Dr. Fred Kapaya

Student ID:

UD76152HPU85336

COURSE NAME:
Fundamentals of Epidemiology

Assignment Title:

Basic Concepts of epidemiology and epidemiological study designs

ATLANTIC INTERNATIONAL UNIVERSITY
April/2022

TABLE OF CONTENTS

1.0 Introduction

Epidemiology is regarded as a fundamental science of preventive and social medicine (Zodpey, 2015). It is a population science that supports health improvement and health care, and involves the distribution, frequency, trends, and aetiologies of disease. Although it is historically ancient, it has evolved rapidly over the past few decades. This evolution has occurred following changes taking place in society and occurrence of new diseases and new area related to epidemiology(Frérot *et al.*, 2018). Therefore, epidemiology goes beyond just the description of disease distribution and causation but also health related states or events. In addition, it also covers determinants of health and health problems.

It is a well-documented truth that population groups have unique and distinctive patterns of health and disease resulting from variations of exposure in individuals in the population. The exposure may be behavioural in nature or environmental and depending on degree of exposure, the individuals in the group may develop the same pattern of the disease or health related state(Decarli, 2003). Characteristic patterns of diseases or health states in populations arise from interaction of individuals in a social setting. For example, cholera occurs only when the community consumes water contaminated with Vibrio cholera. If adequate safe and clean water and improved sanitation and hygiene are provided in the community owing to improvement in social development, cholera is a public health threat would be eliminated in that particular community. Therefore, if society or environment changes, the risk of disease or health related states would also change.

Changes to society and the environment also results in changes in the risk of disease or health related states. Epidemiology as a science and practice aims to understand and use these patterns of disease occurrence, distribution and frequency to improve population health.

2.0 Definitions, Concepts and Epidemiological Approaches

Epidemiology is the basic science and practice of Public Health. An Epidemiologist is the community physician/practitioner who focuses on public health within a population. According to the 4th Edition of John M. Lasts Dictionary of Epidemiology,

"Epidemiology is the ***study*** *of the* ***distribution*** *and* ***determinants*** *of* ***health-related states or events*** *in* ***specified populations****, and the* ***application*** *of this study to the control of health problems"*(Last, 1993).

To understand this definition, it is dissected into five major components and these are(Last, 1993):

i. **Study-** epidemiology is a scientific discipline which uses scientific methods of inquiry. It includes surveillance, hypothesis testing, analytical research and experiments. It is data-driven and depends on structured and objective approach to data collection, analysis and interpretation.
ii. **Distribution**- refers to description by time, place and persons affected. It describes the frequency and pattern of health states in the population.

iii. **Determinants**- all the physical, biological, social, cultural and behavioural factors that affect health. These factors influence health conditions and other well-known characteristics in the affected population.

iv. **Health-related states**- include disability, causes of death, behaviours such as alcohol or drug abuse, amount of exercise, perceptions about prevention interventions, and provision and utilization of healthcare services. It is simply anything that affects the wellbeing of a population.

v. **Specified population**-are those with identifiable characteristics. Epidemiology focuses on identifying the exposure or source that has affected the people including the number of persons who may have been expose and the potential of affecting more people in the community

vi. **Application**-this involves applying the knowledge gained through scientific methods to understand the health status of the community and come up with effective pragmatic and satisfactory public health interventions to control and prevent disease in the community. It is the explicit hallmark of epidemiology to promote, protect and restore health.

2.1 Aims of Epidemiology

The overall aim of epidemiology is prevention of diseases, health events or states and promotion of health of populations. To achieve this purpose, epidemiology has three major aims(Zodpey, 2015)(Brownson, 2011). These are:

i. To describe the distribution and magnitude of health states of populations

ii. To identify the determinants of diseases and health states (aetiological/risk factors) responsible for health problems in populations

iii. To study the natural history and prognosis or disease or health states

iv. To provide the foundation for developing public policy, evaluation of services for the prevention and control of diseases and health problems and making regulatory decisions relating to environmental problems.

2.2 Basic Epidemiological Assumptions

There are two fundamental epidemiological presumptions:

1. Human health problems doesn't occur arbitrarily or at random
2. Human health problems are always caused by something and thus can be prevented by applying some defined measures

2.3 Basic Features of Epidemiology

- Studies are conducted on human population
- It examines patterns of events of people
- Can demonstrate a cause and effect relationship without necessarily understanding the biological mechanism or aetiological agent
- It covers a wide range of conditions
- It is an advancing science

The study of epidemiology relies on a systematic approach which are based on two major foundations(Kundi, 2006)(Zodpey, 2015):

i. Inquiry or conducting in-depth investigation to get answers which lead to more questions. These include questions related to health events (what, when, where, who, how why,); and questions related to health action including what can be done to reduce the problem and its consequences, how to prevent the problem in future, actions needed by all involved stakeholders, what resources are required and what are the anticipated challenges and how to overcome these challenges.
ii. Making comparisons is necessary to help draw conclusions on the differences between two or more groups and/or individuals. This helps to give a clue to the aetiology of the disease or health event. The groups or individuals being compared should be similar.

2.4 The Natural History of Disease

Diseases or health events result from a complex interaction between man (Host), an agent (cause of disease) and the environment. This is called an Epidemiological Triad. The Triad is used to show the link between the causes of diseases or health events, the environment and the host factors. It shows the evolution of disease or health problems from its earliest stage to its termination as either recovery, disability or death in the absence of any intervention. Figure 1 below depicts this link. It is an important framework as it helps to understand the pathogenic chain of events for a particular disease or condition and the application for preventive measures.

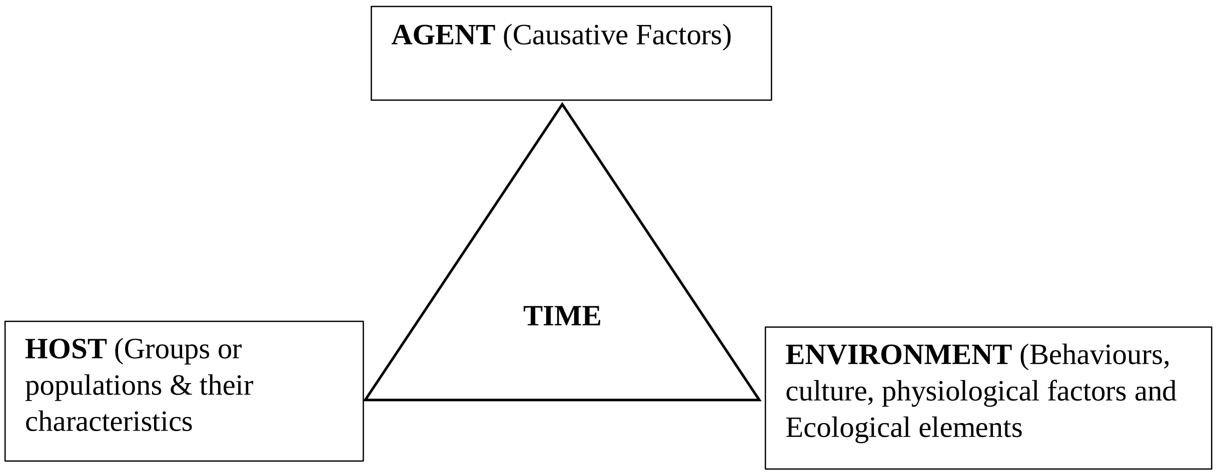

Figure1: Model of Epidemiological Triangle

2.5 **Basic Measurements in Epidemiology**

Epidemiology involves determining the associations between exposures and outcomes. In order to so, exposures and outcome must first be measured accurately. Epidemiology focuses om measurement of mortality and morbidity in human populations. Therefore, the first requirement is to define what needs to be measured and establish the benchmarks or standards by which it can be measured. Clear definitions help to minimize the errors in the classification of the data.

The span of measurement in epidemiology is very extensive and include:

i. Measurements of Mortality, Morbidity, Disability, Natality (birth rate)
ii. Measurement of the presence, absence or distribution of the characteristics or attributes of the disease

iii. Measurement of health needs, health infrastructure and equipment as well as utilization of health services

iv. Measurement of presence, absence or distribution of the environmental or other risk factors believed to be responsible for causing disease

v. Measurement of demographic variables etc.

The basic tools of measurement in epidemiology include the rate, ratio and proportion. A good understanding of these terms is critical for correct interpretation of epidemiological data.

a) Rate

This is a measure of the number of times an event occurs in a defined population over a specified period of time. It is a measure of the frequency of a disease or health event.Rates are particularly useful for comparing disease frequency in different locations, at different times or among different groups of persons with potentially different sized populations. A rate accounts well for the realistic situation in which a population is changing over time. Rates are more flexible, exact and capture the reality of changing population. The formulate for calculating rate is

$$\textbf{Rate} = \frac{\textbf{\# New or Incident cases}}{\textbf{Person-Time}}$$

A rate is made up of the following: numerator, denominator, time specification and multiplier.

Person-Time is the sum of time that a person is at risk for the health outcome and under study observation. A good example of a typical rate is Death Rate.

$$\textbf{Death Rate} = \frac{\text{Number of Deaths in one Year}}{\text{Mid-Year population}} \times \text{Constant multiplier (e.g. } 10^n\text{)}$$

The time is usually calendar year. The rate is usually expressed per 1000 or 10,000, 100,000 etc. The three major categories of rates are:

i. **Crude Rates**: The actual observed rates such as the birth and death rates. Crude rates are also known as unstandardized rates.

ii. **Specific Rates**: These are the actual observed rates due to specific causes (e.g. COVID, Tuberculosis, Malaria, maternal, under-five) or occurring in specific groups (e.g. age-sex groups) or during specific time periods (e.g. annual, monthly or weekly rates)

iii. **Standardized Rates**: These are obtained by direct or indirect method of standardization or adjustment, (e.g. age and sex standardized rates)

b) Ratio

A ratio is a relationship in size between two random quantities. The numerator is not part of the denominator. It is expressed in the form of:

a: b or a/b

A ratio is a relative magnitude of two values. Examples include:

- Sex-ratio, Doctor-Patient ratio, child-woman ratio.

c) Proportion

A proportion involves comparing a part to a whole. It is a special ratio in which the numerator is part of the denominator. A proportion is usually expressed as a percentage.

Proportion = a/a+b

It is often expressed as a percentage

2.5.1 Measurement of Mortality

Mortality is defined as the demographic event of death(Zodpey, 2015). Most epidemiological studies begin with mortality data. Analysis is made easier if there is good quality data on deaths and population. These data are obtained from vital registration systems and population census respectively(Friis and Sellers, 2013).

Uses of Mortality Data

Mortality statistics are a mirror of the situation of health of the population. It depicts the kind of health problems in the population that lead to death(Alderson, 1983). Data on causes of death are important and used widely for a variety of purposes including explaining trends and differentials in overall mortality, pinpointing health priorities for action and the allocation of resources in designing interventions. Further, it helps in the assessment, monitoring and evaluation of public health problems and programs.

1. *Crude Death Rate*

The crude death rate is calculated by dividing the number of registered deaths in a year by the mid-year population for the same year. The rate is expressed as per 1,000 population.

$$\textbf{Crude Death rate} = \frac{\textbf{Total number of deaths}}{\textbf{Total mid-year population}} \times 1{,}000$$

This rate gives the number of deaths that occur, on the average, per 1,000 people in the community. Further, it is quite easy to compute and requires only the total population size and the total number of deaths. It represents an estimate of the chance of dying for a person belonging to the given population, because the whole population is supposedly exposed to the risk of dying of something or the other. However, the CDR is associated with some drawbacks. The CDR assumes that the chance of dying

- is the same for the different age groups and for different genders
- is the same with respect to race, occupation or location

Crude death rate summarises effects of two variables:

1. composition of the population
2. age-specific death rates

2. *Specific Death Rate*

The crude death rates for specific causes of death are calculated in a similar way by selecting deaths due to specific cause as the numerator and mid-year population as the denominator. Thus,

$$\textbf{Cause-specific death rate} = \frac{\textbf{Total number of deaths due to some specific cause}}{\textbf{Total mid-year population}}$$

Examples:

i. Specific death rate due to Tuberculosis $= \frac{\text{Number of deaths from Tuberculosis during calendar year}}{\text{Mid-Year Population}} \times 1{,}000$

ii. Specific death rate for males= $\frac{\text{Number of deaths among males during calendar year}}{\text{Mid-year population of males}} \times 1,000$

iii. Specific DR in age grp 15-20yrs= $\frac{\text{Deaths of persons aged 15-20 during a calendar year}}{\text{Mid-year population of persons aged 15-20}} \times 1,000$

iv. Death rate for December = $\frac{\text{Deaths in January} \times 12}{\text{Mid-year population}} \times 1,000$

v. Weekly Death Rate = $\frac{\text{Deaths in the week} \times 52}{\text{Mid-Year population}} \times 1,000$

The deaths are multiplied by 12 and 52 respectively in order to make the monthly and weekly death rates comparable with the annual death rate

i. Case Fatality Rate (Ratio)

CFR= $\frac{\text{Total number of deaths due to a particular disease}}{\text{Total number of cases due to the same disease}} \times 1,000$

CFR represents the killing power of a disease. It is simply the **ratio** of deaths to **cases**. The time interval is not specified. CFR is typically used in acute infectious diseases (e.g. COVID-19, cholera, measles). It is closely related to virulence and may be used to assess the effectiveness of the intervention such timely and proper patient treatment.

ii. Proportional Mortality Rate (Ratio)-PMR

Sometimes it is desirable to know what proportion of the total deaths are due to a particular cause (e.g. cancer) or even what proportion of deaths are occurring in a particular age group (e.g. above 60 years).

a. **PMR from Specific disease**= $\frac{\text{Number of deaths from specific disease in a year}}{\text{Total deaths from all causes in that year}} \times 100$

b. **PMR for aged ≥60 years** = $\frac{\text{Number of deaths of persons aged 60 yrs and above}}{\text{Total deaths of all age groups in that year}} \times 100$

PMR is usually used for a broad disease group (e.g. communicable disease as a whole) and also for a specific disease of major public health importance e.g. cancer, HIV

iii. Survival Rate

This is a proportion of survivors in a group (e.g. of patients) studied and followed over a period (e.g. a 5-year period). It is a method of describing prognosis in certain disease conditions. Survival experience can be used as a yardstick for the assessment of standards of therapy.

Survival Rate= $\frac{\text{Total number of patients alive after 5 years'}}{\text{Total number of patients diagnosed or treated}} \times 100$

iv. Age Adjusted Mortality Rates or Standardization

Age adjusted mortality rate or Standardized Mortality Rate (SMR) helps to compare mortality rates between different populations. The Crude Death Rate is not the right method to use for this. This is due to the fact that rates are only comparable if the populations upon which they are based are comparable. The age adjustment or standardization removes the confounding effect of different age structures and yields a single standardized or adjusted rate, by which the mortality experience can be compared. The adjustment can also be made for sex, race, parity etc. Standardization is carried out by one of two methods:

- Direct Standardization

- Indirect standardization.

Both methods begin by choosing a standard population.

2.5.2 Measurement of Morbidity

Morbidity is one of the two measures commonly used in epidemiology, the other one being mortality. Morbidity comes from the Latin word “morbidus” which means sick or unhealthy) and refers to a diseased state, disability, or poor health due to any cause. It has been defined as any departure, subjective or objective, from a state of physiological or psychological well-being. According to the World Health Organization (WHO) Expert Committee on Health Statistics, morbidity can be measured by:

1) Persons who are/were ill
2) The illness that these people experienced
3) The duration (days, weeks, etc.) of ill health

Morbidity commonly measures three aspects and these are frequency, duration of illness and severity.

- **Incidence** and **Prevalence rates** are used to measure frequency
- Duration of illness is measured by the **disability** rate
- Severity is measured by the **Case Fatality** Rate

Importance of Morbidity

- Morbidity describes the distribution and nature of the disease common in the community, and thus helps in the development and setting up of priorities
- Provides accurate in-depth information about the disease and patient characteristics
- Serves as a good starting point for studying causes of diseases and health problems including being key in designing prevention programs
- Morbidity is important for monitoring and evaluation of prevention and control activities

a. Incidence

Incidence refers to the occurrence of new cases of disease or injury in a population over a specified period of time. It is given by the formula:

$$\textbf{Incidence} = \frac{\text{Number of new cases of specified disease during a given time period}}{\text{Population at-risk during that period}} \times 1{,}000$$

Incidence rate must include the unit of time used in the final expression. It measures the rate at which new cases are occurring in a population. It is not influenced by the duration of the disease. The use of incidence is generally restricted to acute conditions.

b. Prevalence

Prevalence is the proportion of persons in a population who have a particular disease or attribute at a specified point in time or over a specified period of time. It is sometimes referred to as Prevalence rate. Prevalence differs from incidence in that prevalence includes all cases, both new and pre-existing, in the population at the specified time, whereas incidence is limited to new cases only. Prevalence is of two types:

1. Point prevalence refers to the prevalence measured at a particular point in time. It measures the frequency of all current cases (old & new) existing during a defined period of time (e.g., annual prevalence) expressed in relation to a defined population.

$$\textbf{Period Prevalence} = \frac{\underline{\text{No of all current cases (old+new) of a specified disease existing at a given point in time}}}{\text{Estimated population at the same point in time}} \times 100$$

2. Period prevalence refers to prevalence measured over an interval of time. It is the proportion of persons with a particular disease or attribute at any time during the interval.

$$\textbf{Period Prevalence} = \frac{\underline{\text{No. of existing cases (old+new) of a specified disease per given period of time interval}}}{\text{Estimated mid-interval population at risk}} \mathbf{\times 100}$$

Uses of Prevalence

- Prevalence helps to estimate the magnitude of health/disease problems in the community, and identify potential high risk populations
- Prevalence rates are especially useful for administrative and planning purposes.

2.6 Types of Epidemiology

There are two major types of epidemiology

2.6.1 Descriptive Epidemiology

This defines the frequency and distribution of diseases and other health related events. Descriptive Epidemiology answers four major questions of: how many (numbers), who is affected(**Person**), where is the health event (**Place**) and When did it happen (**Time**). Descriptive studies form the first stage of an epidemiological investigation. They are concerned with observing the distribution of disease or health-related characteristics in human population and identifying the characteristics with which the disease in question seems to be associated. The primary considerations for descriptive epidemiology are frequency and pattern. Frequency evaluates the rate of occurrence, and pattern helps analytical epidemiologists suggest risk factors. Descriptive epidemiology evaluates frequency and pattern by examining the person, place, and time in relationship to health events.

2.6.2 Analytic Epidemiology

Analytic epidemiology analyses determinants of health problems. It tends to answer two major questions of how and why? Analytical epidemiology's main purpose is to test hypotheses which helps to ascertain:

- Existence of a statistical association between the health event and a suspected factor and
- The strength of the association

To do this, two types of analytical studies are conducted. These are Case-Control and Cohort studies.

2.7 Epidemiological Studies

Epidemiologic studies form the foundation for disease control and prevention through understanding disease prevalence, characterizing the natural history, and identifying risk factors or causes of the disease. They help generate evidence for effective preventive and control measures. There are two main types of epidemiological studies(Brownson, 2011). These are:

2.7.1 Observational studies- the investigator does not in any way interfere with the factors being studied but is only concerned with observing:

- What is currently happening – **Cross sectional and prevalence studies**
- What happened in the past – **Case-Control studies and Retrospective cohort**
- What will happen in the future – **Prospective cohort**

2.7.1.1 Cross Sectional Studies

These are also called descriptive studies, surveys and prevalence studies. They are used to describe the general disease pattern or any health event in populations. Measurements are made on a population at a particular point in time (snap shot). The strengths and weaknesses of cross sectional studies are explained in Table 1.

2.7.1.2 Case Control Studies

Case control studies compare a group of people with disease (cases) with one or more groups without the disease (controls) going backwards in time to identify the possible exposures. The sampling starts with cases and then controls. The exposure status is then assessed by going back in time. The measure of association is called the Odds Ratio (OR). OR simply compares the occurrence of an outcome in the exposed group and occurrence of an outcome in the unexposed group. Refer to Table 1 for strengths and weaknesses.

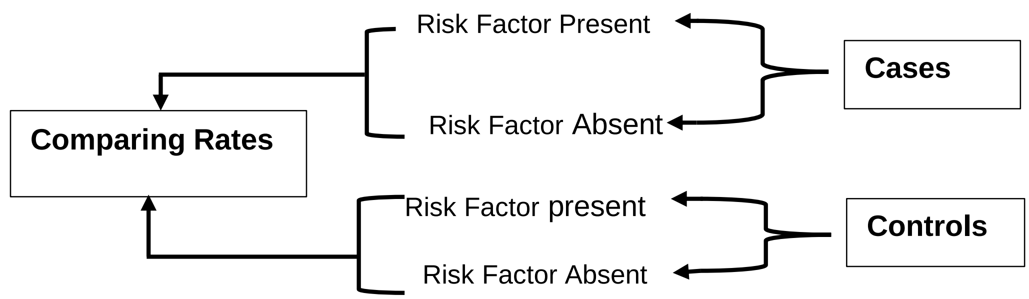

Fig 2: Illustrating backward direction in case control studies

2.7.1.3 Cohort Studies

These are the strongest of all observational studies. This design, measures and compares the incidence of disease in two or more study groups (Cohorts). In cohort studies there is often exposed and unexposed cohorts. The exposed cohort is the group that has been exposed to some event or condition while the unexposed one is the group that has had no exposure to the event or condition. These studies are prospective or forward looking. They are also called longitudinal studies.

Cohorts without disease are used as the basis for exposure status and then are followed up for a defined period of time. This follow up depends on the natural history of the outcome disease or condition and how rare that outcome is. New cases of the outcome disease are picked up during follow up and the incidence of the disease is calculated on the basis of the exposure status. The incidence in the exposed cohort is then compared with the incidence in the unexposed cohort. This ratio is called Relative Risk (RR) or Risk Ratio.

$$\textbf{RR} = \frac{\text{Incidence in the exposed cohort}}{\text{Incidence in the unexposed cohort}}$$

The RR is a measure of association between the exposure and outcome. The larger the RR, the stronger the association. This design helps to estimate the true incidence of a disease.

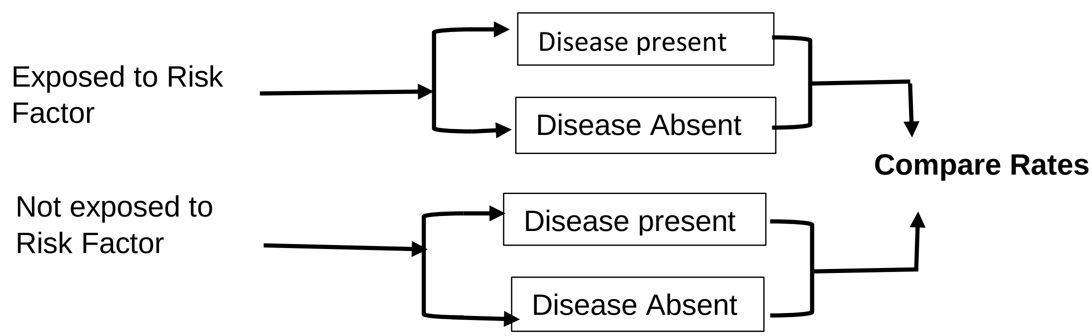

Fig 3: Illustrating cohort studies

2.7.2 ***Experimental (Intervention) studies***- the investigator manipulates the studied exposure. This type of study has ethical implications and is costly. The ethical are particularly pronounced when dealing with human studies.

Owing to the ethical considerations and the cost involved in experimental studies, most studies conducted are observational in nature. There are two types of experimental studies:

i. Clinical Trials also called Randomized Controlled Trials (RCT)

ii. Community Intervention Trials

Randomized Controlled Trials

RCT is the gold standard study design. The subjects are usually chosen from a large number of potential subjects. Sampling includes the use of a set of inclusion and exclusion criteria. Randomization is then done to allocate subjects to either the treatment or the placebo group. Randomization ensures that:

a. Allocation to different groups (treatment or placebo) is done without bias
b. Distributes known and unknown confounders equally between the two groups.

Once randomization is done, intervention can begin.

These studies are illustrated in figure 4 below.

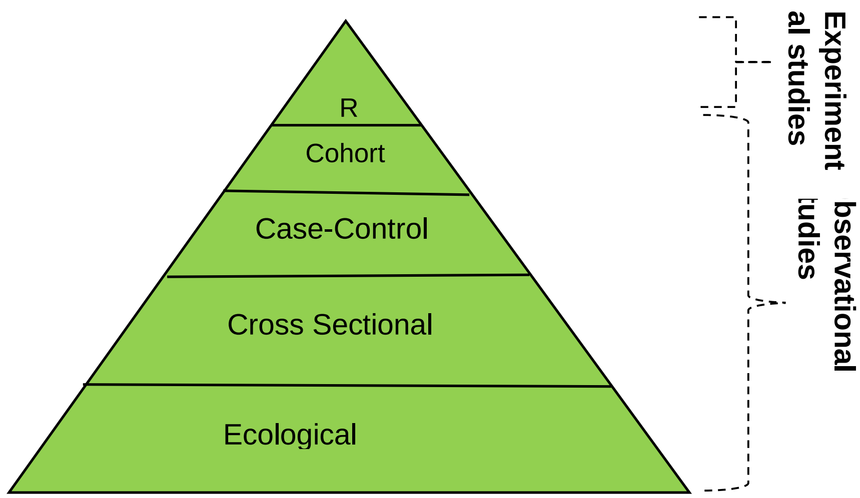

Fig 4: Types of Epidemiological study designs

Table 1: Describing Epidemiological study designs, purpose, strengths and weaknesses

Study Design	Main idea	Purpose	Strengths	Weaknesses
Cross-sectional	Study health & disease conditions in a population at a defined place and time	Measure prevalence of disease, seek associations between disease and related factors	-Easy to do -economical -yield useful data on prevalence needed to assess health situation of studied population -first step of a well-defined epidemiological study	-Causal associations cannot be made -Weakest epidemiological design
Case Control	Look for differences & similarities between cases and non-cases	Seek associations, Generate/test hypotheses	-Simpler & easier to do -Cost and time efficient -Best for investigating aetiology of rare diseases or events	-Occurrence of several types of bias -Challenges of unreliable data
Cohort	Follow-up populations relating to information on risk factor patterns and health states to the outcomes of interest	Study natural history of disease, measure incidence of disease, link disease outcome to possible disease causes; generate/test hypotheses	- Very strong designs - Greater assurance that exposure preceded outcome - Certain exposures can be studied for multiple outcomes at same time	Time consuming
Trial	Intervene with some measure designed to improve health, then follow up people to see the effect	Test understanding of causes, study how to influence natural history of disease, evaluate the benefits and costs of interventions	-Gives researchers a high level of control -Allows researchers to utilize many variations -Can lead to excellent results -Can be used in different fields	-Can take a lot of time and money -Can be affected by errors -Might not be feasible in some situations

3.0 Conclusion

Epidemiology is the fundamental discipline that supports public health practice. It is the cornerstone of public health, and shapes policy decisions and evidence-based practice by identifying risk factors for diseases and health related states for the purpose of effective control. Epidemiology as an interdisciplinary field drawn from biostatistics and social and behavioural sciences as well as medical related fields of toxicology, pathology, virology, genetics, microbiology and clinical medicine.

Epidemiology helps to explain and illustrate the key concepts which underpin its applications to research, policy making, health service planning and health promotion. Disease patterns in populations arise from the differences in the type of individuals they comprise, mode of interaction of the individuals and the environment in which the population lives.

Interactions among the individuals and the environment in specified populations helps us to understand the cause and effect relationship between the agent, host and environment. This is best done through effective data collection approaches, accurate measurements and analysis. The ultimate aim of epidemiology is to prove that a link exists between a disease and its cause and use this to improve interventions for prevention and control of diseases and other health related events.

4.0 Bibliography

1. Alderson, M. (1983) 'The Use of Mortality Statistics', *An Introduction to Epidemiology*, pp. 8–44. doi: 10.1007/978-1-349-17291-7_2.
2. Brownson, R. C. (2011) 'Epidemiology in Public Health Practice: By A. Haveman-Nies, S. C. Jansen, J. A. M. van Oers, and P. van 't Veer', *American Journal of Epidemiology*, 174(7), pp. 871–872. doi: 10.1093/aje/kwr255.
3. Decarli, A. (2003) 'Concepts of epidemiology: an integrated introduction to the ideas, theories, principles and methods of epidemiology: R Bhopal. (Pp 317; price not stated). Oxford University Press, 2002. ISBN 0-19-263155-1', *Journal of Epidemiology & Community Health*, 57(8), pp. 639–639. doi: 1136/jech.57.8.639.
4. Frérot, M. *et al.* (2018) 'What is epidemiology? Changing definitions of epidemiology 1978-2017', *PLoS ONE*, 13(12). doi: 0.1371/journal.pone.0208442.
5. Friis, R. and Sellers, T. (2013) 'Measures of Morbidity and Mortality Used in Epidemiology', *Epidemiology for Public Health Practice*, p. 50. Available at: http://samples.jbpub.com/9781284103717/9781449651589_CH03_Friis.pdf.
6. Kundi, M. (2006) 'Causality and the interpretation of epidemiologic evidence', *Environmental Health Perspectives*, 114(7), pp. 969–974. doi: 10.1289/ehp.8297.
7. Last, J. M. (1993) *Dictionary of epidemiology [1]*, *Journal of Epidemiology and Community Health*. doi: 10.1136/jech.47.5.430.
8. Zodpey, S. P. (2015) 'Principles of epidemiology and epidemiologic methods', *Public Health and Nutrition in Developing Countries Part-I*, pp. 1–24. doi: 10.1533/9780857093905.1.

.