and observe the pathologic changes due to the disease.
- Test the efficacy, efficiency, and effectiveness of preventive and therapeutic measures to control the disease.

Earlier thought to be devoid of ethical issues, animals' rights organizations are now actively protecting the rights of animals. Hence, stringent ethical guidelines need to be followed in experimental research involving animals.

Animal experiments are advantageous as animals can be bred in laboratories and can multiply rapidly, diseases can be produced, and effects of exposure can be seen within a short time. However, all human diseases may not be possible to be produced in animals and also the prevention and control measures that test to be effective in animals may not work in the same manner and to the same extent in humans. One example of this latter condition is the trial of typhoid vaccine where although the alcohol-killed and preserved vaccine was shown to be more effective by laboratory study, subsequent human randomized controlled research proved the heat-killed, phenol-preserved vaccine to be more effective.

Human Experiments

Over ages, doctors in public health or clinical practice have many times applied measures that they have known to be effective through their knowledge and long professional experience. One classical example of this is the treatment resorted to by Ambroise Paré on his war-injured patients in 1537. The standard treatment at that time was cauterization of wounds by application of boiling oil. At one time when oil was not available, Paré used a digestive made of egg yolk, rose oil, and turpentine and found this to work much better than oil, with these patients feeling lesser pain and showing better wound healing. However, this was not a planned trial. Earliest known classical planned human experiment is probably the trial conducted by James Lind in 1757 on 12 sailors suffering from scurvy, where he divided them into six groups and gave different diet to each group and observed them for 6 days. At the end of the trial, he reported complete recovery in the patients taking diet comprising of two oranges and one lemon daily.

Human experiments, named as intervention studies, are the best kinds of research but carry ethical, logistic, and operational problems. With the introduction of research ethics, trials have become more methodical and systematic. Intervention studies are of two broad types: RCTs or true experimental study and nonrandomized trial or quasi-experimental study. These have been discussed in detail as follows.

Randomized Controlled Trial

RCT is done on individuals as the study unit. It can be done for testing new drugs, new procedures, new vaccines, new modes of health care, etc. A group is exposed to the intervention under study and the outcome is compared with another group, known as the control group, which is not given the intervention. Allocation to intervention and control group is done by randomization that helps to eliminate selection bias and makes the groups comparable. Advantage of randomization is more than matching, as in matching only some known variables can be matched, while randomization controls for even unknown or unrecognized factors, which may also be important in influencing the outcome. This is because both intervention and control subjects are selected from one study group. Through randomization, these factors are expected to be distributed evenly in all the groups. Hence, this process minimizes selection bias.

The term "control" in RCT also signifies that the experiment is conducted under totally controlled conditions, beginning from allocation of intervention to follow-up and compliance, which finally ensures unbiased outcome.

Steps of Randomized Controlled Trial

Before initiating the study, the details should be written in the form of a protocol, which should be strictly adhered to throughout the entire phase of conducting the study, so as to avoid occurrence of bias. Protocol is an essential requirement before conducting any research study, more so in intervention studies, as here the manipulation is under

control of the researcher and hence there is likelihood of change in the process in the absence of a protocol.

The steps of conducting the trial are detailed as follows.

Step 1: Selection of Study Subjects

- ***Selection of the reference population*:** This depends on the population for which the result of the trial is most relevant. This population is also known as the target population. This population may be limited to any age group, gender, geographical location, occupational group, etc., depending on the relevance and requirement of the study. For example, testing of efficacy of a childhood vaccine needs to be done on under-5 children, testing of an antihypertensive needs to be done on adult population, etc.
- ***Selection of the study group*:** This is done in similar manner as described under analytical study, following the defined eligibility and diagnostic criteria.

Step 2: Randomization

This is the essence of an RCT, which ensures unbiased estimate of effect of the intervention to be tested, as the entire study population has the chance of being selected as either intervention or control group. It is done only after the entire study population has been selected and all selected subjects have given consent. Randomization can be done in many ways, described here.

- ***Simple randomization*:** This is based on a single sequence of random assignments and can be done by flipping a coin, a die, a pack of cards, using random number table, or computer-generated random numbers. It is simple and easy to follow. However, it may result in unequal number of individuals in the two groups. For example, in deciding intervention and control group as either head or tail, odd number or even number, etc.
- ***Block randomization*:** This is done to ensure balance in size of the intervention and control groups over time, which overcomes the problem faced by simple randomization. In this, total selected study group is divided into small blocks with predetermined group assignments, the block size being multiples of the number of groups. Then all possible combinations of assignments are made within a block and the individuals within the block are assigned accordingly. For example, for allocation into two groups, blocks of size four may be taken and combinations made as AABB, ABAB, BBAA, BABA, ABBA, BAAB, with A and B determined as either intervention or control group before starting the process of randomization.
- ***Stratified randomization*:** The above two types may not ensure equal distribution of covariates, which might act as confounding factors, and hence the groups may not be comparable. This problem is addressed by stratified randomization, where the population is divided into various strata on the basis of relevant and important covariates, e.g., age, gender, presence of comorbidity, etc. Randomization is then done within each stratum. Multiple-stage stratification may also be done, for example, first the entire population is divided into two groups, namely, male and female. Then each group is divided according to age, namely, young and old. Thus, four groups are formed: young male, old male, young female, old female. Finally, randomization is done within each of these four groups. This ensures representation of all groups in both intervention and control population. This method is also known as covariate-adaptive randomization.
- ***Unequal randomization*:** Although normally equal number of subjects is assigned to each group, unequal randomization may be done when two therapies are to be tested, in which one is more costly or has more logistic or operational problems than the other. Hence, the less costly or less problematic therapy may be assigned to a larger group than the more costly or problematic one, the randomization ratio being taken as 2:1. In this situation, loss of statistical power is not much. However, ratio larger than 3:1 results in considerable loss of power and hence should not be used.

Step 3: Intervention

Once the intervention and control groups have been formed, the intervention is applied on the intervention group, which may be in the form of application of preventive or therapeutic measures, or withdrawal of risk factors. The control group may be given a placebo or nothing at all. However, in absence of any intervention on the control group, there may be risk of bias, which has been discussed later. Hence, care should be taken to conceal the identity of the groups to which the individuals belong, which can be done by blinding. For testing a drug, blinding is done by administering to the control group a placebo, which is an inert substance that is identical to the drug in external appearance having the same form, color, texture, taste, and smell as the drug being administered to the intervention group, but lacking in the active pharmacological component. However, in some cases, it may not be possible to give placebo or follow blinding, e.g., comparing medical versus surgical treatment. In any form, care should be taken to prevent contamination of controls. These concepts are described as follows.

Contamination

This means controls also being exposed to the intervention under study in some way or the other. This can commonly occur as follows:

- ***Medication*:** Control subjects may take the medication under trial by way of treatment elsewhere, taking over-the-counter treatment, taking combination drugs, etc. This can be prevented by giving the subjects a list of drugs that they should not take as long as they are participating in the trial.

- ***Health education*:** Control subjects may interact with subjects of the intervention group and get to know the messages and implement those in their lifestyle. This can be prevented by selecting the intervention and control subjects from different areas, which are located at a distance preferably separated by other habitations in between.

Blinding and Allocation Concealment

The word blinding came from blind-folding, which means to keep the participants from knowing or having information regarding implementation of intervention. It may also be called masking, but blinding is a better and more used and acceptable term. There is another term "allocation concealment," which, although sounds similar, is distinctly different from blinding. Allocation concealment is an action taken before intervention is applied, i.e., it protects the randomization sequence list until intervention is assigned to a subject, whereas blinding conceals the randomization sequence after the subjects are allocated to groups and intervention is applied. Also, allocation concealment can and should be carried out in every study, while blinding, although should ideally be done, may not be possible always, e.g., while testing surgical versus medical or oral versus parenteral treatments. Unblinded studies are known as open-label study. Blinding can be done in the following manner, of which the most commonly used method is double-blind trial.

- ***Open-label trial*:** Here, both the data collectors and the participants know to which group the participants belong.
- ***Single-blind trial*:** Here, the participants do not know to which group they belong.
- ***Double-blind trial*:** Here, neither the data collectors nor the participants know to which group the participants belong.
- ***Triple-blind trial*:** Here, neither the data collectors, nor the participants, nor the person analyzing the data know to which group the participants belong. The decoding information is available only with the principal investigator of the research.

Step 4: Follow-up

The subjects of both the intervention and control groups are followed up for the requisite length of time and at required intervals, the two factors being decided on the basis of the concerned research questions and the results expected. Follow-up should be in same manner and intensity for both groups and at all sessions of contact. Care should be taken to prevent loss to follow-up. However, some loss may occur due to various reasons, the most common reason being migration out of the area. Loss of interest may also be another factor, which can be minimized by motivating the subjects to continue to participate in the trial. In trials conducted over a prolonged period of time, death of the enrolled subjects may add to loss. Loss to follow-up is known as attrition. Acceptable rate of attrition is 10% and it should be kept within this limit. Sample size should be increased by 10% to account for the unforeseen attrition. More attrition makes validity of the data questionable. Another problem is noncompliance to the intervention, as follows. Because of this, drop-in and drop-out may occur.

Noncompliance

Subjects in the intervention group, although initially consented to participate, may later refuse to comply with the treatment due to various reasons, resulting in drop-out from the study. Some subjects may stop complying without disclosing the same, which results in error in the study results. Hence, stringent measures should be incorporated in the study protocol, to check for compliance by objective measures. On the other hand, subjects in the control group may inadvertently take the same treatment or medication containing the drug under trial, from elsewhere or for other reasons, resulting in drop-in. To prevent this problem, all subjects of both groups should be given a list of medications containing the drug under trial that they should not take, as long as they are participating in the research.

Step 5: Analysis

Analysis is done by comparing the outcomes of interest in both the groups. Although the research questions are usually framed as positive or beneficial outcomes that are expected to be obtained after the trial, the possible negative effects should also be explored. Differences obtained in both positive and negative outcomes should be tested by means of statistical tests of significance, as applicable. Results may sometimes show error due to presence of various types of bias, which have been explained in Chapter 9: Conducting Epidemiological Research.

Types of Analysis

One problem comes in analysis in the form of change in treatment allotted initially. Due to deteriorating health condition of the subject or inability to tolerate the drug, the intended treatment may be withdrawn and the patient may be put on regular treatment or surgery or placebo. For example, while comparing medical and surgical treatment, a subject randomized in the medical treatment group may subsequently require surgery. Hence, there are two ways in which analysis may be done in such cases, each with its fallacies.

- ***Per-protocol analysis*:** Here, the subjects are considered in the groups according to their intervention they actually receive. In the above-mentioned example, the subject will be considered in the surgical treatment group. This leads to error because the initial randomization sequence is broken. This analysis will include only those

subjects who took at least 80% of the medication and completed the follow-up visits.

- ***Intention-to-treat analysis*:** Here, the subjects are considered in the groups according to the intervention they were intended to receive. In the above-mentioned example, the subject will be considered in the medical treatment group. This leads to error because the subject has not actually received the intended treatment.
- ***Number needed to treat (NNT)*:** This is another important concept in analysis of intervention studies, which indicates the number of patients that are required to be given the treatment under research to produce one additional case of cure as compared with the treatment given to the control group. NNT measures the amount by which the intervention reduces the disease and hence indicates effectiveness of the intervention. It is calculated using the formula: NNT = 1/absolute risk reduction.
- ***Absolute risk reduction*:** It is the difference between incidence in control and intervention groups, calculated using the formula: absolute risk reduction = incidence in control group - incidence in intervention group.

Validity of Results

The ultimate objective of conducting a trial is to generalize the results to population beyond the study group. Two types of validity are of concern, as follows:

- ***Internal validity*:** This indicates that the results are valid for the subjects who were included in the study. Internal validity is usually present if the study is conducted properly.
- ***External validity*:** This indicates generalizability, which means that the results are valid for all similar population who were not included in the study, irrespective of their location. External validity is present if the group included in the study is an adequate representative sample of the entire population, using calculated sample size and proper sampling technique, in addition to the study being conducted properly.

Types of Randomized Controlled Trial Designs

RCT may be conducted in various ways, as described.

- ***Concurrent parallel study design*:** In this design, the subjects are randomly assigned to either of the study group and are continued in the same group throughout the entire duration of study. Here, the number of subjects required is more, as two separate groups should be maintained, i.e., intervention and control groups. However, due to simultaneous administration of treatment/placebo to the groups, time required is less. Ethical issues are involved due to withholding from the control group the privilege of new therapy that has been proven to be of benefit. Still, this is the most commonly used design.
- ***Cross-over study design*:** In this design, the subjects are initially randomly assigned to either of the study group as described in the previous design. However, after observing for a certain predetermined period of time the groups are switched over, i.e., the group that received treatment in the first phase now receives placebo and the group earlier on placebo now receives the treatment. Hence, each subject serves as his/her own control. Some time should be allowed to lapse before the second phase to allow for elimination of the medication from the body, which is termed the washout period. An advantage of this design is that it is ethically better as all patients receive the treatment at some point of time. Also, it is economical as less number of subjects is required. However, time required will be more for each group to go through the intervention and control phases. This design is possible only with drugs that do not leave any residual, carry-over, or long-term effect in the body. Crossover may happen in an unplanned manner also due to reasons such as withdrawal of consent by the subject and inability to tolerate the new drug.
- ***Factorial study design*:** Here, two or more drugs are tested simultaneously. To the control group, either a placebo may be given or no drug given at all. For example, to test two drugs, the total subjects enrolled in the study are randomized into four groups as follows:
  - Group 1: drug A and placebo B, i.e., only drug A.
  - Group 2: drug B and placebo A, i.e., only drug B.
  - Group 3: drug A and drug B, i.e., both drugs.
  - Group 4: placebo A and placebo B, i.e., no drug.

This design is cheaper as only one set of subjects is required for conducting trial for two drugs and time is also less due to simultaneous administration of the two drugs. However, it is feasible only if the two drugs are different in their mechanism of action and outcome. Also, simultaneous administration of the two drugs should not interfere with each other's actions, i.e., there are neither synergistic nor antagonistic effects.

Cluster Randomized Controlled Trial

This is a type of RCT where instead of the individual subjects, communities or groups are randomized. This is done where the intervention cannot be withheld from any of the subjects in a group, either due to ethical or operational reasons. Some examples where classical RCT cannot be undertaken and hence cluster RCT needs to be done are as follows:

- For testing a preventive strategy in a population, e.g., for chlorination of water, it is not operationally feasible to give intervention to a selected group within the population, as the water supply is likely to be the same in the entire community.

- For testing a nutrition supplementation program in under-5 children with malnutrition, it is unethical to not give the intervention to some children, while giving it to others in their presence. Similarly, it is unethical to deny an intervention that is believed to be of benefit.
- For testing a health education package in school children, it is neither ethical nor feasible to withhold information from some students, while giving it to the selected population. It is not feasible because the information is likely to percolate from the intervention group to others, the process being known as contamination, which has been explained earlier.

Types of Randomized Controlled Trial

Broadly, there are three types of RCTs that are commonly undertaken. These are as follows:

- ***Randomized controlled clinical trial*:** In this type of trial, diagnostic or therapeutic interventions are tested on diseased persons. Examples of this type of trial are testing of new drugs, surgical procedures, methods of diagnosis, etc.
- ***Randomized controlled field trial*:** In this type of trial, promotive and preventive interventions are tested on healthy persons. Examples of this type of trials are testing of new vaccines, chemoprophylaxis, cessation of risk factors, healthy lifestyle, etc.
- ***Community trial*:** Here, interventions are tested on entire communities and outcomes are measured in sample group from the communities or from indicators of the communities. This uses the cluster randomization method. Examples of this type of trial are testing of new public health measure, health service, legislation, etc.

Phases in Human Trial for a New Drug

A drug mostly goes through animal trial before being approved for human trial. In the human trial part, there are various phases of trial before being certified for market use, which also is followed by another postmarketing phase. Each phase of trial has some specific objective.

- ***Phase I*:** Clinicopharmacological study with small sample size of usually 20 to 100 patients followed over several months for assessing pharmacological effect, dosage, and toxicity. This phase helps to identify the maximum tolerated dose (MTD).
- ***Phase II*:** Clinical investigation of 100 to 200 patients followed over several months to 2 years for biological activity, efficacy, side effects, relative safety, and adverse events at the MTD.
- ***Phase III*:** Large-scale RCT usually as multicentric study, with 300 to 3,000 patients followed up for 1 to 4 years, for effectiveness and monitoring of adverse reactions. Once this phase is passed, it is followed by licensing for marketing the drug. However, there may be some adverse effect that might require longer duration to manifest, such as carcinogenesis and teratogenesis, for which phase IV trial is essential.
- ***Phase IV*:** This is postmarketing surveillance for monitoring the new drug to identify any long-term adverse effects that might not have manifested during RCT in phase III, which might become evident when the drug is used by large population. One classical example is the thalidomide tragedy where the drug thalidomide given during pregnancy had passed the phase II trial and was in use. However, some years after its extensive use, thousands of babies were born with phocomelia or deformity of long bones of upper limb, along with several other congenital deformities. The drug was withdrawn following this.

Advantages and Disadvantages of Randomized Controlled Trials

RCTs are at the highest level in the hierarchy of primary research. Although this study design has several advantages for which it holds such a position, there are some disadvantages too. These have been discussed here.

Advantages

- These are considered to be gold standards for testing interventions.
- Eliminates selection bias.
- As both arms are selected from the same group, these are balanced with respect to prognostic variables that may be known or unknown.
- Each participant has the same chance of receiving any of the interventions.
- Allocation is carried out using a chance mechanism so that it neither is known in advance nor can be predicted as to which group the individual subjects will be assigned.
- The groups are of similar size and constitution, which can be further assured by conducting stratified randomization.
- Simple and easy to implement.

Disadvantages

- Problem of attrition, as subjects are observed over a long period of time.
- Requires large number of subjects to cover for possible attrition.
- Requires more time as it involves follow-up of subjects.
- Expensive, as more number of subjects and more time is required.
- Ethical concerns present, as subjects cannot be denied treatment known to be beneficial.

- Noncompliance may occur.
- Drop-in and drop-out may happen.
- Treatment method may change over time.
- External validity may not be present.

Examples of Randomized Controlled Trials

Some classical RCTs that have been conducted are briefly described here.

- ***Randomized controlled clinical trial*:** Hypertension detection and follow-up program to investigate benefits of treating mild to moderate hypertension (1979) was conducted as a multicentric RCT with hypertensive patients as participants, where the intervention group received stepped care for hypertension and the comparison group was referred back to the physician treating them. Results showed that the stepped-care group experienced lower mortality over a 5-year follow-up period.
- ***Randomized controlled field trial with factorial design*:** Physician's health study for simultaneous trial of aspirin to reduce CVD and beta-carotene to prevent cancer (1989) was conducted in the United States on male physicians aged 40 to 54 years as a randomized, double-blind, placebo-controlled trial using factorial design. The trial showed conclusive evidence of reduction in risk of myocardial infarction.
- ***Preventive trial for cessation of risk factor*:** Multiple Risk Factor Intervention Trial or MRFIT (1986), conducted in the United States, was a nationwide RCT to test the effect of primary prevention of risk factors of coronary heart disease and death from it. Men aged 35 to 57 years were randomized into special intervention (SI) group or control group receiving usual health care in the community. The SI group received counseling for cessation of cigarette smoking, dietary advice for lowering blood cholesterol level, and stepped-care treatment for hypertension. At the end of follow-up care for average 7 years, risk factor levels declined in both the groups, but more so in the SI group. There was decline in CHD-related mortality too in the SI group, although this was not statistically significant.

Quasi-Experimental Study

These are also known as nonrandomized trials or nonexperimental trials, and are in the lower rung of hierarchy of epidemiological studies compared with RCTs. These are conducted in conditions where RCTs or cluster RCTs are not possible due to ethical or operational reasons. Randomized trial of an intervention may not be the right option for the following reasons:

- Ethical considerations.
- Community interventions.
- Difficult-to-randomize subjects.
- Difficult to randomize by locations.
- Difficult to conduct in field settings
- Long natural history of disease.
- Low disease frequency.
- Small sample size available.

There are different types of nonrandomized or quasi-experimental study designs according to availability of preintervention data and inclusion of control group as outlined in **Table 8.9**.

Uncontrolled Trial

This type of study is conducted when a new screening or therapy method is started in a population. Since there is no control group, findings are compared with results in earlier unscreened or untreated patients who serve as historical controls, e.g., cervical cancer prognosis after introduction of visual inspection with acetic acid for identification of early cases at the peripheral level of health care facility.

Natural Experiment

In this type of study, the study populations are naturally selected, e.g., smoker and nonsmoker, migrant and nonmigrant, disaster affected and not affected.

Before and after Intervention Study

In this type of study, preintervention data are collected, intervention is given, and then the population is followed

Table 8.9 Nonrandomized study designs

		Control group	
		Absent	Present
Preintervention data	Absent	Uncontrolled experiment	Natural experiment
	Present	Before and after comparison study without control	Before and after comparison study with control

up for a defined period of time. This study may be conducted with or without control group.

- ***Before and after intervention study without control*:** Here, there is only one group in which intervention is given. The pre- and postintervention data are compared to see the effect of intervention, e.g., mHealth intervention given to a group of patients of NCDs on lifestyle modification. The problem with this design is that it may be difficult to ascertain that the change was effected by the research intervention and not due to influence of any extraneous factor acting simultaneously.
- ***Before and after intervention study with control*:** In this study design, there are two groups: intervention and control. Data are collected from both the groups, after which intervention is given only to the intervention group. Both the groups are then followed up and postintervention data are collected after a defined period of time. However, ethical issues may arise in depriving the control group from intervention. Hence, this is applicable for conditions that are not serious or life-saving, e.g., mHealth intervention given in a group of elderly healthy subjects on healthy lifestyle.

Three types of comparison are made as follows:

- Baseline data of both the groups are compared to see that both the groups are matching in terms of all possible confounding factors.
- Pre- and postintervention data of the intervention group are compared to assess effect of intervention.
- Postintervention data of both the groups are compared to establish that the effect is due to the intervention only and not due to some extraneous factors that might have influenced both the groups.

Examples of Nonrandomized Trials

Some classical nonrandomized trials that have been conducted are as listed here.

- ***Uncontrolled trial*:** Pap smear test was introduced in 1920 for screening for cervical cancer. Since no woman could be denied the test, there was no control group for such experiment. Hence, data regarding mortality of women due to cervical cancer before introduction of Pap test were compared with the data collected from the uncontrolled trial. It helped to establish evidence that Pap smear test is associated with reduced mortality from cervical cancer.
- ***Natural experiment*:** Smoking ban was in effect in public places, including bars and restaurants, in Helena, Montana, for 6 months in 2002, following which rate of heart attacks dropped. However, the law was suspended after 6 months following which the rate of heart attacks increased to earlier levels. In this experiment, the population was naturally selected into smoker and nonsmoker groups.

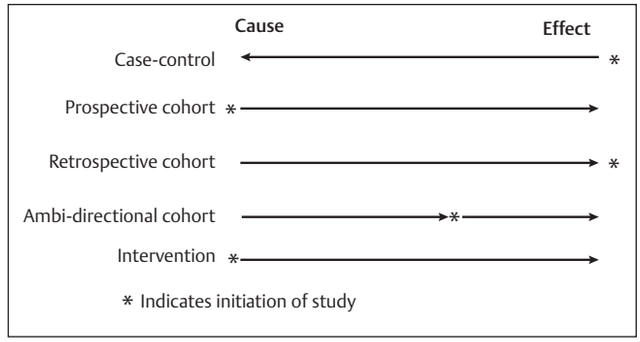

Fig. 8.5 Direction of longitudinal studies.

- ***Before and after intervention study*:** Introduction of legislation in Victoria, Australia, for use of seat belt in motor vehicles for prevention of injuries and deaths caused by motor vehicle accidents is a classic example of before and after intervention study. There was significant decrease of injuries and deaths due to motor vehicle accidents in Victoria compared with the status in the previous year as well as compared with other states of Australia. It used both designs, i.e., without and with controls, as described.
  - ***Without control*:** Data regarding injury occurrence following introduction of legislation were compared with data procured prior to the legislation.
  - ***With control*:** Data regarding injury occurrence following introduction of legislation in Victoria were compared with data of other states of Australia where such legislation was not introduced.

Direction of Longitudinal Studies

Longitudinal studies may be forward- or backward-looking or even ambidirectional. This concept has been explained through **Fig. 8.5**.

Review Studies

All the designs described earlier are primary research, where the investigator collects the data used in the research. There are other ways of data collection for conducting research on any identified problem, as follows:

- ***Secondary research***: Analysis is done on secondary data from existing datasets that have not been collected by the researcher and the researcher has not had any contact with the subjects. Such data are usually collected by some other agency, the source of which may be records of health facilities, reports of organizations, national surveys, etc., many of which may also be available on public domain.
- ***Tertiary research***: This involves collating findings of other published literature, where the research has been

done by other researchers, to arrive at a conclusion based on available evidence. These are called review articles, which may be of many different types. The most commonly used types are described here.

Types of Review Articles

Review articles may be of the following common types:

- Literature review or desk review or narrative review.
- Scoping review.
- Umbrella review or overview of review.
- Rapid review.
- Systematic literature review with or without meta-analysis.

These are described in brief here.

Literature Review

This is the traditional literature review, also known as narrative review or desk review, which examines and reviews existing authentic and scientific published articles, which have been subjected to a peer-review process. These reviews are often topic-based and provide an overview of what is known about the particular topic in the form of a discussion. The results or conclusion of a literature review is likely to be presented in a narrative format rather than statistical methods. Uses of literature review method are to identify and sum up what is known from previous research and identify omissions or gaps in earlier research, which gives a platform to further work on the problem and its solutions. It also avoids duplication of work and thus prevents wastage of resources on areas that have already been worked upon earlier. However, the method of conducting a literature review is arbitrary and hence there are chances of the results being biased. Contrary to systematic review, in a traditional literature review, the research question is often unclear and nonspecific; literature sources, search strategies, and basis of selection are not specified; quality assessment of included studies is also not done meticulously. This type is usually included in postgraduate thesis or research project report as review of literature.

Scoping Review

This is an exploratory study addressing a broad question to assess the extent of the available evidence, organize it into groups, and highlight gaps. If a scoping review finds no studies, this might help researchers to decide that a systematic review is likely to be of limited value and that resources could be better directed elsewhere.

Umbrella Review or Overview of Review

This is a review of multiple systematic reviews, done with explicit and systematic methods to search for systematic reviews on related research questions in the same topic area and synthesize the results of the systematic reviews across important outcomes.

Rapid Review

This is similar to systematic review with the aim to produce a rigorous synthesis quickly based on a predefined research question. It is of use when there is time constraint or urgency, e.g., in epidemic or pandemic conditions such as the recent COVID-19 pandemic. However, it is not as stringent as a formal systematic review. Here, the emphasis is on a replicable prespecified search, restricted to articles published during a specific timeframe and screening methods that minimize the risk of bias.

Systematic Review

This is the best known type of review for systematically searching, reviewing, and synthesizing findings from available research to collate all empirical evidence that fit prespecified eligibility criteria in order to answer a research question. As the name implies, it is systematically done following standard guidelines. Several organizations specialize in conducting, publishing, and funding systematic reviews, of which reviews by the Cochrane Collaboration are internationally considered as the highest quality of evidence-based health research. Earlier systematic reviews were conducted only with RCTs, but recently observational study designs are also being considered for systematic reviews, which enable generation of rich evidence-based findings. A systematic review may be continually updated to incorporate relevant new evidence. Such a review is known as a living systematic review.

The characteristics of a systematic review are as follows:

- Specific research question.
- Clearly stated set of objectives.
- Defined literature sources and search strategies.
- Predefined eligibility criteria for studies to be included.
- Clear and detailed methodology for conducting the review with details regarding process of search, data extraction, and analysis.
- Systematic search to identify all studies, both published and unpublished, scientific or gray literature.
- Explicit quality assessment criteria.
- Assessment of validity of the findings of included studies.
- Collation and synthesis of findings of included studies.
- Systematic presentation of the findings of the review.

A specific clearly formulated research question is framed and a protocol or research plan is drawn up, with explicit and reproducible systematic methods to minimize bias, thus providing reliable findings from which conclusions can be drawn and decisions made.

Systematic review is done as a team work, where at least two individuals or teams search for studies to answer the research question using a highly sensitive search strategy.

The retrieved studies are then screened for eligibility using prespecified inclusion and exclusion criteria. Any disagreement between the two teams is resolved by a third team acting as arbitrator. The data are extracted from all included studies regarding study location, year, design, population, sample size, key findings, and strengths and limitations of the study, and the quality of the included studies is assessed. All studies, both published and unpublished, should be included, ideally in all languages, although this may not be possible always. Studies with findings that are not statistically significant and studies with negative findings should also be included. Finally, the extracted study data are synthesized and the results presented, with conclusion drawn in the end. A systematic review may or may not include meta-analysis.

Meta-analysis

Meta-analysis is the statistical analysis of systematic review that combines the results of quantitative studies and generates precise effect of the results. For conducting meta-analysis, the included studies should be comparable and also of good quality. Meta-analysis pools the results of several studies that were included in systematic review to create a summary statistic, with the use of specialized software. Two graphical presentations are made on analysis.

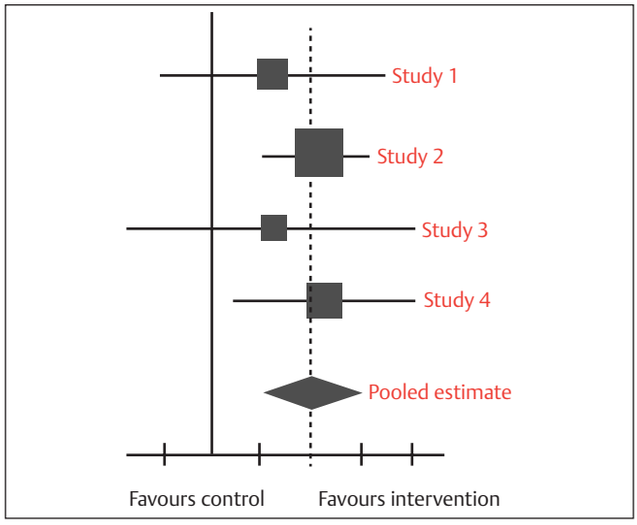

Fig. 8.6 Forest plot.

Forest Plot

This is used to display summary measures of studies included in the systematic review and meta-analysis, as follows (**Fig. 8.6**):

- Summarizes information given in individual studies.
- Depicts visually the heterogeneity of the included studies in terms of methods used, study subjects included, and results obtained.
- Gives a pooled result derived from combining results of the individual studies.

Funnel Plot

This is a graphical representation to examine the possibility of publication bias. It is basically a scatter diagram drawn by plotting the size of trials (y-axis) against the effect size reported (x-axis). The standard error of the effect estimate is usually taken as study size. In case of absence of bias and heterogeneity, the scatter is due to sampling variation. In such a condition, the shape of the plot looks like a symmetrical inverted funnel, indicating equal distribution of studies with positive and negative findings on either side. The more powerful studies, which are fewer in number, are toward the top or the apex, while studies of smaller size, which are more in number, are scattered widely at the bottom or the base. If publication bias is present, there may be a gap in the funnel (**Fig. 8.7**).

Qualitative Research

Qualitative research originated in the field of social and behavioral sciences. With complex problems in the present age, qualitative research methods are used extensively

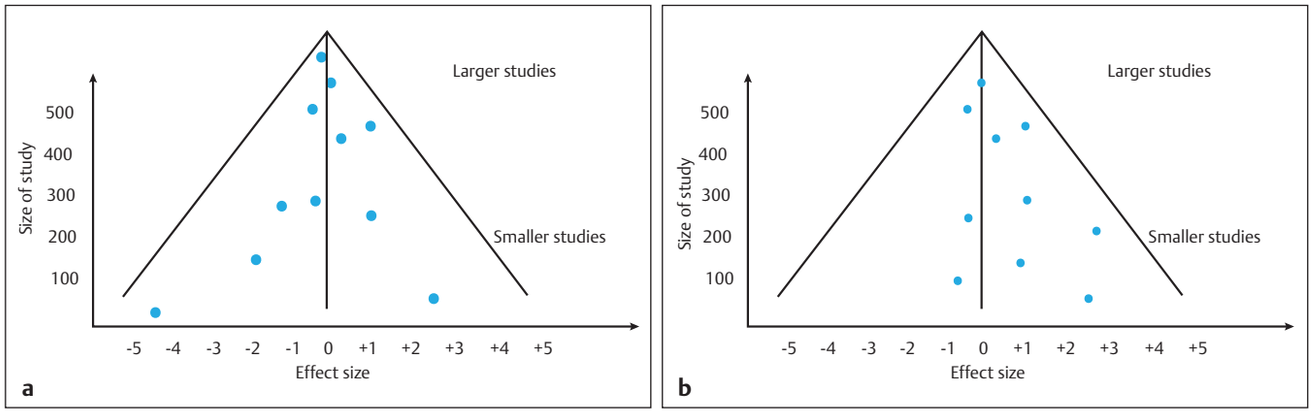

Fig. 8.7 **(a, b)** Funnel plot.

in health to understand people's thoughts, perceptions, ideas, and opinions. Qualitative research methods make it easier to understand as these are more communicative and descriptive. It aims to understand what people think and why they think so, thus helping in revealing the behavior and perception of a target audience, with reference to a particular topic.

Qualitative research focuses on obtaining data through conversational communication. Tool used for data collection is open-ended and not structured. It is known as a topic outline guide for discussion of the various aspects and recording responses of the participants, usually verbatim. Audio recording of the interview may also be done. In addition to the spoken words, the facial expression, gestures, and actions are also noted, which serve as rich nonverbal communication. These can be matched with spoken words to identify any disparity between the two, which can reveal truth or falsity of information provided. The results of qualitative methods are more descriptive, and the inferences can be drawn quite easily from the data that are obtained.

Approaches to Qualitative Research

The aim of qualitative research is to obtain a deeper understanding of the situation; therefore, the approaches to qualitative research tend to be flexible and obtain information-rich data in specific contexts. The common approaches used in qualitative research are described as follows.

Grounded Theory

In this approach, instead of starting the research with a theory, the researchers generate a theory that is "grounded" in data that have been collected and analyzed. This type of research is "inductive," i.e., it begins with observations/narrations, then finds patterns within these observations, and builds a theory based on analysis of these patterns. Grounded theory observes people and, based on the observations, develops a theory to explain a phenomenon of interest, i.e., it proceeds from data to theory.

It helps to understand people's perceptions and experiences related to health in great depth as it collects rich and variety of data to generate a theory or explanation based on views of a large number of study subjects. However, the disadvantages are that it requires collecting large amounts of data to generate a theory/hypothesis. Preconceived notions of the researcher also have the potential to introduce some degree of bias in the results.

Ethnography

This is the most in-depth observational method of a community or a group of people sharing common characteristics to study people in their naturally occurring environment. In ethnography, researchers are interested in understanding shared patterns of behavior, for example, a group, community, or tribe that resides in a geographical area and develops shared patterns of behavior, beliefs, and language. Here, the researchers become a part of the group to study people in their naturally occurring environment, to understand their culture, challenges, motivation, interactions, beliefs, behaviors, their way of functioning, and settings of a community. This helps to understand how these patterns of behavior can explain certain health phenomena.

Researchers can choose to have varying degrees of participation in the group, ranging from "complete observer" to "complete participant." The combination of immersive observations and interviews helps researchers describe the group and their behavior from both "emic" (views of participants) and "etic" (views of researcher) perspectives. This type of research method can continue for a long time and hence is quite challenging and time-consuming.

Phenomenology

In phenomenology, researchers study the experiences of the people regarding an event, concept, or phenomenon. Whereas ethnography studies the day-to-day lives of people to explain their cultural and behavioral patterns, phenomenology aims to capture the experiences of people who have gone through the same event or phenomenon. In health care, phenomenology is used to explain experiences of health-related situations, e.g., experiences of mothers who have had stillbirths or nurses delivering critical care services. Study participants are homogenous with respect to the phenomenon, i.e., they have all experienced the same event or phenomenon. Data are collected through in-depth interviews and sometimes through secondary sources such as documents or written pieces by the participants. Phenomenological studies explain "what" people have experienced and "how" they have experienced it.

Narrative Research

In this approach, researchers examine various people's stories to understand how people perceive and make sense of circumstances and situations. Narratives usually have a "chronology" and are described based on sequence of events narrated by the study participants. While phenomenology explains lived experiences of particular events and describes the essence or meaning of the event, narrative research explores the lives of individuals and stories about individual experiences using a biographical or storytelling format. Narrative research can include either one or more individuals, whereas phenomenology studies several individuals.

Case Study

In qualitative case studies, researchers describe a single problem, issue, or situation in detail. This approach is used to explain a single complex case, issue, event, or phenomenon of interest in detail in real-world settings. Like other

qualitative research approaches, case studies are useful to describe phenomena in real-world contexts. It involves careful and close observation of a community, a group of individuals, or a problem, event, or episode in an individual. It gives an insight into the behavioral pattern of the community or individual under study and helps to identify reasons for the same or any other outcome under study. However, the results cannot be generalized. In health care, the case study approach can be used in the following situations:

- *Intrinsic case study*: To describe a unique or novel phenomenon as, e.g., implementation of a new health service in a community.
- *Instrumental case study*: To gain in-depth understanding of a phenomenon, e.g., male participation in infant and young child nutrition.
- *Collective case study*: To gain a broader appreciation of a particular issue, e.g., implementation of electronic health records in tertiary care hospitals.

Further details about data collection and analysis in qualitative research are explained in Chapter 9: Conducting Epidemiological Research.

Check Your Progress

Scenario-Based Questions

1. **A new disease was reported in a community that was investigated by a research team headed by you. Answer the following questions in this context, with a hypothetical example.**
   A. What study design will be most relevant for this situation?
   B. Draw a map to show place distribution.
   C. How will you describe person distribution for this disease?
2. **A cohort study was conducted on 100 adolescents to find the association between occurrence of eye symptoms and using mobile smart phone. The study cohort consisted of 50 subjects who had a personal smart phone and the control cohort consisted of 50 subjects who did not have access to smart phone, either personal or in family. Results at the end of 6 months showed 20 subjects in the study group and 12 subjects in the control group complained of eye symptoms at any point of time during the total study period. Answer the following questions for this study.**
   A. Draw a 2 × 2 contingency table with these values.
   B. Calculate incidence of eye symptoms in both the groups.
   C. Estimate the possible risk of having eye symptoms in the exposed group.

Short Questions

1. **Explain the hierarchy of quantitative epidemiological research methods, with a diagram.**
2. **State the differences between case–control and cohort study.**
3. **Outline the methods of randomization used in randomized controlled trials.**
4. **Explain the concept of funnel plot in meta-analysis.**
5. **Describe in brief the grounded theory in qualitative research.**

Comprehensive Community Medicine for Undergraduate Students, is the first of its kind textbook, designed meticulously according to the latest National Medical Commission (NMC) Competency-Based Medical Education (CBME) Curriculum 2024. It empowers undergraduate medical students to acquire the knowledge, skills, attitudes, and competencies required of the Indian Medical Graduate. Though designed for undergraduate medical students, this book will also be of immense help to postgraduate students of community medicine. This comprehensive compendium seamlessly integrates all relevant topics and concepts of community medicine and public health. The entire content is spread across 40 chapters and divided into eight broad sections, namely, Health and its Determinants, General Epidemiology, Communicable Diseases, Non-communicable Diseases, Reproductive and Child Health, Health Care for Specific Conditions, Health Care Delivery, and Other Important Topics.

Whether preparing for university examinations, competency assessments, PG entrance preparation, clinical postings, or future medical practice, this book serves as a trusted companion in developing the vision and capabilities of a socially accountable physician. The book will act as an invaluable resource for undergraduate and postgraduate students, teachers, and public health professionals.

Salient Features

- Chapters are logically sequenced to ensure smooth flow while strictly adhering to the curriculum.
- Content is presented in an illustrative, concise, and easy-to-understand format, facilitating better comprehension and retention.
- Each chapter begins with clearly stated CBME competencies and a list of topics covered.
- Scenario-based questions and short questions at the end of every chapter enable self-assessment and performance evaluation.
- The book addresses important contemporary and public health–relevant topics, including:
    - o Hygiene and sanitation in public facilities and mass gatherings
    - o Health risks and their mitigation in international travel
    - o Health concerns of marginalized and vulnerable populations (street children, people with disabilities, sexual minorities)
    - o Emerging concepts such as digital technology in public health including artificial intelligence, organ transplantation, and health tourism
    - o Major 21st-century public health milestones, including policies and goals, eliminated and eradicated diseases, epidemics and pandemics, and newly introduced vaccines
- Chapter-wise compilation of key definitions in community medicine and public health aids quick revision and exam preparation.

B**ratati Banerjee, MBBS, DPH, DNB (SPM), MNAMS, PGDH&HM, WHO Fellowship in Non-Communicable Diseases,** is a fellow of the Indian Public Health Association and the Indian Association of Preventive and Social Medicine. Serving the Government of India as Director Professor of Community Medicine, formerly at Maulana Azad Medical College, New Delhi, and currently at All India Institute of Hygiene and Public Health, Kolkata, she has 30 years of experience in community medicine and public health. She has authored/edited six textbooks, a digital e-book for postgraduate students of community medicine, and contributed 30 chapters in textbooks and training modules. She has written two health education books for children and three chapters for school-level health education books. With over a hundred publications in national and international journals, she has led several funded research projects as Principal Investigator.

Online at MedOne

Access your complimentary online version directly from **https://medone.thieme.com** by using the unique code in the front of this book.

ISBN: 978-81-981265-7-3

www.thieme.in