COURSE GUIDE

PHS803
PRINCIPLES OF EPIDEMIOLOGY/EPIDEMIOLOGY

PHS 803 Fundamental Principles of Epidemiology and Disease Control

Course Team Dr. Mohammed Jimoh Saka, Dr. Aishat Oluwatoyin Saka (Uni-Ilorin) & Dr Oluwatosin Jonadab Akinsola – FUT Owerri – (Course Developer/Writers)
Professor Ukwuoma Abanobi M(Course Editor) – FUT, Owerri
Professor Grace Okoli (Programme Coordinator)- NOUN

NATIONAL OPEN UNIVERSITY OF NIGERIA

National Open University of Nigeria
Headquarters
University Village
Plot 91, Cadastral Zone NnamdiAzikiwe Expressway
Jabi, Abuja

Lagos Office
14/16 Ahmadu Bello Way
Victoria Island, Lagos

e-mail: centralinfo@nou.edu.ng
URL: www.nou.edu.ng

Printed 2021

ISBN:

CONTENTS **PAGE**

INTRODUCTION

PHS803: Epidemiology is a two-credit compulsory course for all students offering Masters of Public Health Science. The health worker must be trained in the same way and work with the same methods as his colleagues who specialise in other areas of Health Sciences. As with all scientific endeavours, the practice of epidemiology relies on a systematic approach. In very simple terms, the epidemiologist: **Counts** cases of a phenomenon or health events, and describes them in terms of time, place, and person variables; **Divides** the number of cases a phenomenon or health event by an appropriate denominator to calculate rates, ratios or proportions (as the case may be) to derive appropriate indictors; and **Compares** these indicators over time or for different groups of people to get relevant interpretations about the occurrence of the referent phenomenon or health event. Application is used in prevention and control of diseases (or any other related health status risks) and ultimately improving the health systems of the country.

WHAT YOU WILL LEARN IN THIS COURSE

In this course, you have the course units and a course guide. The course guide will tell you what the course is all about. It is the general overview of the course materials you will be using and how to use those materials. It also helps you to allocate the appropriate time to each unit so that you can successfully complete the course within the stipulated time limit. The course guide also helps you to know how to go about your Tutor-Marked Assignments (TMAs) which will form part of your overall assessment at the end of the course. Also, there will be regular tutorial classes that are related to this course, where you can interact with your facilitators and other students. Please, I encourage you to attend these tutorial classes.

COURSE AIMS

The course aims to give you detail understanding of Principle of Epidemiology which is an important branch of Public Health.

COURSE OBJECTIVES

To achieve the aims set above, each course unit has a set of specific objectives which are included at the beginning of the unit. These objectives will give you what to concentrate / focus on while studying the unit. Please read the objectives before studying the unit and during your study to check your progress.

Below are the comprehensive objectives of the course as a whole. By meeting these objectives, you should have achieved the aims of the course as a whole. Thus, after going through the course, you should be able to:

- explain the meaning of epidemiology
- describe the fundamental assumptions of epidemiology
- explain different models of disease causation
- discuss the importance of "epidemiology principle"
- define a case definition and explain components of a case or outbreak investigations
- explain COVID -19 case of corona virus; - Surveillance case definition 2020
- identify criteria in case definitions using COVID-19 and Measles as example
- discuss usefulness of counts and rates
- explain variation in case definition
- describe measure of association and risk of diseases in term of attack rate etc.
- explain surveillance systems in Nigeria
- identify Integrated Diseases Surveillance Response (IDSR)
- describe how IDSR contributes to COVID-19 pandemic preparedness & response
- give detailed account of guidelines provision before and during an outbreak
- discuss the role of Village, Health Facility and LGA Level Services
- Discuss core capacity requirements for surveillance and response under IHR.
- define and explain bias (systematic error)
- differentiate types of bias
- discuss selection bias in (Case Control study designs, Cohort Study designs)
- describe "Self Control", "Differential Surveillance", "Referral" or "Diagnosis of subjects" in selection bias
- describe "loss to follow up", "The Healthy Worker Effects", ""Subject Selection Bias in Cohort study designs
- explain differential and non-differential misclassification of disease and exposure and mechanism in information bias
- Understand information bias on recall, interviewer and differences in quality of information
- describe information bias on misclassification of outcome (differential and non-differential)
- identify effect of decreased sensitivity and specificity of detecting diseases subject in non- differential misclassification of information bias

- explain the meaning and conditions that must be present for confounding to occur
- give detailed account of magnitude and method of determine confounding
- identify “residual confounding”, “confounding by indication”, & “reverse causality”
- describe how to control confounding in a study
- describe how to carry out direct standardisation procedure
- describe how to carry out indirect standardisation procedure
- identify the underlying issues in the use of standardisation
- explain the role of descriptive studies for identifying problems and establishing hypotheses
- identify observational studies design in term of case series, cross sectional, cohort studies and ecological studies
- identify and explain experimental studies design
- explain the functions, issues and clinical significance of studies design
- describe rationale for selecting a study design
- discuss the strength and limitation of descriptive studies
- discuss the strength and limitation of analytical studies
- discuss the strength and limitation of experimental studies
- identify the different classes of variables (discrete [dichotomous, categorical, ordinal], continuous, time to event)
- distinguish when to use mean and standard deviation versus median and interquartile range (IQR) to characterise the center and variability for continuous variables data
- use R to compute mean, variance, standard deviation, median, and interquartile range (IQR)
- use R to compute the correlation coefficient for an ecological study
- conduct a narrative case series, present in an abstract format and put in an appropriate table for interpretation
- analysing a cross sectional survey
- computing the Correlation Coefficient
- carry out description and analysis of ecological studies
- conduct calculation of correlation and linear regression using appropriate formula.

WORKING THROUGH THIS COURSE

To complete this course, you are required to read each study unit, read the textbooks and read other materials which may be provided by the National Open University of Nigeria.

Each unit contains self-assessment exercise and at certain points in the course you would be required to submit assignments for assessment purposes. At the end of the course there is a final examination. The course should take you about a total of 12 weeks to complete. Below you will find listed all the components of the course, what you have to do and how you should allocate your time to each unit in order to complete the course on time and successfully.

This course entails that you spend a lot of time to read. We would advise that you avail yourself the opportunity of attending the tutorial sessions where you have the opportunity of comparing your knowledge with that of other people.

THE COURSE MATERIALS

The main components of the course are

1. The course Guide
2. Study chapters
3. References/Further Reading
4. Assignments
5. Presentation Schedule

STUDY UNITS

The study units for this course are made up of six (6) modules and eighteen (18) units as given below:

Module 1 Basic Principles and Methods of Epidemiology

Unit 1 Principle of Epidemiology on Disease Causation
Unit 2 Disease Measures and The Epidemiologic Approach
Unit 3 Integrated Disease Surveillance and Response in Nigeria

Module 2 Systemic Error (Bias), Confounding and Standardisation

Unit 1 Bias in Principle of Epidemiology
Unit 2 Confounding in Principle of Epidemiology
Unit 3 Measure of Standardisation in Epidemiology

Module 3 Valid and Efficient Epidemiologic Studies the Types, Strength & Limitation and Interpreting Results

Unit 1 Designing Valid and Efficient Epidemiologic Studies
Unit 2 Strength and Limitation of Epidemiological Design
Unit 3 Interpret Epidemiologic Results

Module 4 Observational Epidemiology

Unit 1 Descriptive Epidemiology
Unit 2 Analytic Epidemiology I: Case Control Study
Unit 3 Analytic Epidemiology II: Cohort Study

Module 5 Experimental, Screening and Investigative Epidemiology (Interventional Studies in Epidemiology)

Unit 1 Experimental Epidemiology
Unit 2 Screening in Epidemiology
Unit 3 Investigation of Disease Outbreaks

Module 6 Basic Demographic Methods

Unit 1 Introduction to Demography
Unit 2 Measures of Fertility and Mortality
Unit 3 Investigation of Epidemics

Module 1 The first module discusses Principle of Epidemiology on disease causation, it also describes Disease Measures and The Epidemiologic Approach, Using the outbreak of communicable disease. It includes an explanation on Integrated Disease Surveillance and Response in Nigeria. The nodule provides detail understanding of case definition in-terms of suspected, probable and confirmed cases and illustrates them, using the COVID-19 as an example. The module also explains the distinction between integration, coordination and collaboration in case management of disease for prevention, control and eradication. The importance of reporting systems and dissemination of data from source to the National level.

Module 2 examines systemic error (bias), confounding and standardisation. It reviews different type of bias and how to resolve such bias in epidemiologic studies. Unit two critically looks at the sources of confounding and various methods to correct them. The module ends with explanation of measure of standardisation in epidemiology

Module 3: The Unit One of this module focuses on designing valid and efficient different type of epidemiologic studies. Unit Two addresses the strength and limitation of epidemiological design while Unit Three is centered on the applications epidemiologic results to develop hypothesis analysis, presentation and interpretation. There are activities related to the lecture in each study unit which will help your progress and for better comprehension of the unit. You are required to work on these exercises which together with the TMAs will enable you to achieve the objectives of each unit.

Module 4 is concerned with understanding the concepts of observational epidemiology.

Unit one discuss descriptive epidemiology with respect to person, place and time. Unit two first aspect of observational analytic epidemiology termed case-control study. It also enumerates the steps and bias in conducting the study. Unit 3 covers the second aspect of observational analytic study termed cohort study. It also provides the procedure for conducting the study and other estimates that can be derived thereof.

Module 5 examines interventional studies in (experimental) epidemiology, screening and investigation of disease outbreaks. Unit one dissects the types which are divided into animal and human experiments. It also expatiates on the distinction between randomised and non-randomised control trials, as well as describes the basic principles of community trials. Unit two discusses the process and screening procedures critically differentiate between screening tests and diagnostic tests.

Module 6 focuses on basic demographic methods in population studies. Unit one discusses population census, vital statistics and morbidity indicators. Unit two explains basic measures of fertility and mortality. Unit three is centred on standardisation of rates and further distinguish between direct and indirect standardisation.

ASSIGNMENT FILE

There are three types of assessments in this course. First are the Tutor-Marked Assessments (TMAs); second is the Self Assessed Exercises while the third is written examination. In solving the questions in the assignments, you are expected to apply the information, knowledge and experience acquired during the course. The assignments must be submitted to your facilitator for formal assessment in accordance with prescribed deadlines stated in the assignment file.
The work you submit to your facilitator for assessment accounts for 30 percent of your total course mark. At the end of the course, you will be

required to sit for a final examination of 1½ hours duration at your study center. This final examination will account for 70 % of your total course mark.

SELF ASSESSED EXERCISES

References and other resources are provided. The unit directs you to work on exercises related to the required reading. In general, these exercises test you on the materials you have just covered or require you to apply it in some way and thereby assist you to evaluate your progress and to reinforce your comprehension of the material. Together with TMAs and SAEs these exercises will help you in achieving the stated learning objectives of the individual units and of the course as a whole.

PRESENTATION SCHEDULE

Your course materials have important dates for the early and timely completion and submission of your TMAs and attending tutorials. You should remember that you are required to submit all your assignments by the stipulated time and date. You should guard against falling behind in your work.

There is a time-table prepared for the early and timely completion and submission of your TMAs as well as attending the tutorial classes. You are required to submit all your assignments at the stipulated time and date. Avoid falling behind the schedule time. The presentation schedule included in this course guide provides you with important dates for completion of each e-tutor marked assignment (e-TMAs). You should therefore try to meet the deadlines.

ASSESSMENT

There are three aspects to the assessment of the course. First is made up of self-assessment exercises, second consists of the tutor-marked assignments and third is the written examination/end of course examination.

You are advised to do the exercises. In tackling the assignments, you are expected to apply information, knowledge and techniques you gathered during the course. The assignments must be submitted to your facilitator for formal assessment in accordance with the deadlines stated in the presentation schedule and the assignment file. The work you submit to your tutor for assessment will count for 30% of your total course work. At the end of the course you will need to sit for a final or end of course examination of about three-hour duration. This examination will count for 70% of your total course mark.

TUTOR-MARKED ASSIGNMENT (TMAs)

The TMA is a continuous assessment component of your course. It accounts for 30% of the total score. You will be given TMAs questions to answer and these must be answered before you are allowed to sit for the end of course examination. The TMAs would be given to you by your facilitator and returned after you have done the assignment. Assignment questions for the units in this course are contained in the assignment file. You will be able to complete your assignment from the information and material contained in your reading, references and study units. However, it is desirable in all degree levels of education to demonstrate that you have read and researched more into your references, which will give you a wider view point and may provide you with a deeper understanding of the subject.

1. Make sure that each assignment reaches your facilitator on or before the deadline given in the presentation schedule and assignment file. If for any reason you cannot complete your work on time, contact your facilitator before the assignment is due to discuss the possibility of an extension. Extension will not be granted after the due date unless there are exceptional circumstances.

2. Make sure you revise the whole course content before sitting or the examination. The self-assessment activities and TMAs will be useful for this purpose and if you have any comment please do before the examination. The end of course examination covers information from all parts of the course.

FINAL EXAMINATION AND GRADING

The end of course examination for Principle of Epidemiology and Disease Control will equal to or less than 2 hours and it has a value of 70% of Control total course work. The examination will consist of questions, which will reflect the type of self-testing, practice exercise and tutor-marked assignment problems you have previously encountered. All areas of the course will be assessed.
Use the time between finishing the last unit and sitting for the examination to revise the whole course. You might find it useful to review your self-test, TMAs and comments on them before the examination. The end of course examination covers information from all parts of the course.

COURSE MARKING SCHEME

Table 1: Course Marking Scheme

	Marks
Assignment 1 – 3	Three assignments, at 10% each = 30% of
End of Examination Course	70% of overall course marks
Total	100% of course materials

TABLE 2: COURSE ORGANISATION

Unit	**Title of Work**	**Weeks Activity**	**Assessment (End of Unit)**
Unit	**Course Guide Week**	**Week**	
1	Principle of Epidemiology on Disease Causation	Week 1	Assignment 1
2	Disease Measures and The Epidemiologic Approach	Week 2	Assignment 2
3	Integrated Disease Surveillance and Response	Week 3	Assignment 3
4	Bias in Principle of Epidemiology	Week 4	Assignment 4
5	Confounding in Principle of Epidemiology	Week 5	Assignment 5
6	Measure of Standardisation in Epidemiology	Week 6	Assignment 6
7	Designing Valid and Efficient Epidemiologic Studies	Week 7	Assignment 7
8	Strength and Limitation of Epidemiological Design	Week 8	Assignment 8
9	Interpret Epidemiologic Results	Week 9	Assignment 9

HOW TO GET THE MOST OUT OF THIS COURSE

In distance learning, the study units replace the university lecturer. This is one of the huge advantages of distance learning mode; you can read and work through specially designed study materials at your own pace and at a time and place that suit you best. Think of it as reading from the teacher, the study guide tells you what to read, when to read and the relevant texts to consult. You are provided exercises at appropriate points, just as a lecturer might give you an in-class exercise. Each of the study units follows a common format. The first item is an introduction to the subject matter of the unit and how a particular unit is integrated with the other units and the course as a whole. Next to this is a set of learning objectives. These learning objectives are meant to guide your studies. The moment a unit is finished, you must go back and check whether you have

achieved the objectives. If this is made a habit, then you will significantly improve your chances of passing the course. The main body of the units also guides you through the required readings from other sources. This will usually be either from a set book or from other sources. Self-assessment exercises are provided throughout the unit, to aid personal studies and answers are provided at the end of the unit. Working through these self-tests will help you to achieve the objectives of the unit and also prepare you for tutor marked assignments and examinations. You should attempt each self-test as you encounter them in the units.

The following are practical strategies for working through this course

1. Read the Course Guide thoroughly.
2. Organise a study schedule. Refer to the course overview for more details. Note the time you are expected to spend on each unit and how the assignment relates to the units. Important details, e.g. details of your tutorials and the date of the first day of the semester are available. You need to gather together all this information in one place such as a diary, a wall chart calendar or an organiser. Whatever method you choose, you should decide on and write in your own dates for working on each unit.
3. Once you have created your own study schedule, do everything you can to stick to it. The major reason that students fail is that they get behind with their course works. If you get into difficulties with your schedule, please let your tutor know before it is too late for help.
4. Turn to Unit 1 and read the introduction and the objectives for the unit.
5. Assemble the study materials. Information about what you need for a unit is given in the table of contents at the beginning of each unit. You will almost always need both the study unit you are working on and one of the materials recommended for further reading, on your desk at the same time.
6. Work through the unit, the content of the unit itself has been arranged to provide a sequence for you to follow. As you work through the unit, you will be encouraged to read from your set books.
7. Keep in mind that you will learn a lot by doing all your assignments carefully. They have been designed to help you meet the objectives of the course and will help you pass the examination.
8. Review the objectives of each study unit to confirm that you have achieved them. If you are not certain about any of the objectives, review the study material and consult your tutor.

9. When you are confident that you have achieved a unit's objectives, you can start on the next unit. Proceed unit by unit through the course and try to pace your study so that you can keep yourself on schedule.
10. When you have submitted an assignment to your tutor for marking, do not wait for its return before starting on the next unit. Keep to your schedule. When the assignment is returned, pay particular attention to your tutor's comments, both on the tutor marked assignment form and also that written on the assignment. Consult you tutor as soon as possible if you have any questions or problems.
11. After completing the last unit, review the course and prepare yourself for the final examination. Check that you have achieved the unit objectives (listed at the beginning of each unit) and the course objectives (listed in this course guide).

FACILITATORS/TUTORS AND TUTORIALS

There are Sixteen (16) hours of tutorials provided in support of this course. You will be notified of the dates, times and location of these tutorials as well as the name and phone number of your facilitators, as soon as you are allocated a tutorial group.

Your facilitator will mark and comment on your assignments, keep a close watch on your progress and any difficulties you might face and provide assistance to you during the course. You are expected to mail your Tutor Marked Assignment to your facilitator before the schedule date (at least two working days are required). They will be marked by your tutor and returned to you as soon as possible. Do not delay to contact your facilitator by telephone or e-mail if you need assistance.
The following might be circumstances in which you would find assistance necessary, hence you would have to contact your facilitator if:

I. You do not understand any part of the study or the assigned readings.
II. You have difficulty with the self-tests.
III. You have a question or problem with an assignment or with the grading of an assignment.

You should endeavour to attend the tutorials. This is the only chance to have face to face contact with your course facilitator and to ask questions which are answered instantly. You can raise any problem encountered in the course of your study.

To gain much benefit from course tutorials prepare a question list before attending them. You will learn a lot from participating actively in discussions.

SUMMARY

Principle of Epidemiology is a course that intends to provide, describe and apply the basic principles and methods of epidemiology including disease measures, association and causation, bias, confounding in epidemiologic investigations and effect modification and susceptibility; interpret descriptive epidemiologic results in order to develop hypotheses of possible risk factors of a disease; develop a foundation for designing valid and efficient epidemiologic studies to address public health problems including understanding: the strengths and limitations of descriptive, observational and experimental studies; integrated disease surveillance and response in Nigeria; epidemiology of Non-communicable diseases; epidemiology and control of vector-borne diseases; standardisation of rates in epidemiology; epidemiology of water-borne diseases and epidemiology of air-borne diseases.
Upon completing this course, you will be equipped with the basic knowledge of meaning of epidemiology, its usefulness and many other. In addition, you will be able to answer the following:

i. Source of epidemiological status?
ii. How will you describe Epidemiology in relation to health-related events?
iii. What are the interrelationship between Agent, Host and Environment?
iv. Give a detail account of Rothman's Causal Pies
v. What is the difference beteeen risk ratio and relative risk?
vi. Sate the formuula for Risk Ratio, Odd Ratio, using apprarite formula.
vii. What is the difference between Disease Surveillance and Integrated Disease Surveillance?
viii. List Diseases of Public Health Importance under IDSR
ix. Explain how outbreak is reported from the point to the FMoH
x. What are the relationship between IDSR and IHR?

The list of questions that you would be able to answer is not limited to the above list. To gain the most from this course you should endeavour to apply the principles you have learnt to your understanding of Public Health.

We wish you success in this course and we hope you will find it both interesting and useful!

MAIN COURSE

CONTENTS **PAGE**

MODULE 1 BASIC PRINCIPLES AND METHODS OF EPIDEMIOLOGY

UNIT 1 PRINCIPLE OF EPIDEMIOLOGY ON DISEASE CAUSATION

CONTENTS

1.0 INTRODUCTION

Epidemiology based its origin in the idea, first expressed over 2000 years ago by Hippocrates and others, that environmental factors can influence the occurrence of disease. However, it was not until the nineteenth century that the distribution of specific human population groups was measured to any great extent. This work marked not only the formal beginnings of epidemiology but also of its most important achievements; for example, the finding by John Snow that the risk of cholera in London was related, among other things, to the drinking of water supplied.

A subsequent significant land mark in the development of epidemiology is illustrated by the work of Doll, Hill and others who studied the relationship between cigarette smoking and lung cancer in the 1950s. This work which was preceded by clinical observations linked smoking to lung cancer, expanded epidemiological interest to chronic diseases.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- explain the meaning of epidemiology
- describe the fundamental assumptions of epidemiology
- explain different models of disease causation
- discuss the importance of epidemiology principle.

3.0 MAIN CONTENT

3.1 Meaning of Epidemiology

Epidemiology is a fundamental science of Public Health which is concerned with the occurrence of morbidity and mortality in populations. It is defined as the study of the distribution and determinants of health - related states and events in population as well as the application of this study to the control of health problems. Originally epidemiologic principles were employed exclusively for the control of infectious and communicable diseases. However, this concept has been broadened to include a variety of disease and health problems, including chronic and degenerative as well as psycho-behavioural diseases. Thus, Epidemiology provides both a body of knowledge and a formulation of methods for learning about health and disease status with a goal of ultimately finding solution to health problems.

3.2 Fundamental Assumptions in Epidemiology

These assumptions include the following:

i. Disease does not occur in a vacuum.
ii. Disease is not randomly distributed throughout a population.
iii. Epidemiology uses systematic approach to study the differences in disease distribution in subgroups.
iv. Allows for study of causal and preventive factors.

3.3 Causation and Concepts of Disease Occurrence

3.3.1 Causation of Diseases Model

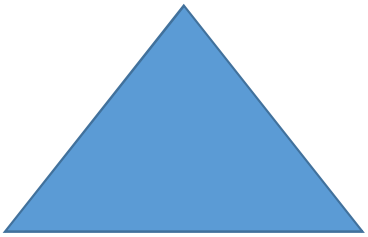
A number of models of disease causation have been proposed. Among the simplest of these is the epidemiologic triad or triangle, the traditional model for infectious disease. The triad consists of an external **agent**, a susceptible **host**, and an **environment** that brings the host and agent together. In this model, disease results from the interaction between the agent and the susceptible host in an environment that supports

transmission of the agent from a source to that host. The two models: the Epidemiologic Triad and Causal Pies Model are described below in Figures 1 and 2.

3.3.2 Epidemiologic Triad Model

Agent, host, and environmental factors interrelate in a variety of complex ways to produce disease. Different diseases require different balances and interactions of these three components. Development of appropriate, practical, and effective public health measures to control or prevent disease usually requires assessment of all three components and their interactions.

Agent



Host Environment

Figure 1 Epidemiologic Triad **Source:** Computation by Saka M.J, (2020).

i. Description of Epidemiologic Triad

Agent originally referred to an infectious microorganism or pathogen: a virus, bacterium, parasite, or other microbe. Generally, the agent must be present for disease to occur; however, presence of that agent alone is not always sufficient to cause disease. A variety of factors influence whether exposure to an organism will result in disease, including the organism's pathogenicity (ability to cause disease) and dose. Over time, the concept of agent has been broadened to include chemical and physical causes of disease or injury. These include chemical contaminants (such as the L-tryptophan contaminant responsible for eosinophilia-myalgia syndrome), as well as physical forces (such as repetitive mechanical forces associated with carpal tunnel syndrome). While the epidemiologic triad serves as a useful model for many diseases, it has proven inadequate for cardiovascular disease, cancer, and other diseases that appear to have multiple contributing causes without a single necessary one.

Host refers to the human who can get the disease. A variety of factors intrinsic to the host, sometimes called risk factors, can influence an individual's exposure, susceptibility, or response to a causative agent. Opportunities for exposure are often influenced by behaviours such as sexual practices, hygiene, and other personal choices as well as by age

and sex. Susceptibility and response to an agent are influenced by factors such as genetic composition, nutritional and immunologic status, anatomic structure, presence of disease or medications, and psychological makeup.

Environment refers to extrinsic factors that affect the agent and the opportunity for exposure. Environmental factors include physical factors such as geology and climate, biologic factors such as insects that transmit the agent, and socioeconomic factors such as crowding, sanitation, and the availability of health services.

The epidemiologic triad aptly illustrates what is also referred to as the Agent-Host-Environment Epidemiological Model of disease occurrence. This model is a useful paradigm for describing patterns of occurrence of a disease in a community. The model is based on the premise that at any time in the endemic occurrence of a disease in a population, the determinants of the disease relative to the agent, host and environment exist in a state of dynamic equilibrium in which they reciprocally influence each other.

In this tripartite relationship, environmental factors stabilise the homeostasis between the disease producing propensities of the agent and the disease-resisting propensities of the host. Any disruption of this equilibrium state will influence the incidence or prevalence of the disease in that population. Thus, any changes in the characteristics of the agent, whether of endogenous or exogenous origin, which potentiate its pathogenicity or virulence will lead to increase in incidence of disease. Equally any factors that increase the susceptibility of the host of a given disease and/or his exposure potential to the agent will lead to increased incidence of disease.

The same principle is applicable with regard to changes in the circumstances of the environments that may influence the harborage of vectors as well as the physiologic activities and overall survival potential of the host and the agent. Environmental conditions can also mediate increased exposure potential of host agent. They can also exacerbate or alternate the pathogenicity of agent and the susceptibility of the host.

This paradigm is particularly useful in the descriptive study of epidemics of disease and injuries. Consistent with changes in contemporary practice of epidemiology, the model is being increasingly applied to non-disease problems such as drug abuse, child abuse, teenage pregnancy, motor-vehicle accidents, fire-arm relate mortality etc.

3.3.3 Causal Pies Model of Diseases causation

In description of the component causes and causal pies, due to the fact that the agent-host-environment model did not work well for many non-

infectious diseases, several **other models** that attempt to account for the multifactorial nature of causation have been proposed. One such model was proposed by Rothman in 1976, and has come to be known as the **Causal Pies Model.** This model is illustrated in Figure 2. An individual factor that contributes to cause disease is shown as a piece of a pie. After all the pieces of a pie fall into place, the pie is complete — and disease occurs. The individual factors are called **component causes**. The complete pie, which might be considered a causal pathway, is called a **sufficient cause**. A disease may have more than one sufficient cause, with each sufficient cause being composed of several component causes that may or may not overlap. A component that appears in every pie or pathway is called a **necessary cause**, because without it, disease does not occur. Note in Figure 2 that component cause A is a necessary cause because it appears in every pie.

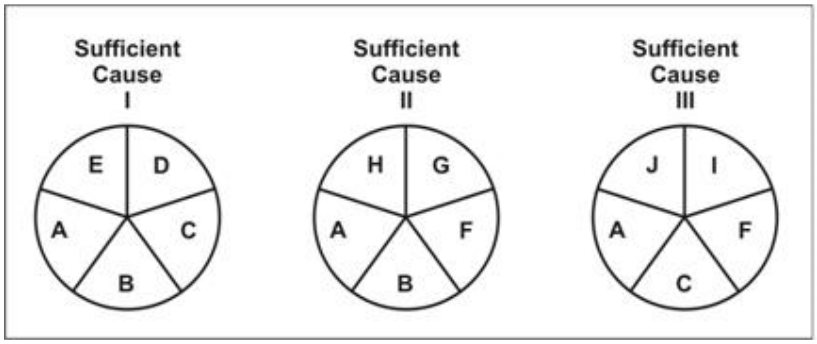


Figure 2: Rothman's Causal Pies Source: Rothman, (2010)

ii. Description of The Diagram

The component causes may include intrinsic host factors as well as the agent and the environmental factors of the agent-host-environment triad. A single component cause is rarely a sufficient cause by itself. For example, even exposure to a highly infectious agent such as measles virus does not invariably result in measles disease. Host susceptibility and other host factors also may play a role.

At the other extreme, an agent that is usually harmless in healthy persons may cause devastating disease under different conditions. Pneumocystis carinii is an organism that harmlessly colonises the respiratory tract of some healthy persons, but can cause potentially lethal pneumonia in persons whose immune systems have been weakened by human immunodeficiency virus (HIV). Presence of Pneumocystis carinii organisms is therefore a necessary but not sufficient cause of pneumocystis pneumonia. In Figure 2, it would be represented by component cause A.

As the model indicates, a particular disease may result from a variety of different sufficient causes or pathways. For example, lung cancer may result from a sufficient cause that includes smoking as a component cause. Smoking is not a sufficient cause by itself, however, because not all smokers develop lung cancer. Neither is smoking a necessary cause, because a small fraction of lung cancer victims have never smoked. Suppose Component Cause B is smoking and Component Cause C is asbestos. Sufficient Cause I includes both smoking (B) and asbestos (C). Sufficient Cause II includes smoking without asbestos, and Sufficient Cause III includes asbestos without smoking. But because lung cancer can develop in persons who have never been exposed to either smoking or asbestos, a proper model for lung cancer would have to show at least one more Sufficient Cause Pie that does not include either component B or component C.

Note that public health action does not depend on the identification of every component cause. Disease prevention can be accomplished by blocking any single component of a sufficient cause, at least through that pathway. For example, elimination of smoking (component B) would prevent lung cancer from sufficient causes I and II, although some lung cancer would still occur through sufficient cause III.

3.4 Why Study Epidemiology?

Epidemiology is therefore studied for the following reasons:

i. to describe the distribution, frequency and magnitude of a health problem;
ii. to identify the probable cause or trigger factors of the health problem, and
iii. to interpret and use the information collected above to promote health and to prevent and control diseases.

Epidemiology provides an understanding of the dynamic interrelationships to describe the occurrence, distribution and causes of disease so as to promote health and reduce diseases on populations.

SELF ASSESSED EXERCISES

What are the aims of Epidemiology?

4.0 CONCLUSION

In this unit, you learnt about the general idea on the meaning of principle of epidemiology. Also, you were introduced to Disease causation model and description factors, limitation of each model. You have been

introduced into the field of epidemiology including the fundamental reason for this study. You were also introduced to the usefulness of epidemiology.

5.0 SUMMARY

Epidemiology basically means the study of epidemics,' in more contemporary times, referring to- the systematic study of the origin and progression of health conditions within a given population. Over time, its scope has gone beyond the etiology of illnesses to focus on the cause, course and correlates of all kinds of health conditions. There are two models of the disease causation: The Epidemiology triad and Causal Pies. Epidemiology is not only about diseases but also include health related events within human population. The knowledge is very useful in addressing health related issues and problems within the population.

6.0 TUTOR-MARKED ASSIGNMENT (TMA)

1. How will you describe Epidemiology in relation to health-related events?
2. What are the interrelationship between Agent, Host and Environment?
3. Give a detailed account of Rothman's Causal Pies with related examples.

7.0 REFERENCE/FURTHER READING

Saka, M.J., Kamaldeen, A.S., Lawal, M.O., Muhammad, A.S. & Saka, A.O. (2012). Evaluation of Analgesics Usage in Pain Management among Physicians. Journal of Applied Pharmaceutical Science, 2 (06):194-198.

Kasius, R.V. (1998). The Challenge of Facts: Selected Public Health Papers of Edgar Sydenstricker. Prodist. New York, NY.

Kermak, W. O., McKendrick A. G. & McKinely., P. L. (1994). Death rates in Great Britain and Sweden: Expression of specific mortality rates as products of two factors, and some consequences thereof. Journal of Hygiene, 33(34):433-451.

Moberg, C. L. (1996). Rene Dubos: A harbinger of microbial resistance to antibiotics. Microb. Drug Resist 2(3):287-297.

Musa, O.I., Akande, T.M. & Saka, M.J. (2002). Epidemiological Investigation of Kerosene Burn Tragedy in Ilorin, Kwara.

State, Nigeria. *Sahel Medical Journal,* 5(4):186-189. Available online at http://www.ajol.info/viewissue.php?jid=73&id=382&ab=0.

Notestein, F. W. (1982). Demography in the United States: A partial account of the development of the field. Population and Development Review 8(4):651-687.

Rothman, K. J. (2010). Causes of Disease on agent host and Environment *Am J* Epidemiol. 104: 587–592.

UNIT 2 DISEASE MEASURES AND THE EPIDEMIOLOGIC APPROACH

CONTENTS

1.0 INTRODUCTION

As with all scientific endeavours, the practice of epidemiology relies on a systematic approach. In very simple terms, the epidemiologist:

- **Counts** cases or health events, and describes them in terms of time, place, and person;
- **Divides** the number of cases by an appropriate denominator to calculate rates (or as may be applicable ratios and proportions) to derive appropriate indicators; and
- **Compares** these rates over time or for different groups of people.

Before counting cases, however, the epidemiologist must decide what a case is. This is done by developing a case definition. Then, using this case definition, the epidemiologist finds and collects information about the occurrence of the disease or health event. The epidemiologist then performs descriptive epidemiology by characterising the cases collectively according to time, place, and person. To calculate the disease rate, the epidemiologist divides the number of cases by the size of the population. Finally, to determine whether this rate is greater than what one would normally expect, and if so to identify factors contributing to this increase, the epidemiologist compares the rate from this population

to the rate in an appropriate comparison group, using analytic epidemiology techniques. These epidemiologic actions are described in more detail below. Subsequent tasks, such as reporting the results and recommending how they can be used for public health action, are just as important, but are beyond the scope of this lesson.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- define "Case Definition" and explain "Components of a Case" and "Outbreak Investigations"
- explain COVID -19 Case of Corona Virus; - Surveillance Case Definition 2020
- identify Criteria in Case Definitions Using COVID-19 and Measles as Example
- discuss the usefulness of Counts and Rates
- explain variation in case definition
- describe Measure of Association and Risk of Diseases in term of Attack Rate etc.

3.0 MAIN CONTENT

3.1 Definition of Case and Disease Measure

Before counting cases, the epidemiologist must decide what to count, that is, what to call a case. For that, the epidemiologist uses a case definition. A case definition is a set of standard criteria for classifying whether a person has a particular disease, syndrome, or other health condition. Some case definitions, particularly those used for national surveillance, have been developed and adopted as national standards that ensure comparability. Use of an agreed-upon standard case definition ensures that every case is equivalent, regardless of when or where it occurred, or who identified it. Furthermore, the number of cases or rate of disease identified in one time or place can be compared with the number or rate from another time or place. For example, with a standard case definition, health officials could compare the number of cases of COVID-19 disease that occurred in African Countries and China in 2019 with the number that occurred there in 2020 Or they could compare the rate of COVID-19 cases in Nigeria in 2000 with the national rate in that same year. When everyone uses the same standard case definition and a difference is observed, the difference is likely to be real rather than the result of variation in how cases are classified.

To ensure that all health departments in the Nigeria use the same case definitions for surveillance, the National Center Disease Control (NCDC)

and State Epidemiologists and other interested parties have adopted standard case definitions for the notifiable infectious diseases. These definitions are revised as needed. In 1999, to address the need for common definitions and methods for state-level chronic disease surveillance, and NCDC adopted standard definitions for major chronic disease indicators.

Other case definitions, particularly those used in local outbreak investigations, are often tailored to the local situation. For example, a case definition developed for an outbreak of viral illness might require laboratory confirmation where such laboratory services are available, but likely would not if such services were not readily available.

3.2 Components of A Case Definition for Outbreak Investigations

A case definition consists of clinical criteria and, sometimes, limitations on time, place, and person. The clinical criteria usually include confirmatory laboratory tests, if available, or combinations of symptoms (subjective complaints), signs (objective physical findings), and other findings. Case definitions used during outbreak investigations are more likely to specify limits on time, place, and/or person than those used for surveillance. Contrast the case definition used for surveillance of COVID-19 cases at first instance (see box below) with the case definition used after three months after during an investigation of a COVID -19 outbreak in Nigeria 2020. Both the national surveillance case definition and the outbreak case definition require a clinically compatible illness and laboratory confirmation of COVID -19 from a normally sterile site, but the outbreak case definition adds restrictions on time and place, reflecting the scope of the outbreak.

3.2.1 Suspected, Probable and Confirm Using COVID -19

At the beginning of the pandemic, the following conditions were used to decide Suspected, Probable and Confirmed COVID -19 patients.

1. Suspected Case:

Any person with acute respiratory illness, (including severely ill-patients who have been hospitalised) presenting with fever, cough, difficulty in breathing **AND** who within 14 days before onset of illness has any one of the following exposures:

i. History of travel to China or other high prevalence country of State 14 days prior to symptoms onset **OR**
ii. Close contact with a confirmed case of nCoOV infection OR

iii. Exposure to healthcare facility in a country where hospital associated nCoV infections have been reported

2. Probable case:

A suspected case for whom testing for 2019-nCoV is inconclusive or for whom testing was positive on a pan-coronavirus assay.

3. Confirmed case:

Any person with laboratory confirmation of 2019nCov* infection with or without signs and symptoms.
After about 1 month, when the spread of the pandemic became more serious, these conditions were reviewed as follows:

1. Suspected Case:

Any person (including severely ill-patients) presenting with fever, cough or difficulty in breathing **AND** who within 14 days before the onset of illness had any of the following exposures:

1. History of travel to any country* with confirmed and ongoing community transmission of SARS-CoV- **OR**
2. Close contact with a confirmed case of COVID-19 OR
3. Exposure to a healthcare facility where COVID-19 case(s) have been reported

2. Probable case:

A suspect case for whom testing for COVID-19 is inconclusive or for whom testing was positive on a pan-coronavirus assay.
3. Confirmed case:

A person with laboratory confirmation of SARS-CoV-2 infection with or without signs and symptoms.

* As at 28/02/2020, countries with ongoing community transmission were China, Republic of Korea, Iran, Italy and Japan.

Clinical description

i. Laboratory criteria for diagnosis
ii. Case classification
iii. Confirmed: a clinically compatible case that is laboratory confirmed

3.2.2 Criteria of Case Definition

A case definition may have several sets of criteria, depending on how certain the diagnosis is. For example, during an investigation of a possible case or outbreak of measles, a person with a fever and rash might be classified as having a suspected, probable, or confirmed case of measles, depending on what evidence of measles is present (see box below).

Measles (Rubella) — Case Definition

Clinical description:

An illness characterised by all the following:

i. A generalised rash lasting greater than or equal to 3 days
ii. A temperature greater than or equal to 101.0°F (greater than or equal to 38.3°C)
iii. Cough, coryza, or conjunctivitis

Laboratory criteria for diagnosis:

i. Positive serologic test for measles immunoglobulin M antibody, or
ii. Significant rise in measles antibody level by any standard serologic assay, or
iii. Isolation of measles virus from a clinical specimen

Case classification

1. Suspected:

Any febrile illness accompanied by rash.

2. Probable

A case that meets the clinical case definition, has noncontributory or no serologic or virologic testing, and is not epidemiologically linked to a confirmed case.

3. Confirmed

A case that is laboratory confirmed or that meets the clinical case definition and is epidemiologically linked to a confirmed case. (A laboratory-confirmed case does not need to meet the clinical case definition.)

Comment: Confirmed cases should be reported to National Centre for Disease Control (NCDC). An imported case has its source outside the country or state. Rash onset occurs within 18 days after entering the jurisdiction, and illness cannot be linked to local transmission. Imported cases should be classified as:

i. International. A case that is imported from another country
ii. Out-of-State. A case that is imported from another state in the

Nigeria. The possibility that a patient was exposed within his or her state of residence should be excluded; therefore, the patient either must have been out of state continuously for the entire period of possible exposure (at least 7-18 days before onset of rash) or have had one of the following types of exposure while out of state: a) face-to-face contact with a person who had either a probable or confirmed case or b) attendance in the same institution as a person who had a case of measles (e.g., in a school, classroom, or day care center).

An indigenous case is defined as a case of measles that is not imported. Cases that are linked to imported cases should be classified as indigenous if the exposure to the imported case occurred in the reporting state. Any case that cannot be proved to be imported should be classified as indigenous. A case might be classified as suspected or probable while waiting for the laboratory results to become available. Once the laboratory provides the report, the case can be reclassified as either confirmed or "not a case," depending on the laboratory results. In the midst of a large outbreak of a disease caused by a known agent, some cases may be permanently classified as suspected or probable because officials may feel that running laboratory tests on every patient with a consistent clinical picture and a history of exposure (e.g., chickenpox) is unnecessary and even wasteful. Case definitions should not rely on laboratory culture results alone, since organisms are sometimes present without causing disease.

3.2.3 Variation in case definitions

Case definitions may also vary according to the purpose for classifying the occurrences of a disease. For example, health officials need to know as soon as possible if anyone has symptoms of COVID, Lasa Fever, or Measles so that they can begin planning what actions to take. For such rare but potentially severe communicable diseases, for which it is important to identify every possible case, health officials use a sensitive case definition. A sensitive case definition is one that is broad or "loose," in the hope of capturing most or all of the true cases. For example, the case definition for a suspected case of rubella (German measles) is "any generalised rash illness of acute onset." This definition is quite broad, and

would include not only all cases of rubella, but also measles, chickenpox, and rashes due to other causes such as drug allergies. Thus, while the advantage of a sensitive case definition is that it includes most or all of the true cases, the disadvantage is that it sometimes includes other illnesses as well.

On the other hand, an investigator studying the causes of a disease outbreak usually wants to be certain that any person included in a study really had the disease. That investigator will prefer a specific or "strict" case definition. For instance, in an outbreak of Salmonella Agona infection, the investigators would be more likely to identify the source of the infection if they included only persons who were confirmed to have been infected with that organism, rather than including anyone with acute diarrhea, because some persons may have had diarrhea from a different cause. In this setting, the only disadvantages of a strict case definition are the requirement that everyone with symptoms be tested and an underestimation of the total number of cases if some people with salmonellosis are not tested.

3.2.4 Using Counts and Rates

As noted, one of the basic tasks in public health is identifying and counting cases. These counts, usually derived from case reports submitted by health-care workers and laboratories to the health department, allow public health officials to determine the extent and patterns of disease occurrence by time, place, and person. They may also indicate clusters or outbreaks of disease in the community.

Counts are also valuable for health planning. For example, a health official might use counts (i.e., numbers) to plan how many infection control isolation units or doses of vaccine may be needed.
Rate: the number of cases divided by the size of the population per unit of time

However, simple counts do not provide all the information a health department needs. For some purposes, the counts must be put into context, based on the population in which they arose. Rates are measures that relate the numbers of cases during a certain period of time (usually per year) to the size of the population in which they occurred.

For example, 108,943 confirmed new cases of COVID 19 cases were reported in Nigeria as at 17th January 2012. This number, divided by the estimated 2021 population, results in a rate of 1.09 cases per 100,000 population. Rates are particularly useful for comparing the frequency of disease in different locations whose populations differ in size. For example, on17th January 2021, Lagos State had 38,723 over twelve times

as many cases (38,723) as its Kano state (2,577). However, Kano state has nearly three times the population of Lagos State. So, a more-fair way to compare is to calculate rates.

Rates are also useful for comparing disease occurrence during different periods of time. For example, 17.5 cases of measles per 100,000 were reported in 2017 compared with 155.8 cases per 100,000 in 2001. In addition, rates of disease among different subgroups can be compared to identify those at increased risk of disease. These so-called high-risk groups can be further assessed and targeted for special intervention. High risk groups can also be studied to identify risk factors that cause them to have increased risk of disease. While some risk factors such as age and family history of breast cancer may not be modifiable, others, such as smoking and unsafe sexual practices, are. Individuals can use knowledge of the modifiable risk factors to guide decisions about behaviors that influence their health.

3.3 Measure of Association and Risk of Diseases

The key to epidemiologic analysis is comparison. Occasionally you might observe an incidence rate among a population that seems high and wonder whether it is actually higher than what should be expected based on, say, the incidence rates in other communities. Or, you might observe that, among a group of case-patients in an outbreak, several reports having eaten at a particular restaurant. Is the restaurant just a popular one, or have more case-patients eaten there than would be expected? The way to address that concern is by comparing the observed group with another group that represents the expected level.

A measure of association quantifies the relationship between exposure and disease occurrence among the two groups. Exposure does not only mean exposure to foods, mosquitoes, a partner with a sexually transmissible disease, or a toxic waste dump, but also inherent characteristics of persons (for example, age, race, sex), acquired characteristics (marital status), biologic characteristics (immune status), activities (occupation, leisure activities) and conditions under which they live (socioeconomic status or access to medical care).
The measures of association described in the following section compare disease occurrence among one group with disease occurrence and another group. Examples of measures of association include risk ratio (relative risk), rate ratio, odds ratio, and proportionate mortality ratio.

3.3.1 Risk Ratio Definition and Method for Calculation

I. Definition of Risk Ratio

A risk ratio (RR) is also called relative risk. It compares the risk of a health event (injury, disease, risk factor, or death) among one group with the risk among another group. It does so by dividing the risk (incidence proportion, attack rate) in the group 1 exposed to the risk factor for the health event or outcome by the risk (incidence proportion, attack rate) in the group II that was not exposed to the risk factor. The two groups normally are differentiated by demographic factors such as sex (e.g., females versus males) or by exposure to a suspected risk factor (e.g., did or did not consume vegetable soup). Usually, the group of primary interest that had contact with the causative agent is labelled the exposed group while the comparison group that had no contact is labelled the unexposed group.

II. Method for Calculating Risk Ratio and Attack Rate

The formula for risk ratio or relative risk (RR) is:

$$\textbf{Risk Ratio} = \frac{\text{Risk of disease (incidence proportion, attack rate) in group of primary interest } \textit{(Exposed persons)}}{\text{Risk of disease (incidence proportion, attack rate) in comparison group } \textit{(Non-exposed persons)}}$$

- A risk ratio of 1.0 indicates identical risk among the two groups.
- A risk ratio greater than 1.0 indicates an increased risk for the group in the numerator, usually the exposed group.
- A risk ratio less than 1.0 indicates a decreased risk for the exposed group, indicating that perhaps exposure actually protects against disease occurrence.

EXAMPLES: Calculating Risk Ratios

Example 1: In an outbreak of cough among prison inmates in Oko State in 2006, 29 of 160 inmates residing in dormitory A developed cough, compared with 5 of 140 inmates residing in dormitory B. These data are summarised in a two-by-two table in Table 1. A two-by-two table has two rows for the exposure and two columns for the outcome. Here is the general format and notation.

Table 1: General Format and Notation for a Two-by-Two Table

	Ill	**Well**	**Total**
Exposed	A	B	$a + b = H_1$
Unexposed	C	D	$c + d = H_0$
Total	$a + c = V_1$	$b + d = V_0$	T

In this example, the exposure is the dormitory wing and the outcome is cough) illustrated in Table 2. Calculate the risk ratio.

Table 2: Incidence of cough Infection among prison inmates in Oko State in 2006

	Developed Cough?		
	Yes	**No**	**Total**
Isolation Center A	a = 29	b = 129	$H_1 = 158$
Isolation Center B	c = 6	d = 129	$H_0 = 135$
Total	35	258	T = 293

To calculate the risk ratio, first calculate **the risk or attack rate** for each group. Here are the formulas:
Attack rate for Exposed = a/a+b
Attack rate for unexposed = c/c+d
For this example:

Risk of cough in dormitory A inmates = 29/158 = 0.1835 = 18.3 %
Risk of cough in dormitory B inmates = 6/135 = 0.0444 = 4.4 %
The risk ratio is simply the ratio of these two risks:

Risk ratio = 0.1835/0.0444 = 4.132

The risk ratio is more than 1 so the inmates in dormitory A were 4.1 times as likely to develop cough as those inmates in dormitory B to developed cough. They are more at risk.

Example 2: Similarly, in an outbreak of COVID-19 infection among congregated, corona virus infected person in isolation center in Eke State in 2020, 30 of 160 infected persons residing in isolation center A developed Covid-19, compared with 5 of 140 infected persons residing in isolation center B. Calculate the relative risk.

Table 3: Incidence of COVID-19 Infection among congregated, corona virus infected persons in isolation center in ABC State, 2020

	Developed COVID-19?		
	Yes	**No**	**Total**
Isolation Center A	a = 30	b = 130	$H_1 = 160$
Isolation Center B	c = 5	d = 135	$H_0 = 140$

Total	35	265	T = 300

To calculate the risk ratio, first calculate **the risk or attack rate** for each group. Here are the formulas:
Attack rate for Exposed = a/a+b
Attack rate for unexposed = c/c+d
For this example:

Risk of COVID-19 in Isolation center A residents = 30/160 = 0.1875 = 18.7 %

Risk of COVID-9 in Isolation centre B residents = 5/140 = 0.0357 = 3.5 %

The risk ratio is simply the ratio of these two risks:

Risk ratio = 0.1875/0.0357 = 5.25
The risk ratio is more than 1 so the Isolation centre A residents were 5.3 times as likely to develop CORONA virus as those in Isolation centre B to developed CORONA virus.

Example 3: Calculating Risk Ratio

In an outbreak of measles among vaccinated children in Nigeria in 2016, measles was diagnosed in 18 of 153 of the children vaccinated compared with 3 out of 7 unvaccinated children. Calculate the risk ratio.

Table 3: Incidence of Measles among children in Nigeria in 2016

	Measles Exposed	**Non-case Non-exposed**	**Total**
Vaccinated	a = 18	b = 135	153
Unvaccinated	c = 3	d = 4	7
Total	21	139	160

Risk of measles among vaccinated children = 18/153 = 0.1176 = 11.8 %
Risk of measles among unvaccinated children = 3/7 = 0.4285 = 42.8 %

Risk ratio = 0.1176/0.4285 = 0.274
The risk ratio is less than 1.0, indicating a decreased risk or protective effect for the exposed (vaccinated children). The risk ratio of 0.27 indicates that vaccinated children were only approximately one-fourth as likely (27 %) to develop measles as were unvaccinated children.

3.3.2 Rate Ratio

A rate ratio compares the incidence rates, person-time rates, or mortality rates of two groups. It is closely related to risk ratio. As with the risk ratio, the two groups are differentiated by demographic factors or by exposure to a suspected causative agent. Rate ratio is computed as incidence rate in an exposed group ($IR_{exposed}$), i.e. the primary interest group divided by the incidence rate in the non-exposed group ($IR_{non\text{-}exposed}$) which is the comparison group.

$$\text{Rate Ratio} = \frac{\text{Incidence Rate for exposed group}}{\text{Incidence Rate for not-exposed group}}$$

$$\text{Rate Ratio} = \frac{IR_{exposed}}{IR_{non\text{-}exposed}}$$

The interpretation of the value of a rate ratio is similar to that of the risk ratio. That is:

- A rate ratio of 1.0 indicates equal rates in the two groups,
- A rate ratio greater than 1.0 indicates an increased risk for the group in the numerator, and
- A rate ratio less than 1.0 indicates a decreased risk for the group in the numerator.

EXAMPLE 4: Calculating Rate Ratios

Health officials were invited to investigate perceived increased visitors to ships' infirmaries for Acute Respiratory Illness (ARI) by passengers of cruise ships in Nigeria in 1967. The officials compared passenger visits to ship infirmaries for ARI from March – June 1967 and the same period in 1968. They recorded 12.6 visits for ARI per 1,000 tourists per week in 1967 compared with 6.3 visits per 1,000 tourists per week in 1968. Calculate the rate ratio.

$$\text{Rate Ratio} = \frac{\text{Incidence Rate for exposed group in 1967}}{\text{Incidence Rate for non-exposed group in 1968}}$$

$$\text{Rate Ratio} = \frac{IR_{exposed}}{IR_{non\text{-}exposed}}$$

$$\text{Rate Ratio} = 12.6 / 6.3 = 2.0$$

Interpretation: Passengers on cruise ships in Nigeria during March – June 1967 were twice more likely to visit their ships' infirmaries for ARI compared to passengers in 1968.

Example 5: A prospective cohort study investigated the effects of hormone replacement therapy (HRT) on coronary artery disease (CAD)

in post-menopausal women. The investigators computed the incidence rate of coronary artery disease in post-menopausal women who were taking HRT and compared it to the incidence rate in post-menopausal women who were not taking HRT. Calculate the rate ratio.

Table 4: A prospective cohort study investigated the effects of hormone replacement therapy (HRT) on coronary artery disease (CAD) in post-menopausal women

Post-menopausal hormone use	Number with CAD	Person-years of diseases-free Follow-up	Rate Ratio Calculation
Yes	40	65,409	40 / 75,409 = 0.000530 = 53.0 per 100, 000 person years
No	80	120,578	80 / 72,578 = 0.001102 = 110.2 per 100, 000 person years

Rate Ratio = 53.0/ 110.2 = 0.4809

Interpretation: Women who used post-menopausal hormones had 0.48 times the rate of CAD compared to women who did not use post-menopausal hormones. Note the similarity between the interpretation of Risk Ratio and Rate Ratio.

3.3.3 Odds Ratio

An odds ratio (OR) is another measure of association that quantifies the strength of association between two events - an exposure and an outcome, X and Y. It compares the intervention group with the comparism, control or placebo group. Odds ratio signifies the odds that an outcome will occur given an exposure, compared to the odds that the outcome will occur in the absence of the exposure.

Odds ratio is used to compare the relative odds of the occurrence of the outcome of interest such as disease or illness or given exposure to variable of interest. It can also be used to determine if a particular exposure is a risk factor for a specific outcome. Hence, odds ratio can be used to compare the magnitude of different risk factors for that particular outcome.

Table 5: Table showing how to calculate odds ratio

	Disease	**No Disease**	**Total**	**Risk (%)**	**Odds Ratio**
Exposed	A	b	a + b	a / a + b	ad⁄bc
Not Exposed	C	d	c + d	c / c + d	
Total	a + c	b + d	(a + b) + (c + d)		

Risk Ratio = (a / a + b) / (c / c + d)
Odds ratio is calculated as:
Odds Ratio = (a x d) / (b x c) = ad⁄bc
Where:

The odds ratio is sometimes referred to as the **cross-product ratio** because the numerator is the product of values of cell "a" and cell "d," (a x d) whereas the denominator is the product of cell "b" and cell "c" (b x c) as shown in the formular above. A line from cell "a" to cell "d" (for the numerator) and another from cell "b" to cell "c" (for the denominator) creates an x or cross on the two-by-two table (CDC, 2012).

Interpretation: If the:

- Odds ratio =1: there is no difference between the two arms of the study: intervention and control.
- Odds ratio is more than 1 (> 1): The control is better than the intervention.
- Odds ratio is less than 1 (< 1): The intervention is better than the control.

Example 6: Calculating Odds Ratios

Table 6: Exposure and Disease in a Hypothetical Population of 10,000 Persons

	Disease	**No Disease**		**Total**	**Risk**		**Odds Ratio**
Exposed	a = 200	b 3,800	=	4,000	0.05 5.0%	=	5.21
Not Exposed	c = 160	d 15,840	=	16,000	0.01 1.0%	=	
Total	**360**	**19,640**		**20,000**			

From the data above, the risk and odds ratios are calculated as follows:

1. *Risk ratio =* 5.0/1.0 = 5.0
2. Odds Ratio = (a x d) / (b x c) = ad/bc

= (200 × 15,840)/(3800 × 160) = 3168000 / 608,000 = 5.21

From the above results, the odds ratio of 5.21 is close to the risk ratio of 5.0. This is one of the advantages of odds ratio. The advantages of odds ratio are as follows:

i. For rare health outcome, the odds ratio provides a reasonable approximation of the risk ratio.
ii. The odds ratio can be calculated with data from a case-control study, but relative risk and rate ratio cannot be calculated. In a case-control study, researchers enroll a group of case-patients (distributed in cells and c of the two-by-two table), and a group of non-cases or controls (distributed in cells b and d) (CDC, 2012).

Only odds ratio can be calculated in a case-control study. A case-control study operates with two groups: diseased group (case-patients) and not-diseased group (control or a comparable group). The number of persons in the control group is usually decided by the investigator. Often, the size of the population from which the case-patients came is not known. Consequently, risks, rates, risk ratios and rate ratios cannot be calculated from a case-control study. But you can calculate odds ratio and interpret it as an approximation of the risk ratio, especially when the disease is rare in the population (CDC, 2012).

SELF ASSESSED EXERCISE

i. What is the difference between risk ratio and relative risk?
ii. State the formulas for Risk Ratio, Odds Ratio, using approprite formula.

4.0 CONCLUSION

In this unit, you have learnt the broad description of measure of association in epidemiology systems and approach. You also learnt what a case definition is and variation in case definition, during pandemic. Specifically, a case of pandemic COVID -19 was used for illustrative purposes. You were taught about the fundamental meaning / interpretations of measure of association and risk of diseases. Finally, you also learnt about the calculation method of attack rate and ratio.

5.0 SUMMARY

In all scientific endeavours, the practice of epidemiology relies on a systematic approach. In very simple terms, the epidemiologist c**ounts** cases or health events, and describes them in terms of time, place, and person; **Divides** the number of cases by an appropriate denominator to calculate rates; and **Compares** these rates over time or for different groups of people. A risk ratio (RR), also called **relative risk**, compares the risk of a health event (disease, injury, risk factor, or death) among one group with the risk among another group.

6.0 TUTOR-MARKED ASSIGNMENT

Using Table 7 below on the incidence of COVID-19 cases,

1. What is the formula for Risk Ratio?
2. What is total number of COVID-19 cases
3. Complete the table matrix
4. What is the risk Ratio?

Table 7: Incidence of COVID -19 infection among congregated, Corona virus infected persons by isolation centre in Kwara State, 2020

	Developed COVID 19?		
	Yes	**No**	**Total**
Isolation Centre A	a = 48	b = 229	H_1 =
Isolation Center B	c = 14	d = 233	H_0 =
Total			T =

Source: Saka M.J. and Mohammed, Bolarinwa Unpublished data

7.0 REFERENCES/FURTHER READING

CDC, (2012). Principles of Epidemiology in Public Health Practice. An Introduction to Applied Epidemiology and Biostatistics. Lesson 3: Measures of Risk: Section 5: Measures of Association. Third Edition https://www.cdc.gov/csels/dsepd/ss1978/lesson3/section5.html.Accessed 3/1/2021

Centres for Disease Control and Prevention. (2020). Case definitions for infectious conditions under public health surveillance. MMWR Recomm. Rep. 46(10):1–55.

Centres for Disease Control and Prevention. (2020). Indicators for chronic disease surveillance. MMWR Recomm. Rep. 53(11):1–6.

Centres for Disease Control and Prevention. (2019). Outbreak of severe acute respiratory syndrome–worldwide. MMWR, 52**:**226–8.

Centres for Disease Control and Prevention. (2019). Revised U.S. surveillance case definition for severe acute respiratory syndrome (SARS) and update on SARS cases–United States and worldwide, December. MMWR 52:1202–6.

Centres for Disease Control and Prevention. (2018). Summary of notifiable diseases–United States. MMWR, 5 (1):5-31.

MacDonald, P., Boggs, J., Whitwam, R., Beatty, M., Hunter, S., & MacCormack, N. (2001). Listeria-associated birth complications linked with homemade Mexican-style cheese, North Carolina, 50th Annual Epidemic Intelligence Service Conference, 23–27.

Saka, M.J. (2020). Report of Supportive Supervision of COVID 19 Cases at the Isolation Centre at Sobi Infectious Disease Centre Kwara State Nigeria. 1- 10.

Tugwell, B.D., Lee, L.E., Gillette, H., Lorber, E.M., Hedberg, K., & Cieslak, P.R. (2014). Chickenpox outbreak in a highly vaccinated school population. *Pediatrics*, 113(3): 455–459.

UNIT 3 INTEGRATED DISEASE SURVEILLANCE AND RESPONSE IN NIGERIA

CONTENTS

1.0 INTRODUCTION

Nigeria is situated in the West African sub region. Administratively, the country is divided into thirty-six States and one Federal Capital Territory (FCT) Abuja. There are 774 Local Government Areas (LGAs) which is the lowest administrative level. The 2006 National Population Census estimated the Nigerian population at 140 million, with an annual growth rate of 3.2%. In 2000, the World Health Organisation ranked Nigeria's overall health system performance as 187th among 191 member states. The health indicators for Nigeria are currently worse than the average for Sub-Saharan Africa; for example, infant mortality rate (IMR) is 78 out of 1000, under 5 years mortality rate is 147 out of 1000, and the maternal mortality rate (MMR) is 640 out of 100,000. Diseases such as malaria, diarrheal diseases, acute respiratory infections, and vaccine preventable diseases account for at least 90% of childhood morbidity and mortality and other childhood health problems in Nigeria. Other diseases like Lassa fever, Cerebrospinal Meningitis (CSM) and Measles continue to occur with increased frequency in epidemic proportion and produce highest case fatality rate. Nigeria, like all other countries in the region, is affected by the HIV/AIDS pandemic with a national prevalence rate of 4.4% (2005). This prevalence in 2006, the country experienced outbreak of highly pathogenic Avian Influenza (H5N1) in poultry and in 2007, a

human case was recorded. However, there has been a tremendous decline in HIV/AIDs prevalence in Nigeria. The current adult (15-49 years) HIV/AIDS prevalence rate is 1.4 % (NACA, 2020).

In September 1998, the 48th World Health Organisation Regional Committee for Africa met in Harare, Zimbabwe. Through resolution AFRO/RC48/R2, Member States adopted integrated disease surveillance as a regional strategy for strengthening weak national surveillance systems in the African region. Until 2008, the diseases under the Integrated Disease Surveillance and Response (IDSR) were mainly those diseases that are targeted for eradication, elimination, epidemic prone diseases and some communicable diseases of public health importance.
With the epidemiologic transition, non-communicable diseases are now contributing a significant burden of morbidity and mortality in Africa. Nigeria, like other developing countries, is facing a double burden of both communicable and Non-Communicable Diseases (NCDs). This guideline was revised to include non-communicable diseases that are of public health importance (e.g. Diabetes mellitus, High blood pressure), Neglected Tropical Diseases (Noma, Buruli ulcer), emerging infectious diseases such as H5N1 and SARS, and other diseases under International Health Regulations (IHR). On 23rd of May 2005, the Fifty-eighth World Health Assembly adopted the International Health Regulations in Geneva, Switzerland through Resolution WHA58.3. The International Health Regulations entered into force on 15th June, 2007.

By 2014, Nigeria had outbreak of Ebola Virus which was wide spread through the Nation and other countries. In December 2016, there was an outbreak of Lassa Fever and in February, 2020, Nigeria recorded its first case of COVID-19 till end of 2020 when it became a Pandemic disease. The availability of accurate, up-to-date, reliable, and relevant health data and information is essential for strengthening and managing the health system. Currently, there is paucity of relevant health data for policy decision and planning. The implementation of the health reform agenda, including strategies and action plans, is limited by the dearth of reliable data on health parameters at all levels of the health system. Where information flow exists, it remained exclusively vertical, from the periphery to the center, with little feedback.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- explain Surveillance Systems in Nigeria
- understand Integrated Diseases Surveillance Response (IDSR)
- describe how IDSR can contribute to COVID-19 Pandemic preparedness & response
- give detailed account of guideline provision before and during outbreak
- discuss the role of Village, Health Facility and LGA Level Services
- discuss Core Capacity Requirements for Surveillance and Response Under International Health Regulation (IHR).

3.0 MAIN CONTENT

3.1 What Is Disease Surveillance?

Surveillance is the ongoing systematic collection, analysis, and interpretation of health data. It includes the timely dissemination of the resulting information to those who need them for action. Surveillance is also essential for planning, implementation, and evaluation of public health practice. Data collected at health facility level is compiled and sent to the next level and regular feedback is shared with the lower level. A standard case definition is used to identify such priority diseases or events and the laboratory is recognised as an important component of public health surveillance. Several types of surveillance are used in national programs. The choice of method depends on the purpose of the surveillance action. In general, types of surveillance methods describe:

i. Focused location for surveillance (such as health facility-based surveillance or community-based surveillance).
ii. Designated or representative health facility or reporting site for early warning of epidemic or pandemic events (sentinel surveillance).
iii. Surveillance conducted at laboratories for detecting events or trends not necessarily evident at other sites.
iv. Disease-specific surveillance involving activities aimed at targeted health data for a specific disease.

Regardless of the type of surveillance, the important issue is that the health data is used for public health action.

3.2 What Is Integrated Disease Surveillance and Response?

The Integrated Disease Surveillance and Response (IDSR) is a strategy and a tool to promote rational use of resources by integrating and streamlining common surveillance activities. Many intervention programs still rely on their own disease surveillance systems. Each program has made efforts through the years to improve its ability to obtain reliable data on time in order to use information for taking action. Disease control and prevention objectives are successfully met when resources are dedicated to improving the ability of health officials to detect the targeted diseases, obtain laboratory confirmation of the disease, and use thresholds to initiate action at the LGA level. Building on these successes, the World Health Organization (WHO) Regional Office for Africa (AFRO) proposes an Integrated Disease Surveillance and Response (IDSR) strategy for improving disease surveillance in Nigeria linking community, health facility, LGA, State and National levels. Additionally, IDSR takes into account the ***One World-One Health*** perspective ***which*** is a strategy that addresses events at the intersection of human, domestic animal, wildlife, and the environment. For example, 75% of recently emerging and re-emerging diseases affecting human health are of animal origin (HIV/AIDS, avian influenza, etc.).

One World-One Health is an interdisciplinary, holistic and integrated approach to health. Diseases and other threats resulting from climate change, food safety, and chemical hazards constitute a complex set of challenging events involving human, animal and environmental health. The ***One World-One Health*** strategy promotes the integration and coordination within and across sectors for disease surveillance, outbreak investigation and response activities undertaken by professionals from various fields. It is a strategy that ensures the strengthening of each sector and enhances inter-sectoral linkages. This facilitates efficient utilisation of scarce resources, effective and prompts leveraging of various sectors capabilities for a better disease prevention and control. In an integrated system:

i. The LGA level is the focus for integrating surveillance functions. This is because the LGA is the first level in the health system with full-time staff dedicated to all aspects of public health such as monitoring health events in the community, mobilising community action, encouraging national assistance and accessing regional resources to protect the LGA's health.
ii. All surveillance activities are coordinated and streamlined. Rather than using scarce resources to maintain separate vertical activities, resources are combined to collect information from a single focal point at each level.
iii. Several activities are combined into one integrated activity and

take advantage of similar surveillance functions, skills, resources and target populations. For example, surveillance activities for acute flaccid paralysis (AFP) can address surveillance needs for neonatal tetanus, measles and other diseases. Thus, health workers who routinely monitor AFP cases can also review LGA and health facility records for information about other priority diseases.

iv. Surveillance focal points at the LGA, state and national levels collaborate with epidemic response committees at each level to plan relevant public health response actions and actively seek opportunities for combining resources.

Integration refers to the ***harmonisation of different*** methods, software, data collection forms, standards and case definitions in order to prevent inconsistent information as well as maximise efforts put in by stakeholders in disease prevention and control programmes. Integration involves training and supervision, use of common feedback bulletin, and other resources such as computers and vehicles are shared. IDSR involves nearly full-time coordination of surveillance activities and joint action (planning, implementation, monitoring, evaluation) whenever it is possible and useful.

Coordination refers to ***working or acting together effectively*** for the rational and efficient use of available but limited resources such as Health Management Information System (HMIS) and various disease programs. Coordination involves information sharing, joint planning, monitoring and evaluation in order to provide accurate, consistent and relevant data and information to policy-makers and stakeholders at LGA, state and national levels.

3.3 Goal and Objectives of Integrated Disease Surveillance and Response (Idsr)

The goal of Integrated Disease Surveillance and Response (IDSR) is to improve the ability of LGAs to detect and respond to diseases and conditions that cause high levels of death, illness and disability in the LGA's catchment area. Strengthening skills and resources for integrated disease surveillance and response will result in improved health and well-being of communities in the LGA. The general objective of the IDSR strategy is to provide a rational basis for decision-making and implementation of public health interventions that are efficacious in responding to priority diseases and events. The specific objectives of IDSR are to:

i. Strengthen the capacity *to conduct effective surveillance activities*: train personnel at all levels; develop and carry out plans of action; and advocate and mobilise resources.

ii. Integrate multiple surveillance *systems* so that forms, personnel and resources can be used more efficiently.
iii. *Improve the use of information* to detect changes in time in order to conduct a rapid response to suspect epidemics and outbreaks; monitor the impact of interventions: for example, declining incidence, spread, case fatality, and to facilitate evidence-based response to public health events; health policy design; planning; and management
iv. *Improve the flow of surveillance information* between and within levels of the health system.
v. *Strengthen laboratory capacity and involvement* in confirmation of pathogens and monitoring of drug sensitivity.
vi. *Increase involvement of clinicians* in the surveillance system.
vii. *Emphasise community participation* in detection and response to public health problems including event-based surveillance and response in line with IHR
viii. *Trigger epidemiological investigations* in detection, investigation and reporting of public health problems, and in the implementation of effective public health interventions.

3.4 How Can Integrated Disease Surveillance and Response Contribute to Covid-19 Pandemic Preparedness & Response?

When an outbreak of infectious disease occurs or is detected, there is usually no time to conduct initial training or assemble supplies. All efforts will be focused on meeting the needs of patients promptly and containing the outbreak in the community. Being prepared for an emergency situation can ultimately save lives. In cases where epidemic preparedness plans are in place, timely detection of outbreaks is followed by prompt and appropriate response actions.

Epidemiologic surveillance provides skills and information for early detection of outbreaks leading to enhanced preparedness for emergency situations because it collects data for describing and analysing health events. For example, a LGA's Epidemic Management Committee (EMC) can define relevant roles in outbreak response in advance. Limited resources are maximised by combining resources for training, demonstrations and setting aside adequate supplies of equipment, vaccines, drugs and supplies.

3.4.1 How is surveillance functions described in these guidelines?

These guidelines assume that all levels of the health system are involved in conducting surveillance activities for detecting and responding to

priority diseases and conditions (even though the different levels do not perform identical functions). These activities include the following core functions:

Step 1 - Identify cases and events: Use standard case definitions, identifying priority diseases, conditions and events.

Step 2 – Report: suspected cases or conditions or events to the next level. If this is an epidemic prone disease or a potential Public Health Emergency of International Concern (PHEIC), or a disease targeted for elimination or eradication, respond immediately by investigating the case or event and submit a detailed report. For events to be notified under IHR, use the decision instrument (Annex 2 of IHR) to identify any potential PHEIC.

Step 3 - Analyse and interpret findings: Compile the data, and analyse it for trends. Compare information with previous periods and summarise the results.

Step 4 - Investigate and confirm suspected cases, outbreaks or events: Take action to ensure that the case, outbreak or event is confirmed including laboratory confirmation wherever it is feasible.
Gather evidence about what may have caused the outbreak or event and use it to select appropriate control and prevention strategies.

Step 5 – Prepare: Take steps in advance of outbreaks or public health events so that teams may respond quickly and essential supplies and equipment are available for immediate action. It is also premised on the availability of surveillance data.

Step 6 – Respond: Coordinate and mobilise resources and personnel to implement the appropriate public health response.

Step7 - Communicate/Provide feedback: Encourage future cooperation by communicating with levels that provided data, reported outbreaks, cases and events about the investigation outcome and success of response efforts.

Step 8 - Evaluate and improve the system: Assess the effectiveness of the surveillance and response systems, in terms of timeliness, quality of information, preparedness, thresholds, case management and overall performance. Take action to correct problems and make improvements.
There is a role for surveillance function at each level of the health system. The levels are defined as follows:

i. **Community**
Represented by basic village-level services such as trained birth attendants, community or village health agents, or similar care providers, village leaders (religious, traditional or political) or school teachers, veterinarians and/or health extension workers, pharmacists, traditional healers.

ii. **Health facility**
Defined by each country. For surveillance purposes, all institutions (public, private, NGOs or other governmental) with out-patient and/or in-patient facilities are defined as a health facility.

iii. **Local Government Area (LGA)**
The LGA is the lowest administrative unit and is the level responsible for primary health care implementation.

iv. **State**
The intermediate level of government is responsible for supervision and provision of technical support to the LGA.

v. **National level**
This is the Federal level where policies are set with coordination of technical support to States and LGAs.

vi. **Laboratory**
In an integrated system, some laboratory services are available at each level guided by a national level system of quality assurance and linked to reference laboratories for specific diseases.

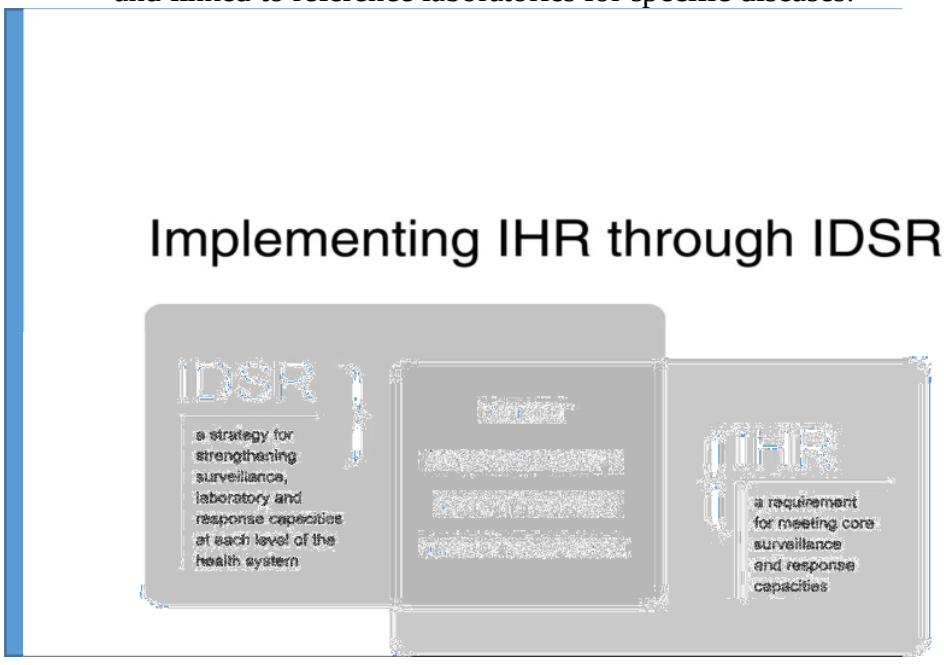


Figure 1: Implementing International Health Regulation (IHR) through Integrated Disease Surveillance and Response (IDSR)
Source: FMoH IDSR Guidelines

3.5 Core Capacity Requirements for Surveillance and Response Under International Health Regulation (Ihr)

According to International Health Regulation (IHR), member states shall use existing national structures and resources to meet their core capacity requirements. These requirements include capacity for surveillance, reporting, notification, verification, response and collaboration activities. Each part is expected to assess the ability of existing national structures and resource to meet the minimum requirements. Based on the results of the assessment, each member state should develop and implement action plan to ensure that these core capacities are present and functioning throughout the country.

Appendix 2 of the regulations defines the core capacity requirements for surveillance and response. The regulations recognise the following three levels of the health care system.

I. Local Government public health response level
II. State public health response level
III. National public health response level

I. Local Government public health response level

At the local government public health response level, the capacities are:

(a) To detect events involving disease or death above expected levels for the particular time and place in all areas within the LGA
(b) To report all available essential information immediately to the appropriate level of healthcare response. At the community level, reporting shall be to local community health-care institutions or the appropriate health personnel. At the LGA public health response level, reporting shall be to the State or national response level, depending on organisational structures.

For the purposes of these guidelines, essential information includes:

i. clinical descriptions
ii. laboratory results
iii. sources and type of risk
iv. numbers of human cases and deaths
v. conditions affecting the spread of the disease and the health measures employed

(c) To implement preliminary control measures immediately.

II. State Public Health response level

The core capacity requirements at State level are:

(a) To confirm the status of reported events and to support or implement additional control measures; and
(b) To assess reported events immediately and if found urgent, to report all essential information to the national level. For the purpose of this (Appendix 1C), the criteria for urgent events include serious public health impact and/or unusual or unexpected nature with high potential for spread.

III. National Level: Assessment and Notification

The capacity requirements at National level are:

(a) To assess all reports of urgent events within 48 hours; and
(b) To notify WHO immediately through the National IHR Focal Point when the assessment indicates the event is notifiable pursuant to Article 6,7 and 9 (Appendices 2 and 3).

National Level Public health response

The capacities are:

i. To determine rapidly the control measures required to prevent domestic and international spread
ii. To provide support through specialised staff, laboratory analysis of samples (domestically or through collaborating centres) and logistic assistance (e.g. equipment, supplies and transport)
iii. To provide on-site assistance as required to supplement local investigations
iv. To provide a direct operational link with senior health and other officials to approve rapidly and implement containment and control measures
v. To provide direct liaison with other relevant government ministries
vi. To provide by the most efficient means of communication available, links amongst stakeholders (hospitals, clinics, airports, sea ports, ground crossings, laboratories and other key operational areas) for the dissemination of information and recommendations received from WHO regarding events in the Country and other Countries
vii. To establish, operate and maintain a national public health emergency response plan, including the creation of multidisciplinary/multi-sectoral teams to respond to events that

may constitute a PHEIC; and

viii. To provide the foregoing on a 24-hour basis.

SELF ASSESSED EXERCISE

i. What is the difference between Disease Surveillance and Integrated Disease Surveillance?
ii. List diseases of Public Health Importance under IDSR.

4.0 CONCLUSION

This unit described disease surveillance and it is implication in the management of disease outbreak. In the unit you have learnt the meaning of surveillance system with integration with disease of public health importance. You should also have learnt relationship between Integrated Disease Surveillance and Response (IDSR) and International Health Regulation (IHR).

5.0 SUMMARY

This guideline was revised to include Non-Communicable Diseases that are of public health importance (e.g. Diabetes Mellitus, High Blood Pressure), Neglected Tropical Diseases (Noma, Buruli Ulcer), emerging infectious diseases such as H5N1 and SARS, and other diseases under International Health Regulations (IHR).

6.0 TUTOR-MARKED ASSIGNMENT (TMA)

1. Explain how outbreak is reported from the point of outbreak to the FMoH.
2. What are the relationships between IDSR and IHR?

7.0 REFERENCES/FURTHER READING

Federal Ministry of Health. (2019). Guideline of Integrated Disease Surveillance Response
Nigeria Center for Disease Control. (2020). (NCDC) www.ncdc.ng Registrar-General. (2018). Report on the mortality of cholera in England 1848–49. London, UK: Her Majesty's Stationery Office, 18-52.

Tauxe, R.V., Tormey, M.P., Mascola, L., Hargrett-Bean, N.T., & Blake, P.A. (1987). Salmonellosis outbreak on transatlantic flights; foodborne illness on aircraft. Am J Epidemiol, 125,150–7.

Tobler, W. A. (1997). Computer movie simulating urban growth in the Detroit region. Econ Geogr, 46(Suppl), 234–40.

Vazquez-Prokopec, G.M., Kitron, U., Montgomery, B., Horne, P., & Ritchie, S.A. (2010). Quantifying the spatial dimension of dengue virus epidemic spread within a tropical urban environment. *PLoS Negl Trop Dis* 4:920-930.

National Agency for Control of AIDS (NACA), 2020. Nigeria Prevalence Rate. https://naca.gov.ng/nigeria-prevalence-rate/. Accessed 4/1/2021.

MODULE 2 SYSTEMATIC ERROR (BIAS), CONFOUNDING AND STANDARDISATION

UNIT 1 BIAS IN PRINCIPLES OF EPIDEMIOLOGY

CONTENTS

1.0 INTRODUCTION

Before concluding that an individual study's conclusions are valid, one must consider three sources of error that might provide an alternative explanation for the findings. These are random error, bias and confounding. If a determination is made that the findings of a study were not due to any one of these three sources of error, then the study is considered **internally valid**. In other words, the conclusions reached are likely to be correct for the circumstances of that particular study. This does not necessarily mean that the findings can be generalised to other circumstances (**external validity**). For example, the Physicians' Health Study concluded that aspirin use reduced the risk of myocardial infarction in adult male physicians in the Kwara State, Nigeria. The study was carefully done, and the study was internally valid, but it was not clear that the results could be extrapolated to women, or even to non-physicians (whose risk of myocardial infarction is generally lower than that of the population overall). However, internally validity must be established before one can consider whether the results are externally valid. Therefore, investigators must first ensure that a study is internally valid even if that means that the generalisability of the findings will be compromised.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- define and explain bias (systematic error or systematic bias)
- identify different types of bias
- discuss selection bias in (Case Control, Cohort Study)
- describe self-control, Differential Surveillance, Referral or Diagnosis of Subjects in Selection Bias.
- describe loss to follow up, The Healthy Worker Effects, Subject Selection Bias in Cohort Study
- explain Differential and Non-differential misclassification of disease and exposure and mechanism in Information Bias
- identify Information Bias on recall, Interviewer and Differences in Quality of Information
- describe Information Bias on Misclassification of Outcome (Differential and Non-Differential)
- identify Effect Decreased Sensitivity and Specificity of Detecting Diseases Subject in Non- Differential Misclassification of Information Bias

3.0 MAIN CONTENT

3.1 Definition of Bias (Systematic Error)

In contrast to random error, **bias** refers to systematic errors in any type of epidemiologic study that result in an incorrect estimate of the association between exposures and outcomes. Investigators can introduce bias into a study as a result of the procedures for identifying and enrolling subjects or from the procedures for collecting or analysing information. Bias can also be introduced by errors in classification of outcomes or exposures. It is important for investigators to be mindful of potential biases in order to reduce their likelihood when they are designing a study, because once bias has been introduced, it cannot be removed. The two major types of bias are:

i. Selection Bias
ii. Information Bias

In addition, many epidemiologists think of confounding as a type of bias. While **confounding** also produces incorrect estimates of the association, one can often adjust for confounding in the analysis in order to remove its distorting effects to obtain a more accurate measure of association. Typically, confounding effects in epidemiologic studies is a function of the complex interrelationships between various exposures/susceptibility factors and disease. Confounding can be controlled in the design by randomisation, restriction and matching. Confounding can also be controlled in the analysis by stratification, multivariable analysis and matching. The best that can be done about unknown confounders in the conduct of epidemiologic studies is to use a randomised design. *The problem of confounding will be addressed in a separate unit of the module.*

When reading a study, one should be aware of potential biases that might have affected the conclusions and be able to predict what effect a given bias would be expected to have on the estimate of effect.

3.2 Selection Bias in Case Control Studies

3.2.1 Meaning of Selection Bias

Selection bias can be defined as a distortion in the estimate of association between risk factor and disease that results from how the subjects are selected for the study. Selection bias is an important problem in case-control and retrospective cohort studies. It is less likely to occur in a prospective cohort study design. It cannot be completely excluded in a

case-control study because non participation between cases and control subjects may have differed.

In general, there are two basic conditions that bring about the phenomenon of selection bias. It could occur, either because the sampling frame is sufficiently different from the target population *or* because the sampling procedure cannot be expected to deliver a sample that is a mirror image of the sampling frame. Either way, selection bias has serious implications for both internal validity of findings and their generalisability to the population, that is, the external validity of findings of the study.

Suffice it to say that selection bias can result when the selection of subjects into a study or their likelihood of being retained in the study leads to a result that is different from what you would have gotten if you had enrolled the entire target population. If one enrolled the entire population and collected accurate data on exposure and outcome, then one could compute the true measure of association. We generally don't enroll the entire population; instead we take samples. However, if one sampled the population in a fair way, such the sampling from all four cells was fair and representative of the distribution of exposure and outcome in the overall population, then one can obtain an accurate estimate of the true association (assuming a large enough sample, so that random error is minimal and assuming there are no other biases or confounding). Conceptually, this might be visualised by equal sized ladles (sampling) for each of the four cells.

Table 8: Sampling using Contingency Tables

Fair Sampling	**Diseased**	**Non-diseased**
Exposed		
Non-exposed		

The contingency table showed columns of diseased and non-diseased while the rows labeled exposed and non-exposed. In this example, the 4 exposure / disease categories have equal-sized star in them to convey the idea of unbiased sampling.

However, if sampling is not representative of the exposure-outcome distributions in the overall population, then the measures of association will be biased, and this is referred to as selection bias. Consequently, selection bias can result when the **selection of subjects** into a study or **their likelihood of being retained in a cohort study** leads to a result that is different from what you would have gotten if you had enrolled the entire target population. One example of this might be represented by the table below, in which the enrollment procedures resulted in disproportionately large sampling of diseased subject

who had the exposure.

Table 9: Sampling showing Selection Bias

Selection Bias	Diseased	Non-diseased
Exposed	★	★
Non-exposed	★	★

This contingency table has a larger Star in the cell tabulating the number of exposed subjects with disease. This is to indicate that there was a tendency to over-sample this category, for example, a case-control study in which cases were more likely to be selected if they had been exposed. There are several mechanisms that can produce this unwanted effect:

1. Selection of a comparison group ("controls") that is not representative of the population that produced the cases in a case-control study. (Control selection bias)
2. Differential loss to follow up in a cohort study, such that the likelihood of being lost to follow up is related to outcome status and exposure status. (Loss to follow-up bias)
3. Refusal, non-response, or agreement to participate that is related to the exposure and disease (Self-selection bias)
4. Using the general population as a comparison group for an occupational cohort study ("Healthy worker effect")
5. Differential referral or diagnosis of subjects

3.2.2 Control Section Bias

In a case-control study selection bias occurs when subjects for the "control" group are not truly representative of the population that produced the cases. Remember that in a case-control study the controls are used to estimate the exposure distribution (i.e., the proportion having the exposure) in the population from which the cases arose. The exposure distribution in cases is then compared to the exposure distribution in the controls in order to compute the odds ratio as a measure of association.
In the module on Overview of Analytic Studies and in the module on Measures of Association we considered a rare disease in a source population that looked like this:

Table 10: Control Section Bias

	Diseased	Non-diseased	Total	Risk Ratio	
Exposed	6	**1,002**	1,008	6/1008	= 0.00595

Non-exposed	5	**5,633**	5,638	5/5638 = 0.00088	

Risk Ratio = 0.00595/0.00088 = 6.76

The calculated the risk ratio for the entire population is = 6.76. However, when the outcome is rare, a case-control study is conducted because it is much more efficient. Consequently, in order to estimate the risk ratio, we could use the relative distribution of exposure in a sample of the population, provided that these controls are selected by procedures such that the sample is a representative estimate of the exposure distribution in the overall population.

If a control sample was selected appropriately, i.e. such that is was representative of exposure status in the population, then the case-control results might look like the table below.

Table 11: Proper selection of a control sample

	Cases	**Controls**
Exposed	6	**10**
Non-exposed	5	**56**

Note that the sample of controls represents only 1% of the overall population, but the exposure distribution in the controls (10/56) is representative of the exposure status in the overall population (1,002/5,633).

2. Odds Ratio = (a x d) / (b x c) = ad⁄bc = 6 x 56/10 x 5 = 6.72
As a result, the odds ratio = 6.72 gives an unbiased estimate ratio of the risk ratio.

In contrast, suppose that in the same hypothetical study controls were somewhat more likely to be chosen if they had the exposure being studied. The data might look something like this:

Table 12: A hypothetical study controls

	Cases	**Controls**
Exposed	7	**16**
Non-exposed	6	**50**

Here we have the same number of controls, but the investigators used selection procedures that were somewhat more likely to select controls who had the exposure. As a result, the estimate of effect, the odds ratio, was biased (OR = 3.65).

Conceptually, the bias here might be represented by the table below in which the large ladle indicates that non-diseased subjects with the exposure were over sampled.

Table 13: Selection bias

Selection Bias	Diseased	Controls
Exposed		
Non-exposed		

In this table the greater tendency to enroll non-diseased controls who had been exposed is represented by a larger Star in that cell.

Depending on which category is over or under-sampled, this type of bias can result in either an underestimation or an overestimation of the true association.

Example: A hypothetical case-control study was conducted to determine whether lower socioeconomic status (the exposure) is associated with a higher risk of cervical cancer (the outcome). The "cases" consisted of 250 women with cervical cancer who were referred to Isolo General Hospital for treatment for cervical cancer. They were referred from all over the State. The cases were asked a series of questions relating to socioeconomic status (household income, employment, education, etc.). The investigators identified control subjects by going from door–to-door in the community around Isolo General Hospital from 8:00 am to 6:00 pm. Many residents were not at home, but they persisted and eventually enrolled enough controls. The problem was that the controls were selected by a different mechanism than the cases (immediate neighbourhood for controls compared to statewide for cases), AND the door-to-door recruitment mechanism may have tended to select individuals of different socioeconomic status, since women who were at home may have been somewhat more likely to be unemployed. In other words, the controls were more likely to be enrolled (selected) if they had the exposure of interest (lower socioeconomic status).

The "Would" Criterion

Epidemiologists sometimes use the "**would" criterion** to test for the possibility of selection bias; they ask "If a control had had the disease, would they have been likely to be enrolled as a case?" If the answer is 'yes', then selection bias is unlikely.

3.2.3 Self Selection Bias

Selection bias can be introduced into case-control studies with low response or participation rates if the likelihood of responding or participating is related to both the exposure and the outcome.
Table 14 shows a scenario with differential participation rates in which diseased subjects who had the exposure had a participation rate of 80%, which the other three categories had participation rates of 60%.

Table 14: Differential participation rates

Selection Bias	Diseased	Non-diseased
Exposed		
Non-exposed		

In this contingency table greater participation by subjects who had the exposure and the outcome of interest is represented by the larger star in that cell.

3.2.4 Differential Surveillance, Referral, or Diagnosis of Subjects

Aschengrau and Seage (2018) give an example in which investigators conducted a case-control study to determine whether use of oral contraceptives increased the risk of thromboembolism. The case group consisted of women who had been admitted to the hospital for venous thromboembolism. The controls were women of similar age who had been hospitalised for unrelated problems at the same hospitals. The interviews indicated that 70 % of the cases used oral contraceptives, but only 20 % of the controls used them. The odds ratio was 10.2, but in retrospect, this was an overestimate. There had been reports suggesting such an association. As a result, health care providers were vigilant of their patients on oral contraceptives and were more likely to admit them to the hospital if they developed venous thrombosis or any signs or symptoms suspicious of thromboembolism. As a result, the study had a tendency to over sample women who had both the exposure and the outcome of interest.

Table 15: Over-sampling

Selection Bias	Diseased	Non-diseased
Exposed	X	X
Non-exposed	X	X

As seen in Table 15 Over-sampling of women with the exposure and the outcome is represented by a larger ladle for that category.
Aschengrau and Stage (2014) suggest that this selection bias could have been minimised by more restrictive case selection criteria, such that only women who clearly required hospitalisation would be enrolled in the case group.

3.2 Selection Bias in Cohort Studies

A cohort is defined as a group of subjects (persons) who share a common experience within a define time period. In cohort study designs, this purposively defined group of subjects are considered to be free of a given disease but vary in exposure potential to supposed risk factor(s) or causative agent(s) of the disease. They are defined or recruited on the basis of assumed exposure to risk factors rather than on the basis of disease status. Cohort studies are therefore operationally defined as epidemiological0 research studies in which a group of subjects (persons) who share a common experience within a define time period (a cohort) and who are considered to be free of a given disease but who vary in exposure to supposed risk factors are followed over time in order to determine differences in the rate at which disease develops in relation to exposure to the factor(s).

Depending on the time when the cohort study is initiated relative to occurrence of the disease(s) to be studied, it can be distinguished between prospective (concurrent) or retrospective (historical) cohort studies: In a concurrent cohort study, the data concerning exposure are assembled prior to the occurrence of disease —the concurrent cohort design thus representing a true prospective study. In a historical cohort study, sometimes referred to as retrospective cohort studies, data on exposure and occurrence of disease are collected after the events have taken place—the cohorts of exposed and non-exposed subjects are assembled from existing records, or health care registries. In a prospective cohort study, individuals are selected into the study at a time prior to development of symptoms (pre-pathogenesis stage) based on their exposure status and are followed for a specified period of time to determine the number of people who develop disease.

3.3.1 Subject Selection Bias

Factors affecting enrollment of subjects into a **prospective cohort study** would not be expected to introduce selection bias. In order for bias to occur, selection has to be related to both exposure and outcome. Subjects are enrolled in prospective cohort studies before they have experienced the outcome of interest. Therefore, while it is easy to see how enrollment might be related to exposure (exposed might be more or less likely to

enroll), it is difficult to imagine how either investigators or enrollees could be influenced by awareness of an outcome that hasn't yet occurred.

This form of selection bias could be more common in a **retrospective cohort study**, especially if individuals have to provide informed consent for participation. Since a retrospective cohort study starts after all cases of disease have occurred, subjects generally would know both their exposure and outcome status. It is not hard to imagine that those with the most interest in participation would both have been exposed and have the disease, a dynamic that would only be accentuated if the study question were a controversial one and/or there were potential liability and monetary consequences tied to the results of the study. Another less common mechanism of selection bias in a retrospective cohort study might occur if retention or loss of records of study subjects (e.g., employment, medical) were related to both exposure and outcome status. Selection bias can occur if selection or choice of the exposed or unexposed subjects in a retrospective cohort study is somehow related to the outcome of interest.

Example: Consider a hypothetical investigation of an occupational exposure (e.g., an organic solvent) that occurred 15-20 years ago in factory. Over the years there were suspicions that working with the solvent led to adverse health events, but no definitive data existed. Eventually, a retrospective cohort study was conducted using the employee health records. If all records had been retained the results might have looked like those shown in the first contingency Table 14 below.

Table 14: Employee health records

Unbiased Results	**Diseased**	**Non-diseased**	**Total**
Solvent exposure	100	900	1000
Unexposed	50	950	1000

This unbiased data would give a risk ratio as follows:

$$RR_{unbiased} = \frac{100/1000}{50/1000} = 2.0$$

However, suppose that many of the old records had been lost or discarded, but, given the suspicions about the effects of the solvent, the records of employees who had worked with the solvents and subsequently had health problems were more likely to be retained. Consequently, record retention was 99 % among workers who were exposed and developed health problems**,** but recorded retention was only 80% for all other

workers. This scenario would result in data shown in the next contingency Table 15.

Table 15: Contingency Table

Biased Results	Diseased	Non-diseased	Total
Solvent exposure	99	720	819
Unexposed	40	760	800

$$RR_{unbiased} = \frac{99/819}{40/600} = 2.42$$

Differential loss of records results in selection bias and an overestimate of the association in this case, although depending on the scenario, this type of selection bias could also result in an underestimate of an association

Prospective cohort studies will not have selection bias as they enroll subjects, because the outcomes are unknown at the beginning of a prospective cohort study. However, prospective cohort studies may have differential retention of subjects over time that is somehow related to exposure status and outcome, and this differential loss to follow up is also a type of selection bias that is analogous to what we saw above in the retrospective study on solvents in a factory.

3.3.2 Loss to Follow Up (Attrition) Bias

In cohort study design loss to follow-up becomes a threat to the internal validity of research findings when study participants that are lost to follow-up or who withdraw from the study re different from those who are followed for the entire duration of the study. Loss to follow up is said to be differential when is not equally distributed between comparison groups.

As noted above, the enrollment of subjects will not bias a prospective cohort study, because the outcome has not yet occurred. Therefore, choice cannot be related to both exposure status and outcome status. However, ***retention*** of subjects may be differentially related to exposure and outcome, and this has a similar effect that can bias the results, causing either an overestimate or an underestimate of an association. In the hypothetical cohort study below investigators compared the incidence of thromboembolism (TE) in 10,000 women on oral contraceptives (OC) and 10,000 women not taking OC. TE occurred in 20 subjects taking OC

and in 10 subjects not taking OC, so the true risk ratio was (20/10,000) / (10/10,000) = 2.

Table 16: Hypothetical cohort study

Unbiased Results	Thromboembolism	Non-diseased	Total
Oral Contraceptives	20	9980	10,000
Unexposed	10	9990	10,000

This unbiased data would give a risk ratio as follows:

$$\mathbf{RR_{unbiased}} = \frac{20/10{,}000}{10/10{,}000} = 2.0$$

However, suppose there were substantial loses to follow-up in both groups, and a greater tendency to loose subjects taking oral contraceptives who developed thromboembolism. In other words, there was differential loss to follow up with loss of 12 diseased subjects in the group taking oral contraceptives, but loss of only 2 subjects with thromboembolism in the unexposed group. This might result in a contingency table like the one shown below.

Table 17: Contingency table

Biased Results	Thromboembolism	Non-diseased	Total
Oral Contraceptives	8	5980	5988
Unexposed	8	5984	5992

This biased data would give a risk ratio as follows:

$$RR = \frac{8/5988}{8/5992} = 1.0$$

So, in this scenario both exposure groups lost about 40 % of their subjects during the follow up period, but there was a greater loss of diseased subjects in the exposed group than in the unexposed group, and it was this differential loss to follow up that biased the results.

Table 18: Bias due to differential loss to follow up

Selection Bias	Diseased	Non-diseased

In essence, the differential loss to follow up resulted in a relative under-sampling of exposed subjects who developed the outcome, as shown in this table with the small star in the upper right-hand cell.
Again, depending on which category is underreported as a result of differential loss to follow-up, either an underestimate or overestimate of effect (association) can occur.

Preventing Loss to Follow-up

The only way to prevent bias from loss to follow-up is to maintain high follow up rates (>80%). This can be achieved by:

i. Enrolling motivated subjects
ii. Using subjects who are easy to track
iii. Making questionnaires as easy to complete as possible
iv. Maintaining the interest of participants and making them feel that the study is important
v. Providing incentives

3.3.3 The "Healthy Worker" Effect

The "health worker" effect is a type of noncomparability bias that occurs in the design aspects of a research. It is really a special type of selection bias that occurs in cohort studies of occupational exposures when the general population is used as the comparison group. The general population consists of both healthy people and unhealthy people. Those who are not healthy are less likely to be employed, while the employed work force tends to have fewer sick people. Moreover, people with severe illnesses would be most likely to be excluded from employment, but not from the general population. As a result, comparisons of mortality rates between an employed group and the general population will be biased.
Suppose, for example, that a given occupational exposure truly increases the risk of death by 20% (RR=1.2). Suppose also that the general population has an overall risk of death that is 10% higher than that of the employed workforce. Given this scenario, use of the general population as a comparison group would result in an underestimate of the risk ratio, i.e. RR=1.1.

Another possibility is that the exposure being tested is not associated with any difference in risk of death (i.e., true RR=1.0). If the general

population is used as a comparison group, the estimated RR might be around 0.9. Aschengrau and Seage (2018) that found a 16% lower mortality rate (standardised mortality rate = 0.84 in radiation-exposed workers at the Portsmouth Shipyard. It was noted, however, that the radiation workers had to undergo a special physical examination in order to be eligible to work in this particular programme. Consequently, it is likely that their baseline health was significantly better than that of the population at large.

i. Information Bias (Observation Bias)

From the previous section it should be clear that, even if the categorisation of subjects regarding exposure and outcome is perfectly accurate, bias can be introduced differential selection or retention in a study. The converse is also true: even if the selection and retention into the study is a fair representation of the population from which the samples were drawn, the estimate of association can be biased if subjects are incorrectly categorised with respect to their exposure status or outcome. These errors are often referred to as ***misclassification***, and the mechanism that produces these errors can result in either ***non-differential*** or ***differential misclassification***. Ken Rothman distinguishes these as follows:

> "For exposure misclassification, the misclassification is no differential if it is unrelated to the occurrence or presence of disease; if the misclassification of exposure is different for those with and without disease, it is differential. Similarly, misclassification of disease [outcome] is no differential if it is unrelated to the exposure; otherwise, it is differential."

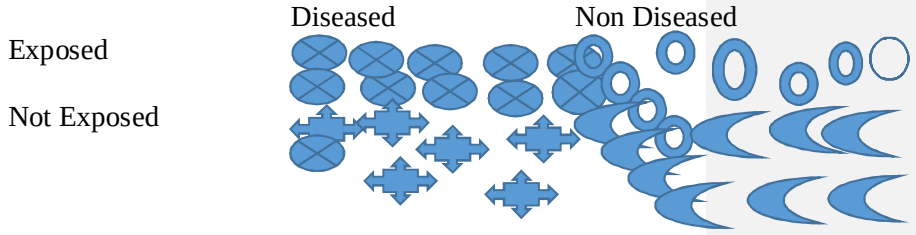
3.4 Non-Differential Misclassification of Exposure

Non-differential misclassification means that the frequency of errors is approximately the same in the groups being compared. Misclassification of exposure status is more of a problem than misclassification of outcome, but a study may be biased by misclassification of either exposure status, or outcome status, or both.

3.4.1 Non-differential Misclassification of Exposure

Non-differential misclassification of a dichotomous exposure occurs when errors in classification occur to the same degree regardless of outcome. Non-differential misclassification of exposure is a much more pervasive problem than differential misclassification (in which errors occur with greater frequency in one of the study groups). The figure 2 below illustrates a hypothetical study in which all subjects are correctly

classified with respect to outcome, but some of the exposed subjects in each outcome group were incorrectly classified as 'non-exposed.'



Non differential Misclassification of Exposure

Figure 2: A hypothetical study

Suppose a case-control study was conducted to examine the association between a high fat diet and coronary artery disease. Subjects with heart disease and controls without heart disease might be recruited and asked to complete questionnaires about their dietary habits in order to categorise them as having diets with high fat content or not. It is difficult to assess dietary fat content accurately from questionnaires, so it would not be surprising if there were errors in classification of exposure. However, it is likely that in this scenario the misclassification would occur with more or less equal frequency regardless of the eventual disease status. Non-differential misclassification of a dichotomous exposure always biases toward the null. In other words, if there is an association, it tends to minimise it regardless of whether it is a positive or a negative association.

The figure above depicts a scenario in which disease status is correctly classified, but some of the exposed subjects are incorrectly classified as non-exposed. This would result in bias toward the null. Rothman gives a hypothetical example in which the true odds ratio for the association between a high fat diet and coronary heart disease is 5.0, but if about 20 % of the exposed subjects were misclassified as 'not exposed' in both disease groups, the biased estimate might give an odds ratio of, say, 2.4. In other words, it resulted in bias toward the null.

However, now consider what would happen in the same example if 20 % of the exposed subjects were misclassified as 'non exposed' in both outcome groups, AND 20 % of the non-exposed subjects were misclassified as 'exposed' in both groups - in other words a scenario that looked something like this:

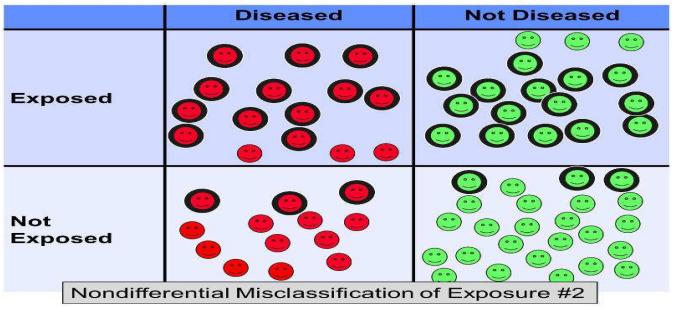


Figure 3: Rothman's hypothetical example

This additional nondifferential misclassification would result in even more severe bias toward the null, giving an odds ratio of perhaps 2.0.
Note that if there are multiple exposure categories, i.e. if the exposure is not dichotomous, then nondifferential misclassification may bias the estimate either toward the null or away from it, depending on the categories into which subjects are misclassified.

3.4.2 Mechanisms for Nondifferential Misclassification

Non-differential misclassification can occur in a number of ways. Records may be incomplete, e.g., a medical record in which none of the healthcare workers remember to ask about tobacco use. There may be errors in recording or interpreting information in records, or there may be errors in assigning codes to disease diagnoses by clerical workers who are unfamiliar with a patient's hospital course, diagnosis, and treatment. Subjects completing questionnaires or being interviewed may have difficulty in remembering past exposures. **Note** that if difficulty in remembering past exposures occurs to the same extent in both groups being compared, then there is nondifferential misclassification, which will bias toward the null. However, if one outcome group in a case-control study remembers better than the other, then there is a ***differential*** misclassification which is called "recall bias." Recall bias is described below under differential misclassification of exposure.

3.5 Differential Misclassification of Exposure

If errors in classification of exposure status occur more frequently in one of the groups being compared, then differential misclassification will occur, and the estimate of association can be overestimated or under estimated. There are several mechanisms by which differential misclassification of exposure can occur.

3.5.1 Recall Bias

Recall bias occurs when there are systematic differences in the way subjects remember or report exposures or outcomes. Recall bias can occur in either case-control studies or retrospective cohort studies. In a case-control study: subjects with disease may remember past exposures differently (more or less accurately) than those who do not have the disease.

Example 7: Mothers of children with birth defects are likely to remember drugs they took during pregnancy differently than mothers of normal children. In this particular situation the bias is sometimes referred to as maternal recall bias. Mothers of the affected infants are likely to have thought about their drug use and other exposures during pregnancy to a much greater extent than the mothers of normal children. The primary difference arises more from under reporting of exposures in the control group rather than over reporting in the case group. However, it is also possible for the mothers in the case group to under report their past exposures. For example, mothers of infants who died from SIDS may be inclined to under report their use of alcohol or recreational drugs during pregnancy.

Recall bias occurs most often in case-control studies, but it can also occur in retrospective cohort studies. For example, those who have been exposed to a potentially harmful agent in the past may remember their subsequent outcomes with a different degree of completeness or accuracy.

Example 8: In the retrospective portion of the Ranch Hand Study which looked at effects of exposure to Agent Orange (dioxin). Pilots who had been exposed may have had a greater tendency to remember skin rashes that occurred during the year following exposure.

Pitfall: In a case-control study, if both cases and controls have more or less ***equal difficulty*** in remembering past exposures accurately, it is nondifferential, and it is a form of ***nondifferential*** misclassification. In contrast, if one group remembers past exposures more accurately than the other, then it is called "recall bias" which is a ***differential*** type of misclassification.

Ways to Reduce Recall Bias

i. Use a control group that has a different disease (that is unrelated to the disease under study).
ii. Use questionnaires that are carefully constructed in order to maximise accuracy and completeness. Ask specific questions.

iii. For socially sensitive questions, such as alcohol and drug use or sexual behaviours, use a self-administered questionnaire instead of an interviewer.
iv. If possible, assess past exposures from biomarkers or from pre-existing records.

3.5.2 Interviewer Bias (Also Recorder Bias)

Differential bias can be introduced into a study when there are systematic differences in soliciting, recording, or interpreting information on exposure (in a case-control study) or outcome (in retrospective and prospective cohort studies and in intervention studies (clinical trials)). This type of bias can also occur when data is collected by review of medical records if the reviewer (abstractor) interprets or records information differently for one group or if the reviewer searches for information more diligently for one group. Since this introduces a differential misclassification, it can cause bias either toward or away from the null, depending on the circumstances.
Ways to Reduce Interviewer Bias

i. Use standardised questionnaires consisting of closed-end, easy to understand questions with appropriate response options.
ii. Train all interviewers to adhere to the question and answer format strictly, with the same degree of questioning for both cases and controls.
iii. Obtain data or verify data by examining pre-existing records (e.g., medical records or employment records) or assessing biomarkers.

3.5.3 Differences in the Quality of Information

Obviously, if the data for each of the groups being compared comes from different sources, the accuracy of the data may be better in one group, and this will introduce differential misclassification. For example, if exposure data for a case group were obtained from a facility specialising in the care of that condition and data from the comparison group were obtained from another source, there might be significant differences in the completeness and accuracy of the exposure data.

3.6 Misclassification of Outcome

Misclassification of outcomes can also introduce bias into a study, but it usually has much less of an impact than misclassification of exposure. First, most of the problems with misclassification occur with respect to exposure status since exposures are frequently more difficult to assess and categorise. We glibly talk about smokers and non-smokers, but what do these terms really mean? One needs to consider how heavily the

individual smoked, the duration, how long ago they started, whether and when they stopped, and even whether they inhaled or whether they were exposed to environmental smoke. In addition, as illustrated above, there are a number of mechanisms by which misclassification of exposure can be introduced. In contrast, most outcomes are more definitive and there are few mechanisms that introduce errors in outcome classification.

Another important consideration is that most of the outcomes that one studies are relatively uncommon; even when an association does exist, the majority of exposed and non-exposed subjects do not experience the outcome. As a result, there is much less potential for errors to have a major effect in distorting the measure of association.
Certainly, there may be clerical and diagnostic errors in classification of outcome, but compared to the frequency of exposure misclassification, errors in outcome classification tend to be less common and have much less impact on the estimate of association. In addition to having little impact on the estimate of effect, misclassification of outcome will generally bias toward the null, so if an association is demonstrated, if anything the true effect might be slightly greater.

Example 9: Consider the case-control conducted by Doll and Hill in 1947. This was one of the first analytic studies that examined the association between smoking and lung cancer. The study gathered data from more than twenty hospitals in the London area. Cases with a recent diagnosis of lung cancer were identified and interviewed about their past exposures, including a detailed history of smoking tobacco. Non-cancer control patients in the same hospitals were also interviewed. The study was quite extensive, but the bottom line was that statistically significant associations between smoking and lung cancer were found in both males and females (although the association was not as strong in females.

The investigators took steps to verify the diagnoses whenever possible by checking operative findings, pathology reports, and autopsy findings. Given the nature of the disease and the efforts to verify the diagnosis, it is likely that the diagnosis was correct in the vast majority of subjects. However, far more problematic was the classification of the degree of exposure to tobacco. The assessment of exposure could be influenced not only by misclassification as a result of trying to remember the details of smoking exposure over a lifetime, but the potential problems with recall bias and interviewer bias.

3.6.1 Differential Misclassification of Outcome

To illustrate differential misclassification of outcome Rothman uses the following example" "Suppose a follow-up study were undertaken to compare incidence rates of emphysema among smokers and nonsmokers.

Emphysema is a disease that may go undiagnosed without unusual medical attention. If smokers, because of concern about health effects of smoking (such as bronchitis), seek medical attention to a greater degree than nonsmokers, then emphysema might be diagnosed more frequently among smokers than among nonsmokers simply as a consequence of the greater medical attention. Unless steps were taken to ensure comparable follow-up, an information bias would result. An 'excess' of emphysema incidence would be found among smokers compared with nonsmokers that is unrelated to any biologic effect of smoking. This is an example of differential misclassification, since the under diagnosis of emphysema, a misclassification error, occurs more frequently for nonsmokers than for smokers."

3.6.2 Non-differential Misclassification of Outcome

Non-differential misclassification of a dichotomous outcome will ***generally*** bias toward the null, but there are situations in which it will not bias the risk ratio. Bias in the risk difference depends upon the sensitivity (probability that someone who truly has the outcome will be identified as such) and specificity (probability that someone who does not have the outcome will be identified as such).

This is additional detail on the effects of non-differential misclassification of outcome that is not required in the introductory course, although it is required in Intermediate Epidemiology.

3.6.2.1 Effect of Decreased Sensitivity of Detecting Diseased Subjects

"Consider a cohort study in which 40 cases actually occur among 100 exposed subjects and 20 cases actually occur among 200 unexposed subjects. Then, the actual risk ratio is (40/100) / (20/200) = 4, and the actual risk difference is 40/100-20/200 = 0.30. Suppose that specificity of disease detection is perfect (there are no false positives), but sensitivity is only 70% in both exposure groups (that is sensitivity of disease detection is nondifferential and does not depend on errors in classification of exposure). The expected numbers detected will then be 0.70(40) = 28 exposed cases and 0.70(20) = 14 unexposed cases, which yield an expected risk-ratio estimate of (28/100)/ (14/200) = 4 and expected risk-difference estimate of 28/100 - 14/200 = 0.21. Thus, the disease misclassification produced no bias in the risk ratio, but the expected risk-difference estimate is only 0.21/0.30 = 70 % of the actual risk difference.

"This example illustrates how independent nondifferential disease misclassification with perfect specificity will not bias the risk-ratio estimate, but will downwardly bias the absolute magnitude of the risk-difference estimate by a factor equal to the false-negative probability

(Rogers and MacMahon, 1995). With this type of misclassification, the odds ratio and the rate ratio will remain biased toward the null, although the bias will be small when the risk of disease is low (<10 %) in both exposure groups. This approximation is a consequence of the relation of the odds ratio and the rate ratio to the risk ratio when the disease risk is low in all exposure groups."
The scenario described above could be summarised with the following contingency Tables 19 and 20.
First, consider the true relationship:

Table 19: True Relationship A:

	Diseased	**Non-diseased**	**Total**
Exposed	40	60	100
Unexposed	20	180	200

Sensitivity = 100% (all disease cases were detected)
Specificity = 100% (all non-cases correctly classified)
Risk Ratio = (40/100)/ (20/200) = 4
Risk Difference = 40/100-20/200 = 0.30

Then consider:

Table 20: Misclassification of Outcome #1

	Diseased	**Non-diseased**	**Total**
Exposed	28	72	100
Unexposed	14	186	200

Sensitivity = 70% (30% false negative rate)
Specificity = 100% (all non-cases correctly classified)
Risk Ratio = (28/100)/ (14/200) = 4
Risk Difference = 28/100-14/200 = 0.21

This illustrates the effect when all of the non-diseased subjects are correctly classified, but some of the diseased subjects are misclassified as non-diseased. As you can see the risk ratio is not biased under these circumstances, but the risk difference is. The reason for this is that decreased sensitivity results in a proportionate decrease in the cumulative incidence in both groups, so the ratio of the two (the risk ratio) is unchanged. However, the groups are of unequal size, so the absolute difference between the groups (the risk difference) does change.

3.6.2.2 Effect of Decreased Specificity of Detecting Diseased Subjects

It is also possible for non-diseased subjects to be incorrectly classified as diseased, i.e., specificity <100%. For the scenario above, suppose that sensitivity had been 100% (all of the truly diseased subjects were identified), but the specificity was only 70%, i.e., 70% of the non-diseased people were correctly categorised as non-diseased, but 30% of them were incorrectly identified as diseased. In that case the scenario would give a contingency table as illustrated below in Table 21.

Table 21: Misclassification of Outcome #2

	Diseased	**Non-diseased**	**Total**
Exposed	58	42	100
Unexposed	74	126	200

Sensitivity = 100% (all disease cases were detected)
Specificity = 70% (30% of non-cases incorrectly classified)
Risk Ratio = (58/100)/ (74/200) = 1.57
Risk Difference = 58/100-74/200 = 0.58-0.37= 0.21

Here, the specificity is 70% in both groups, but there are more non-diseased subjects in the unexposed, so the result is a disproportionate increase in the apparent number of diseased subjects in the unexposed group, and both the risk ratio and the risk difference are underestimated. This is also true when the number of subjects in the exposed group is larger as illustrated in the example below. First, consider true relationship B:

Table 22: True Relationship B

	Diseased	**Non-diseased**	**Total**
Exposed	40	160	200
Unexposed	10	90	100

Sensitivity = 100% (all disease cases were detected)
Specificity = 100% (all non-cases correctly classified)
Risk Ratio = (40/200)/ (10/100) = 0.2/0.1 = 2
Risk Difference = 40/200-10/100 = 0.20-0.10 = 0.21

In contrast, consider the next table with misclassification of outcome, but a larger number of exposed subjects.

Table 23: Misclassification of Outcome #3

	Diseased	Non-diseased	Total
Exposed	88	112	200
Unexposed	37	63	100

Sensitivity = 100% (all disease cases were detected)
Specificity = 70% (30% of non-cases incorrectly classified)
Risk Ratio = (88/200)/ (40/100) = 0.44/0.40 = 1.1
Risk Difference = 88/200-40/100 = 0.44-0.40= 0.04

In this example, sensitivity is again 100% and specificity is 70%. As a result, 0.30*160 = 48 diseased subjects in the exposed group are incorrectly classified as diseased and move from cell B to cell A. Similarly, in the unexposed group, 0.30*90 = 27 non-diseased people are incorrectly classified as diseased and move from cell D to cell C. Again, the risk ratio and the risk difference are biased toward the null.

SELF ASSESSED EXERCISES

Can self-selection bias occur in prospective cohort studies?

4.0 CONCLUSION

In this unit, you have learnt the meaning of Bias (systemic error) in epidemiology and importance of bias and various bias types encountered in an epidemiology study. You were sufficiently exposed to the implication of each type of bias mechanism of control technique. You were also exposed to various methods of bias misclassification of exposure and outcome in differential and non-differential with effect of decreased sensitivity or specificity of detecting diseases subject.

5.0 SUMMARY

You learnt that Bias is systemic error encountered in a research study. When listening to a presentation or reading an article in which data is presented to support a conclusion, one must always consider alternative explanations that may threaten the validity of the conclusions. Specifically, one needs to consider whether random error, bias or confounding could have undermined the conclusions to a significant extent. Virtually all studies have potential flaws, but carefully done studies are designed and conducted in a way that minimises these problems so that they don't have any important effect on the conclusions. However, in other studies that are conducted in difficult circumstances (e.g., a prospective cohort study in a homeless population in which one would expect difficulty maintaining follow-up) or in poorly designed

studies, biases may have a major impact and produce large overestimates or underestimates of the true association.

In view of this, it is always important to ask oneself:

1. Given the conditions of the study, could bias have occurred?
2. Is it likely that bias actually was present?
3. If there were bias, would it bias the results toward the null or away from the null?
4. Is the magnitude of distortion likely to be small and inconsequential or large?

6.0 TUTOR-MARKED ASSIGNMENT (TMA)

1. Explain differential misclassification in Bias Study.
2. Discuss the effect of decreased sensitivity and specificity of detecting diseased subjects

7.0 REFERENCE/FURTHER READING

Ann Aschengrau, (2018). Essential Epidemiology in Public Health 44 pages: Jones & Bartlett Learning; 4th edition

Gordis, L. (2014). Epidemiology (5th ed.). Philadelphia, PA: Elsevier Saunders. Epidemiology principle http://ocw.jhsph.edu/courses/FundEpiII/PDFs/Lecture18.pdf accessed 2020.

Lash, Brian D Bradbury, Jemma B Wilk, Ann Aschengrau (2005), Breast Cancer Research journal, Volume 7 Issue 3 Pages R385. BioMed Central

Tauxe, R.V., Tormey, M.P., Mascola, L., Hargrett-Bean, N.T., & Blake, P.A. (1987). Salmonellosis outbreak on transatlantic flights; foodborne illness on aircraft. Am J Epidemiol, 125,150–7. Epid. Studies http://www.ump.edu.pl/files/8_483_errors_in_epidemiological_studies.pdf Accessed 2020.

Tobler, W. A. (1997). Computer movie simulating urban growth in the Detroit region. Econ Geogr, 46(Suppl), 234–40.

Vazquez-Prokopec, G.M., Kitron, U., Montgomery, B., Horne, P., & Ritchie, S.A. (2010). Quantifying the spatial dimension of dengue virus epidemic spread within a tropical urban environment. PLoS Negl Trop Dis, 4:920-930.

White, F., Stallones, L. & Last, J. M. (2013). Global public health: Ecological foundations. New York, NY: *Oxford University Press*,23-30.

Xie, S., Zeng, G., & Lei, J. (2013). A highly efficient transmission of SARS among extended family and hospital staff in Beijing, China, April 2003. Presented at the 2nd Southeast Asian and Western Pacific Bi-Regional TEPHINET Scientific Conference, November 24–28, 2003, Borocay, Philippines.

UNIT 2 CONFOUNDING AND EFFECT MEASURE MODIFICATION

CONTENTS

1.0 INTRODUCTION

Confounding is a distortion of the association between an exposure and an outcome that occurs when the study groups differ with respect to other factors that influence the outcome. Unlike selection and information bias, which can be introduced by the investigator or by the subjects, confounding is a type of bias that can be adjusted for in the analysis, provided that the investigators have information on the status of study subjects with respect to potential confounding factors.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- explain the meaning and conditions that must be present for confounding to occur
- give detail account of Magnitude and Method of Determining Confounding
- identify Residual Confounding, Confounding by Indication, & Reverse Causality
- describe how to control confounding in a study

3.0 MAIN CONTENT

3.1 What Is Confounding?

Confounding is a distortion (inaccuracy) in the estimated measure of association that occurs when the primary exposure of interest is mixed up with some other factor that is associated with the outcome. In the diagram below, the primary goal is to ascertain the strength of association between physical inactivity and heart disease. Age is a confounding factor because it is associated with the exposure (meaning that older people are more likely to be inactive), and it is also associated with the outcome (because older people are at greater risk of developing heart disease).

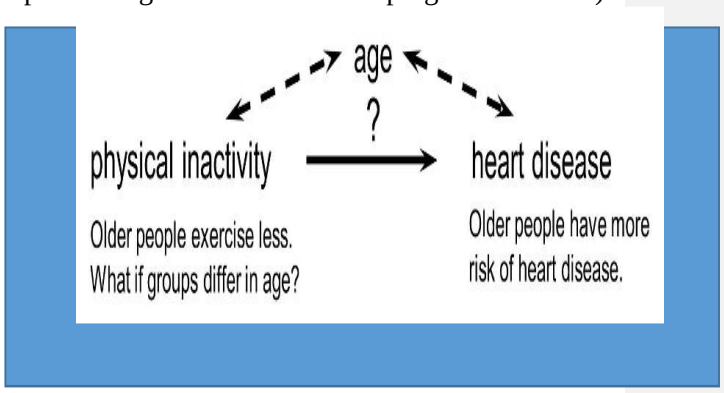
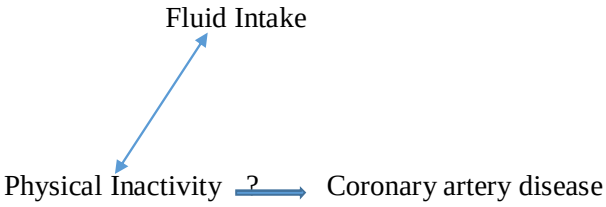
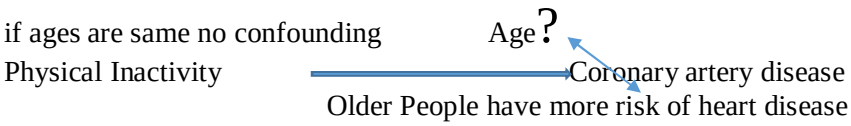


Figure 4: Age as a confounding factor for heart disease Source: Naim, (2005)

In order for confounding to occur, the extraneous factor must be associated with both the primary exposure of interest and the disease outcome of interest. For example, subjects who are physically active may drink more fluids (e.g., water and sports drinks) than inactive people, but drinking more fluid has no effect on the risk of heart disease, so fluid intake is not a confounding factor here.



Or, if the age distribution is similar in the exposure groups being compared, then age will not cause confounding.



3.1.1 Refining Our Understanding of Confounding

Rothman and others (2016 use a study by Stark and Mantel (2008) to illustrate the key features of confounding. These authors investigated the association between birth order and the risk of Down syndrome. The first graph to the right shows a clear trend toward increasing prevalence of Down syndrome with increasing birth order, or an association between increasing birth order and risk of Down syndrome.

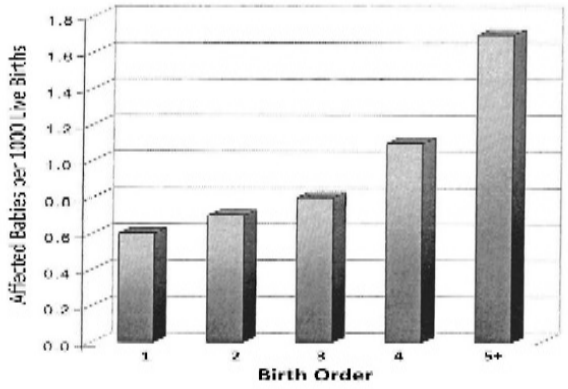


Figure 5: Association between increasing birth order and risk of Down syndrome.
[Data and graphs adapted from Rothman K: Epidemiology: An Introduction using data from Stark CR and Mantel N: Effects of maternal age and birth order on the risk of mongolism and leukemia. J. Natl. Cancer Inst. 37(5):687-98, 1966]

A 5th born child appears to have roughly a 4-fold increase in risk of being born with Down syndrome. Results like this also invite us to think about the mechanisms by which this occurred. Why might birth order cause a greater risk of Down syndrome? Keep in mind that this analysis does not consider any other "risk factors" besides birth order.

However, consider also that the order in which a women's children are born is also linked to her age at the time of her child's birth. When Stark and Mantel examined the relationship between maternal age at birth and risk of the child having Down syndrome, they observed the relationship depicted in the bar graph below. This shows an even more striking relationship between maternal age at birth and the child's risk of being born with Down syndrome.

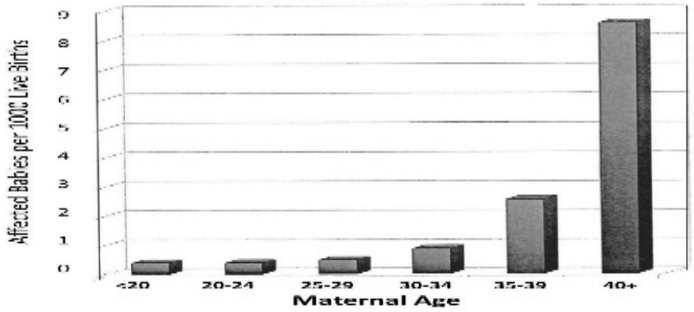


Figure 6: Relationship between maternal age at birth and the child's risk of being born with Down syndrome. (Source Rothman K 1996)

Obviously, women giving birth to their fifth child are on average, older than women giving birth to their first child. In other words, birth order of

children is mixed up with maternal age when a child is born. The correlation between maternal age and prevalence of Down syndrome is much stronger than the correlation with birth order, and a woman having her 5th child is clearly older than when she gave birth to her previous children. In view of this, the relationship between birth order and prevalence of Down syndrome is confounded by age. In other words, the association between birth order and Down syndrome is exaggerated by the confounding effect of maternal age.

But is the converse also true? Is the effect of maternal age confounded by birth order? It is possible, but only if birth order really has some independent effect on the likelihood of Down syndrome, i.e. an effect independent of the fact that birth order is linked to maternal age. Rothman points out that a good way to sort this out is to look at both effects simultaneously, as in the graph below.

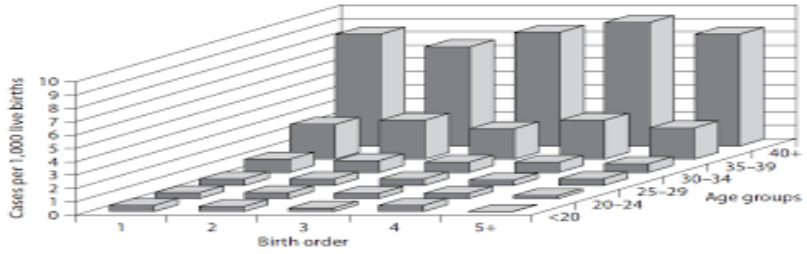


Figure 7: Effect of birth order on maternal age (Source: Wayne W. LaMorte, MD, PhD, MPH Boston University School of Public Health June 2016

In a sense this graph shows the relationships by stratifying the prevalence of Down syndrome by both birth order and maternal age. If one focuses on how prevalence changes within any particular maternal age group looking from side to side, it is clear that increasing birth order does not correlate with the prevalence of Down syndrome. In other words, if one "controls for maternal age," there is no evidence that birth order has any impact. On the other hand, if one now examines changes in prevalence within each of the birth order groups by looking from front to back within a given birth order, there is clearly a marked increase in prevalence as maternal age increases within all five levels of birth order. In other words, even after taking birth order into account (i.e., controlling for birth order) the strong association with maternal age persists.

Based on this analysis one can conclude that the association between birth order and Down syndrome was confounded by age. The different birth order groups had different age distributions, and maternal age is clearly associated with prevalence of Down syndrome. As a result, the apparent association between birth order and Down syndrome that was seen in the first figure was completely due to the confounding effect of age. On the

other hand, the association between maternal age and Down syndrome was NOT confounded by birth order, because birth order has no impact on the prevalence of Down syndrome, and the association between age and Down was not distorted by differences in birth order.

3.2 Conditions Necessary for Confounding

There are three conditions that must be present for confounding to occur:

1. The confounding factor must be associated with **both** the risk factor of interest and the outcome.
2. The confounding factor must be distributed unequally among the groups being compared.
3. A confounder cannot be an intermediary step in the causal pathway from the exposure of interest to the outcome of interest.

For example, it is known that modest alcohol consumption is associated with a decreased risk of coronary heart disease, and it is believed that one of the mechanisms by which alcohol causes a reduced risk is that alcohol raises blood levels of HDL, the so called "good cholesterol." Higher levels of HDL are known to be associated with a reduced risk of heart disease. Consequently it is believed that modest alcohol consumption raises HDL levels, and this, in turn, reduces coronary heart disease. In a situation like this HDL levels are not confounder of the association between alcohol and heart disease, because it is part of the mechanism by which alcohol produces this beneficial effect. If increased HDL is a consequence of alcohol consumption and part of the mechanism by which it lowers the risk of heart disease, then it is not a confounder.

Not surprisingly, since most diseases have multiple contributing causes (risk factors), there are many possible confounders.

i. A confounder can be another ***risk factor*** for the disease. For example, in the hypothetical cohort study testing the association between exercise and heart disease, age is a confounder because it is a risk factor for heart disease.
ii. Similarly, a confounder can also be a ***preventive factor*** for the disease. If those people who exercised regularly were more likely to take aspirin, and aspirin reduces the risk of heart disease, then aspirin use would be a confounding factor that would tend to exaggerate the benefit of exercise.

iii. A confounder can also be a ***surrogate or a marker*** for some other cause of disease. For example, socioeconomic status may be a confounder in this example because lower socioeconomic status is a marker for a complex set of poorly understood factors that seem to carry a higher risk of heart disease.

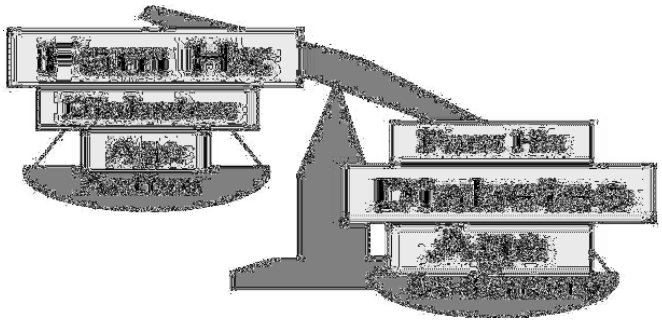


Figure 8: Confounding effects of socioeconomic status (Source Shah 2005)

As a result, there may be many possible confounding factors that could influence an association. For example, in looking at the association between exercise and heart disease, other possible confounders might include age, diet, smoking status and a variety of other risk factors that might be unevenly distributed between the groups being compared.

Aside from their physical inactivity, sedentary subjects may be more likely to smoke, to have high blood pressure and diabetes, and to consume diets with a higher fat content; all of these factors would tend to increase the risk of coronary heart disease. On the other hand, subjects who go to a gym regularly (active) may be more likely to be males and perhaps more likely to have a family history of heart disease, i.e., factors that might increase the risk of active subjects. Consequently, there may be many confounders that can distort the estimate of association in one direction or another.

Table 24: Confounders and their percentage effects

	Active	Sedentary
Age (years)	46 ± 1.4	59 ± 1.5
Dietary Fat (%)	29 ± 5.0	42 ± 7.0
Current Smokers	5%	24%
Hypertension	8%	17%
Diabetes	2%	9%
Family History of Heart Disease	25%	5%
Males	60%	40%

1. Identifying Confounding

a. A simple, direct way to determine whether a given risk factor caused confounding is to compare the estimated measure of association before and after adjusting for confounding. In other words, compute the measure of association both before and after adjusting for a potential confounding factor. If the difference between the two measures of association is 10% or more, then confounding was present. If it is less than 10%, then there was little, if any, confounding. How to do this will be addressed in greater detail below.

b. Other investigators will determine whether a potential confounding variable is associated with the exposure of interest and whether it is associated with the outcome of interest. If there is a clinically meaningful relationship between the variable and the risk factor and between the variable and the outcome (regardless of whether that relationship reaches statistical significance), the variable is regarded as a confounder.

c. Still other investigators perform formal tests of hypothesis to assess whether the variable is associated with the exposure of interest and with the outcome.

3.3 Quantified Magnitude of Confounding

The magnitude confounding

The magnitude confounding can be quantified by computing the percentage difference between the crude and adjusted measures of effect. There are two slightly different methods that investigators use to compute this, as illustrated below.

Percent difference is calculated by calculating the difference between the starting value and ending value and then dividing this by the starting value. Many investigators consider the crude measure of association to be the "starting value".

Method Favoured by Biostatisticians

$$\text{Magnitude of confounding} = \frac{RR_{crude} - RR_{adjusted}}{RR_{crude}}$$

Other investigators consider the adjusted measure of association to be the starting value, because it is less confounded than the crude measure of association.

Method Favoured by Epidemiologists

$$\text{Magnitude of confounding} = \frac{RR_{crude} - RR_{adjusted}}{RR_{adjusted}}$$

While the two methods above differ slightly, they generally produce similar results and provide a reasonable way of assessing the magnitude of confounding. Note also that confounding can be negative or positive in value.

3.4 Residual Confounding, Confounding by Indication, & Reverse Causality

3.4.1 Residual Confounding

Residual confounding is the distortion that remains after controlling for confounding in the design and/or analysis of a study. There are three causes of residual confounding:

1. There were additional confounding factors that were not considered, or there was no attempt to adjust for them, because data on these factors was not collected.
2. Control of confounding was not tight enough. For example, a study of the association between physical activity and age might control for confounding by age by a) restricting the study population to subject between the ages of 30-80 or b) matching subjects by age within 20 year categories. In either event there might be persistent differences in age among the groups being compared. Residual differences in confounding might also occur in a randomised clinical trial if the sample size was small. In a stratified analysis or in a regression analysis there could be residual confounding because data on confounding variable was not precise enough, e.g., age was simply classified as "young" or "old".
3. There were many errors in the classification of subjects with respect to confounding variables.

3.4.2 Confounding by Indication

Confounding by indication is a special type of confounding that can occur in observational (non-experimental) pharmaco-epidemiologic studies of

the effects and side effects of drugs. This type of confounding arises from the fact that individuals who are prescribed a medication or who take a given medication are inherently different from those who do not take the drug, because they are taking the drug for a reason. In medical terminology, such individuals have an "indication" for use of the drug. Even if the study population consists of subjects with the same disease, e.g., osteoarthritis, they may differ in the severity of their disease and may therefore differ in the need for medication. Aschengrau and Seage *(2018)* give the example of studies of the association between antidepressant drug use and infertility. The use of antidepressant medications may appear to be associated with an increased risk of infertility. However, depression itself is a known risk factor for infertility. As a result, there would appear to be an association between antidepressants and infertility. One way of dealing with this is to study the association in subjects who are receiving different treatments for the same underlying disease condition. A variation on this might be dubbed "***confounding by contraindication***." For example, in the case-control study by Perneger and Whelton (2005) examining the association between analgesic drug use and kidney failure the authors compared prior analgesic use between patients receiving kidney dialysis and population controls without known kidney disease. Suppose that patients on dialysis had been advised to avoid taking aspirin because of its effects on blood clotting; they may have been advised to take acetaminophen (Tylenol) instead). If the group of dialysis cases included a number of people who had been on long-term dialysis, this would result in a decreased frequency of aspirin use and increased use of Tylenol in the case group. As a result, an association with aspirin would be underestimated, while an association with Tylenol would be overestimated.

3.4.3 Reverse Causality

Reverse causality occurs when the probability of the outcome is causally related to the exposure being studied. For example, Child feeding recommendations of the World Health Organisation include breastfeeding for two years or more, because of evidence that breast fed children have a reduced risk of infectious agents and are less likely to die. However, some studies have produced conflicting concerns. One possibility is that in communities with very poor resources the children who are at greatest risk and perhaps have the least access to other food sources are more likely to be breast fed for at least two years. A comparison of growth and development between these children and more advantaged children would likely find less progress in the breast fed group. (See "Association of Breastfeeding and Stunting in Peruvian Toddlers: An Example of Reverse Causality" by Marquis GS, et al.: International Journal of Epidemiology 1997; 26: 349–356.

The case-control study by Perneger and Whelton (2005) may also have been affected by reverse causality. Diabetes is a leading cause of renal failure in the US, and chronic diabetes is associated with a number of other health problems such as cardiovascular diseases and infections that could result in a greater use of analgesics. If so, the dialysis cases whose renal failure resulted from diabetes might have taken more analgesics because of their diabetes. Nevertheless, it would appear that analgesic use was associated with an increased risk of renal failure rather than vice versa.

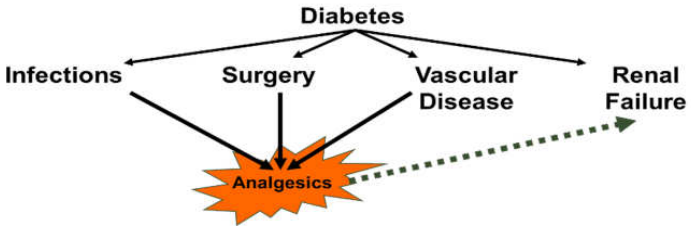


Figure 9: Associated between analgesic use and an increased risk of renal failure: (Source Whelton 2005)

3.5 Control of Confounding in Study Design

3.5.1 Restriction

One of the conditions necessary for confounding to occur is that the confounding factor must be distributed unequally among the groups being compared. Consequently, one of the strategies employed for avoiding confounding is to restrict admission into the study to a group of subjects who have the same levels of the confounding factors. For example, in the hypothetical study looking at the association between physical activity and heart disease, suppose that age and gender were the only two confounders of concern. If so, confounding by these factors could have been avoided by making sure that all subjects were males between the ages of 40-50. This will ensure that the age distributions are similar in the groups being compared, so that confounding will be minimised.

This approach to controlling confounding is simple and effective, but it has several limitations:

i. It reduces the number of subjects who are eligible (may cause sample size problem).
ii. Residual confounding can occur if you don't restrict narrowly enough. For example, in the study on exercise and heart disease, the investigators might have restricted the study to men aged 40-65. However, the age-related risk of heart disease still varies widely within this range as do levels of physical activity.

iii. You can't evaluate the effects of factors that have been restricted for. For example, if the study is limited to men aged 45-50, you can't use this study to examine the effects of gender or age (because these factors don't vary within your sample).
iv. Restriction limits generalisability. For example, if you restrict the study to men, you may not be able to generalise the findings to women.

3.5.2 Matching

Instead of restriction, one could also ensure that the study groups do not differ with respect to possible confounders such as age and gender by matching the two comparison groups. For example, for every active male between the ages of 40-50, we could find and enroll an inactive male between the ages of 40-50. In this way, the groups we are comparing can artificially be made similar with respect to these factors, so they cannot confound the relationship. This method actually requires the investigators to control confounding in both the design and analysis phases of the study, because the analysis of matched study groups differs from that of unmatched studies. Like restriction, this approach is straightforward, and it can be effective. However, it has the following disadvantages:

i. It can be time-consuming and expensive.
ii. It limits sample size.
iii. You can't evaluate the effect of the factors you that you matched for.

Nevertheless, matching is useful in the following circumstances:

i. When one needs to control for complex, multifaceted variables (e.g., heredity, environmental factors)
ii. When doing a case-control study in which there are many possible controls, but a smaller number of cases (e.g., 4:1 matching in the study examining the association between DES and vaginal cancer)

3.5.3 Randomisation in Clinical Trials

You previously studied randomisation in the online module on Clinical Trials. Given the more detailed discussion in this current module of the conditions necessary for confounding to occur, it should be obvious why randomisation is such a powerful method to control prevent confounding. If a large number of subjects are allocated to treatment groups by a random method that gives an equal chance of being in any treatment group, then it is likely that the groups will have similar distributions of age, gender, behaviours, and virtually all other known and as yet unknown possible confounding factors. Moreover, the investigators can

get a sense of whether randomisation has successfully created comparability among the groups by comparing their baseline characteristics.

3.5.4 Control of Confounding in the Analysis - Stratified Analysis

One way of identifying confounding is to examine the primary association of interest at different levels of a potential confounding factor. The side by side tables below examine the relationship between obesity and incident CVD in persons less than 50 years of age and in persons 50 years of age and older, separately.

Table 25: Obesity and Incident of Cardiovascular Diseases by Age Group

	Age ‹ 50				**Age ≥ 50**		
	CVD	No CVD	Total		CVD	No CVD	Total
Obese	10	90	100	Obese	36	164	200
Not Obese	35	465	500	Not Obese	25	175	200
Total	45	555	600	Total	61	339	400

The ***stratum-specific risk ratios*** are as follows:

- Among those < **50**, the risk ratio is:

$$RR = \frac{10/100}{35/500}$$

$$= 0.10/0.07 = 1.43$$

- Among those ≥ **50**, the risk ratio is:

$$RR = \frac{36/200}{25/200}$$

$$= 0.18/0.125 = 1.44$$

Recall that the risk ratio for the total, combined sample was RR = 1.79; this is sometimes referred to as the "crude" measure of association, because it is not adjusted for potential confounding factors. The risk ratios for the age-stratified analysis are similar (RR = 1.43 and 1.44, respectively), but less than the crude risk ratio. This indicates that there was confounding by age in the overall sample. We saw that obese subjects were more likely to be 50 and older, and we also saw that those over age 50 had a greater risk of CVD. As a result, the crude analysis

overestimated the true association between obesity (per se) and CVD, because of the greater proportion of older subjects among the obese group.

Several things are noteworthy in this example. First, if you compare the cumulative incidence in young versus old active subjects, you can see that older subjects had a higher risk of CVD than younger subjects; this was true for both obese and non-obese subjects. Therefore, age and CVD (the outcome of interest) are associated. In addition, obesity was more common in older subjects, meaning that age and obesity were also associated. Finally, there is no reason to think that age is an intermediary variable in the causal chain between obesity and CVD. Therefore, these observations satisfy all three of the requirements for a confounder.

Comparing the crude and stratum-specific measures of association is a very practical way to determine whether confounding is present and how bad it is. You calculate an overall crude (unadjusted) relative risk (or odds ratio) and compare it to the stratum-specific relative risks (or odds ratios). If the stratum-specific measures of association are similar to the crude measure of association, then there is no confounding by that factor, and you can just use the crude measure of association. However, if the stratified estimates of association differ from the unadjusted estimate by 10% or more, then there is evidence of confounding.

3.5.4.1The Cochran-Mantel-Haenszel Method

In the example above we saw that the relationship between obesity and CVD was confounded by age. When the data was pooled, it appeared that the risk ratio for the association between obesity and CVD was 1.79. However, when we stratified the analysis into those age <50 and those age 50+, we saw that both groups had a risk ratio of about 1.43. The distortion was due to the fact that obese individuals tended to be older, and older age is a risk factor for CVD. Consequently, in the analysis using the combined data set, the obese group had the added burden of an additional risk factor.

The Cochran-Mantel-Haenszel method is a technique that generates an estimate of an association between an exposure and an outcome after adjusting for or taking into account confounding. The method is used with a dichotomous outcome variable and a dichotomous risk factor. We stratify the data into two or more levels of the confounding factor (as we did in the example above). In essence, we create a series of two-by-two tables showing the association between the risk factor and outcome at two or more levels of the confounding factor, and we then compute a weighted average of the risk ratios or odds ratios across the strata (i.e., across subgroups or levels of the confounder).

3.5.4.2Data Layout for Cochran-Mantel-Haenszel Estimates

Before computing a Cochran-Mantel-Haenszel Estimate, it is important to have a standard layout for the two by two tables in each stratum. We will use the general format depicted here:

Tables 26: Data Layout for Cochran-Mantel-Haenszel Estimates

	Outcome Present	**Outcome Absent**	**Total**
Risk Factor Present (Exposed)	A	B	a+b
Risk Factor Absent (Unexposed)	C	D	c+d
	a+c	b+d	N

Commented [U1]:

Using the notation in this table estimates for a risk ratio or an odds ratio would be computed as follows:

- **Risk Ratio: RR= a/(a+b)**
  c/(c+d)
- **Odds Ratio: OR = a/b= ad**
- **c/d bc**

3.5.4.3Cochran-Mantel-Haenszel Equations

To explore and adjust for confounding, we can use a stratified analysis in which we set up a series of two-by-two tables, one for each stratum (category) of the confounding variable. Having done that, we can compute a weighted average of the estimates of the risk ratios or odds ratios across the strata. The weighted average provides a measure of association that is adjusted for confounding. The weighted averages for risk ratios and odds ratios are computed as follows:

i. Cochran-Mantel-Haenszel Estimate for a Risk Ratio

$$RR_{cmh} = \frac{\Sigma \frac{a_i(c_i+d_i)}{n_i}}{\Sigma \frac{a_i(c_i+d_i)}{n_i}}$$

ii. Cochran-Mantel-Haenszel Estimate for an Odds Ratio

$$OR_{cmh} = \frac{\sum \frac{a_i d_i}{n_i}}{\sum \frac{b i c_i}{n_i}}$$

Where a_i, b_i, c_i, and d_i are the numbers of participants in the cells of the two-by-two table in the w^{ith} stratum of the confounding variable, and n_i represents the number of participants in the i^{th} stratum.
To illustrate the computations, we can use the previous example examining the association between obesity and CVD, which we stratified into two categories: those with age ‹50 and those who were ≥50 at baseline:

Table 27: Obesity and Incidence of Cardiovascular Diseases by Age Group

	Age < 50				**Age ≥ 50**		
	CVD	No CVD	Total		CVD	No CVD	Total
Obese	10	90	100	Obese	36	164	200
Not Obese	35	465	500	Not Obese	25	175	200
Total	45	555	600	Total	61	339	400

i. Among those <**50**, the risk ratio is:

$$RR = \frac{10/100}{35/500}$$

= **1.43**

ii. Among those ≥ **50**, the risk ratio is:

$$RR = \frac{36/200}{25/200}$$

= **1.44**

From the stratified data we can also compute the Cochran-Mantel-Haenszel estimate for the risk ratio as follows:
Using above formula the answer will be equal to RR_{cmh} $\frac{8.33+18.00}{5.83+12.50} = 1.44$

If we chose to, we could also use the same data set to compute a **crude odds ratio** (crude OR = 1.93) and we could also compute **stratum-specific odds ratios** as follows:

i. Among those ‹**50**, the risk ratio is:

$$OR = \frac{4650}{} = 1.48$$

3150

ii. Among those ≥ **50**, the risk ratio is:

$OR = \frac{a/c}{b/d} = \frac{ad}{bc} = \frac{36(175)}{164(25)} = \frac{6300}{4100} = 1.54$

And, using the same data we could also compute the **Cochran-Mantel-Haenszel estimate for the odds ratio** as follows:

$OR_{cmh} = \frac{775+15.75}{5.25+10.25} = 1.52$

The Cochran-Mantel-Haenszel method produces a single, summary measure of association which provides a weighted average of the risk ratio or odds ratio across the different strata of the confounding factor. Notice that the adjusted relative risk and adjusted odds ratio, 1.44 and 1.52, are not equal to the unadjusted or crude relative risk and odds ratio, 1.78 and 1.93. The adjustment for age produces estimates of the relative risk and odds ratio that are much closer to the stratum-specific estimates (the adjusted estimates are weighted averages of the stratum-specific estimates).

3.5.4.4 Cochran-Mantel-Haenszel for Incidence Rates

Note that there is also a Cochran-Mantel-Haenszel equation which can be used when dealing with incidence rates in prospective studies in which incidence rates are computed.
The general format is depicted here:

Table 28: Cochran-Mantel-Haenszel for Incidence Rates

	Outcome Present	**Person-Time**
Risk Factor Present (Exposed)	A	PT_e
Risk Factor Absent (Unexposed)	C	PT_0
Total		PT_T

Using the notation in this table estimates for an incidence rate ratio would be computed as follows:

$$\text{Pooled Rate Ratio} = \frac{\Sigma a_i \frac{(PT0)}{PTi}}{\Sigma c_i \frac{(PTei)}{PT_{Ti}}}$$

Where for each stratum, a_i= number of exposed cases, ci=number of unexposed cases, PTei and PT0i are the person-time for exposed and unexposed groups respectively, and PT_{Ti} is the total person-time in each stratum.

3.5.4.5 More Than Two Sub-strata

In the examples above we used just two levels or sub-strata or of the confounding variable, but one can use more than two sub-strata. This is particularly important when using stratification to control for confounding by a continuously distributed variable like age. In the example above looking at the relationship between obesity and CVD we stratified the analysis by age, looking at the relationship in subjects <50 years and those who were 50+. However, subjects <50 years are likely to vary substantially with respect to BMI and rates of CVD; the same is true for subjects of age 50+. By stratifying into just two broad age groups, we would likely have a problem with residual **confounding**. To deal with this, we could stratify by age at 5-year intervals.

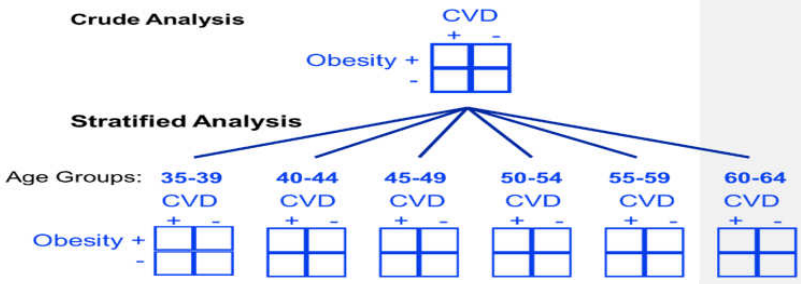


Figure 10: More Than Two Sub-strata by age; Source Cochran-Mantel-Haenszel 2016

3.5.4.6 Stratification to Control for Two or More Factors

In looking at the relationship between exercise and heart disease we were also concerned about confounding by other factors, such as gender and the presence of a family history of heart disease. We could also stratify by these factors to see if they were confounders and to adjust for them.

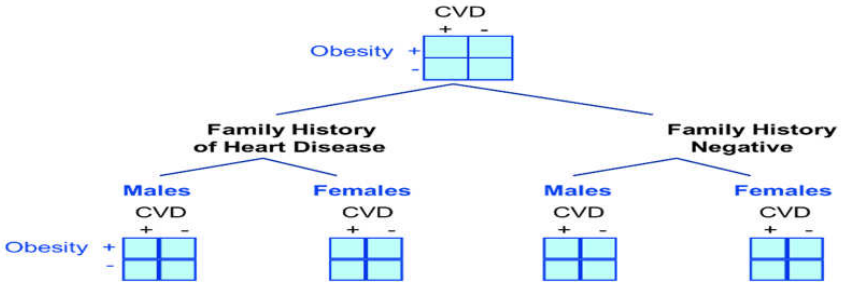


Figure 26: Stratification to Control for gender and family history: Source Cochran-Mantel-Haenszel 2016

3.5.4.7 Limitations of Stratified Analysis

A stratified analysis is easy to do and gives you a fairly good picture of what's going on. However, a major disadvantage to stratification is its inability to control simultaneously for multiple confounding variables. For example, you might decide to control for gender, 3 levels of smoking exposure, 4 levels of age, and 4 levels of BMI. This would require 96 separate strata to control for all of these variables simultaneously, and as you increase the number of strata, you keep whittling away at the number of people in each stratum, so sample size becomes a major problem, since many of the strata will contain few or no people.

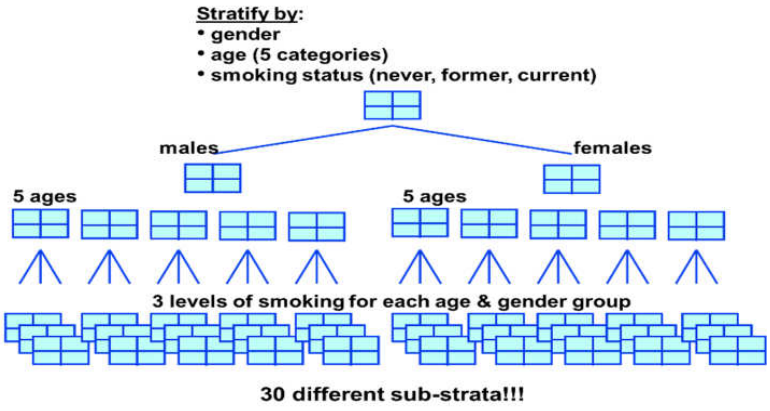


Figure 27: Stratification by gender, age and smoking status: (Source Cochran-Mantel-Haenszel 2016)

3.6 Effect Measure Modification

The term Effect Measure Modification (EMM) is applied to situations in which the magnitude of the effect of an exposure of interest differs depending on the level of a third variable. Reye's syndrome is a rare, but severe condition characterised by the sudden development of brain

damage and liver dysfunction after a viral illness. The syndrome is most commonly seen in children between the ages of 4-14 who have been treated with aspirin while recovering from a viral illness, most commonly chickenpox or influenza. Fortunately, Reye's syndrome has become very uncommon since aspirin is no longer recommended for routine use in children. While Reye's syndrome can occur in adults, it is distinctly more common in children. Thus, the effect of aspirin treatment for a viral illness is very clearly modified by age.

In this situation, computing an overall estimate of association is misleading. One common way of dealing with effect modification is examine the association separately for each level of the third variable. For example, if one were to calculate the odds ratio for the association between aspirin treatment during a viral infection and development of Reye's syndrome, the odds ratio would be substantially greater in children than in adults. As another example, suppose a clinical trial is conducted and the drug is shown to result in a statistically significant reduction in total cholesterol. However, suppose that with closer scrutiny of the data, the investigators find that the drug is only effective in subjects with a specific genetic marker and that there is no effect in persons who do not possess the marker. The effect of the treatment is different depending on the presence or absence of the genetic marker. This is an example of effect modification or "statistical interaction".

3.6.1 Effect Modification with a Continuous Outcome: Evaluation of a Drug to Increase HDL Cholesterol

Consider the following clinical trial conducted to evaluate the efficacy of a new drug to increase HDL cholesterol (the "good" cholesterol). One hundred patients are enrolled in the trial and randomised to receive either the new drug or a placebo. Background characteristics (e.g., age, sex, educational level, income) and clinical characteristics (e.g., height, weight, blood pressure, total and HDL cholesterol levels) are measured at baseline, and they are found to be comparable in the two comparison groups. Subjects are instructed to take the assigned medication for 8 weeks; at which time their HDL cholesterol is measured again. The results are shown in the Table 29 below.

Table 29: Evaluation of a Drug to Increase HDL Cholesterol

	Sample Size	**Mean HDL**	**Standard Deviation of HDL**
New Drug	50	40.16	4.46
Placebo	50	39.21	3.91

On average, the mean HDL levels are 0.95 units higher in patients treated with the new medication. A two-sample test to compare mean HDL levels between treatments has a test statistic of Z = -1.13 which is not statistically significant at α=0.05.

Based on their preliminary studies, the investigators had expected a statistically significant increase in HDL cholesterol in the group treated with the new drug, and they wondered whether another variable might be masking the effect of the treatment. Other studies had, if fact, suggested that the effectiveness of a similar drug was different in men and women. In this study, there are 19 men and 81 women. Table 30 below shows the number and percent of men assigned to each treatment.

Table 30: number and percent of men assigned to each treatment

	Sample Size	**Number of Men (%)**
New Drug	50	0 (20%)
Placebo	50	9 (18%)

There is no meaningful difference in the proportions of men assigned to receive the new drug or the placebo, so sex cannot be a confounder here, since it does not differ in the treatment groups. However, when the data are stratified by sex, they find the following:

Table 31: Stratification by sex

WOMEN	**Sample Size**	**Mean HDL**	**Standard Deviation of HDL**
New Drug	40	38.88	3.97
Placebo	41	39.24	4.21
MEN	**Sample Size**	**Mean HDL**	**Standard Deviation of HDL**
New Drug	10	45.25	1.89
Placebo	9	39.06	2.22

On average, the mean HDL levels are very similar in treated and untreated women, but the mean HDL levels are 6.19 units higher in men treated with the new drug. This is an example of effect modification by sex, i.e., the effect of the drug on HDL cholesterol is different for men and women. In this case there is no apparent effect in women, but there appears to be a moderately large effect in men. (Note, however, that the comparison in men is based on a very small sample size, so this difference should be interpreted cautiously, since it could be the result of random error or confounding.

When there is effect modification, analysis of the pooled data can be misleading. In this example, the pooled data (men and women combined), shows no effect of treatment. Because there is effect modification by sex, it is important to look at the differences in HDL levels among men and women, considered separately. In stratified analyses, however, investigators must be careful to ensure that the sample size is adequate to provide a meaningful analysis.

3.6.2 Effect Modification with a Dichotomous Outcome

Consider the following hypothetical study comparing hospitalization after a motor vehicle collision for male and female drivers.

Table 32: Crude Data

	Hospitalised	Not Hospitalized	Total
Male	1330	7018	8348
Female	798	6400	7198

Crude risk ratio=1.44
Age-Stratified: Age ‹40

Table 32: Age-Stratified: Age ‹40

	Hospitalised	Not Hospitalized	Total
Male	966	3146	4112
Female	460	3000	3450

Stratum-specific risk ratio=1.80

Table 33: Age-Stratified: ≥40

	Hospitalised	Not Hospitalised	Total
Male	364	3872	4236
Female	348	3400	3748

Stratum-specific risk ratio=0.93

In this case, the crude analysis suggests an association between male gender and frequency of hospitalisation for motor vehicle collisions. However, if we stratify this by age, we see a strong association with male gender in subjects <40 years old, but no association in subjects 50+. Perhaps males <40 years old driver more recklessly than their female counterparts, but after age 40 driving aggression becomes similar in males and females.

Another good example of effect modification is seen with skin cancers. It is well established that excessive exposure to UV irradiation increases

one's risk of skin cancer. However, the risk of UV-induced skin cancer is 1,000 times greater in people with xeroderma pigmentosum. This is a rate hereditary defect (autosomal recessive) in the enzyme system that repairs UV-induced damage to DNA. It is characterised by photosensitivity, pigmentary changes, premature skin aging, and greatly increased susceptibility to malignant tumor development.

If effect modification is present, it is **NOT** appropriate to use Mantel-Haenszel methods to combine the stratum-specific measures of association into a single pooled measurement. Effect modification is a biological phenomenon that should be described, so the stratum-specific estimates should be reported separately. In contrast, confounding is a distortion of the true association caused by an imbalance of some other risk factor.

i. **If there is only confounding:** The stratum-specific measures of association will be similar to one another, but they will be different from the overall crude estimate by 10% or more. In this situation, one can use Mantel-Haenszel methods to calculate a pooled estimate (RR or OR) and p-value.

ii. **If there is neither confounding nor effect modification:** The crude estimate of association and the stratum-specific estimates will be similar. They don't have to be identical, just similar.

iii. **If there is only effect modification:** The stratum-specific estimates will differ from one another significantly. Whether they are "significantly different" can be tested by using a chi-square test of homogeneity, as described in the Aschengrau & Seage textbook.

iv. **If there is both effect modification and confounding:** Here, you need to consider two possibilities:

1) If the stratum-specific estimates differ from one another, and they are both less than the crude estimate or if they are both greater than the crude estimate, then there is both confounding and effect modification.

2) If the stratum-specific estimates differ from one another, and the crude estimate is between the two stratum-specific estimates, then you need to pool the stratum-specific estimates (with a Mantel-Haenszel equation) to determine whether the pooled estimate is more than 10% different from the crude estimate.

Note that in this situation you are only pooling the stratum-specific estimates in order to make a decision about whether confounding is present; you should **not** report the pooled estimate as an "adjusted" measure of association if there is effect modification

3.6.3 Statistical Interaction versus Biological Interaction

While the discussion above provides a standard description of effect modification, but on closer scrutiny the concept of effect modification is more complicated than this. Consider the figure below (adapted from KJ Rothman: Epidemiology - An Introduction, Oxford University Press, 2002.) We see two scenarios in which incidence rates in exposed and unexposed individuals are assessed at different ages. Rate ratio and rate difference are both measures of effect, but depending on which we use, our conclusions about effect modification differ.

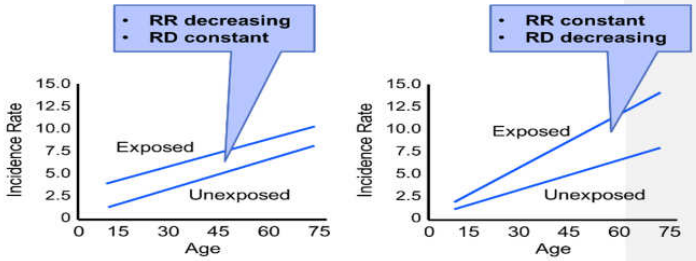


Figure 28: Statistical Interaction versus Biological Interaction
(Source: Rothman, 2002)

In the first scenario the rate difference remains constant across the spectrum of age, suggesting no effective modification. However, the rate ratio decreases with increasing age (RR=3 at age 15; RR=1.5 at age 75). In the second scenario the rate ratio remains relatively constant, but the rate difference increases with age. Our conclusion regarding whether or not there is effect modification will depend on which measure of effect we use.

Consider also the hypothetical data on the risk of lung cancer in smokers and non-smokers, both with and without exposure to asbestos (also adapted from Rothman).

Table 34: Hypothetical 1-Year Risk of Lung Cancer per 100,000

	Without Asbestos	**With Asbestos Exposure**
Smokers	5	50
Non-smokers	1	10

First consider the effect of asbestos on the risk associated with smoking. The risk ratio is 5 both with and without asbestos exposure, suggesting no effect modification. However, the risk difference 4 per 100,000 without asbestosis and 40 per 100,000 with asbestosis exposure. This

effect measure is clearly modified by asbestos. We can also look at the effect of smoking on the risk associated with asbestos. The risk ratio for asbestos exposure compared to no asbestos exposure is 10 in both smokers and non-smokers, suggesting an absence of effect modification. However, the risk difference is 45 per 100,000 in the presence of smoking, but only 9 per 100,000 in the absence of smoking. Thus, the risk ratios suggest no effect modification, but the risk differences suggest substantial effect modification.

Rothman (2002) argues that this ambiguity regarding effect measure modification and statistical interaction makes it important to make a distinction between statistical interaction (which is ambiguous) and biological interaction (which is not ambiguous; it is either present or absent.) Biological interaction between two causes occurs if the effect of one is dependent on the presence of the other. For example, exposure to the measles virus is a component cause of developing measles, but it is dependent on another factor, i.e., the immune status of the exposed individual. Someone who is immune because of vaccination or having already had measles will not experience any effect from exposure to the measles virus. A discussion of the methods for measuring biological interaction is beyond the scope of this module. Those who are interested should refer to the discussion in Rothman's excellent text.

SELF ASSESSED EXERCISE

The director of the surgical trauma service at Boston Medical Center suspected that elderly drivers (age 70+) had inordinately poor outcomes compared to younger drivers after being in a motor vehicle collision (MVC). His research hypothesis was tested using data from the Boston Medical Center Trauma registry and data from the National Trauma Data Bank.

1. Are there any factors that might confound the association between being an elderly driver and the risk of death after a motor vehicle collision? If so, what factors would you consider? How would you deal with these potential confounders?
2. The figure below summarises some of the data obtained from the Boston Medical Center Trauma registry. The upper contingency table shows deaths among the 74 elderly drivers hospitalised after an MVC and the 960 younger drivers who had been hospitalised after an MVC. The lower two tables summarise the findings after stratifying based on whether the drivers had the benefit of safety devices (seat belt buckled and/or air bag in the vehicle. Do these findings suggest the presence of effect modification? Why or why not?

Table 35: Crude Analysis

	Died	Lived	Total
Age ≥70	13	61	74
Age‹70	25	935	960

Table 36: Stratified by Use of a Safety Restraint: Unrestrained (no seatbelt or air bag)

	Died	Lived	Total
Age ≥70	8	16	24
Age‹70	13	359	372

Table 37: Restrained with Seatbelt, Air bag, or Both

	Died	Lived	Total
Age ≥70	5	45	50
Age‹70	12	576	588

4.0 CONCLUSION

In the unit you have learn detail on confounding as a form of bias, you also know and learn how it can be introduce in a study when there is distortion of the association between an exposure and an outcome that occurs when the study groups differ with respect to other factors that influence the outcome. In the unit you learnt how it can be adjusted can be adjusted for in the analysis, provided that the investigators have information on the status of study subjects with respect to potential confounding factors.

5.0 SUMMARY

Confounding is a distortion (inaccuracy) in the estimated measure of association that occurs when the primary exposure of interest is mixed up with some other factor that is associated with the outcome. For example, in heart disease in an elderly person. Age, is a confounding factor because it is associated with the exposure (meaning that older people are more likely to be inactive), and it is also associated with the outcome (because older people are at greater risk of developing heart disease). Three conditions that must be present for confounding to occur include the confounding factor must be associated with **both** the risk factor of interest and the outcome. The confounding factor must be distributed unequally among the groups being compared and A confounder cannot be an intermediary step in the causal pathway from the exposure of interest to the outcome of interest.

6.0 TUTOR-MARKED ASSIGNMENT (TMA)

1. List three conditions necessary for a confounding to take place in

a study.

2. How will you measure the magnitude of confounding?
3. Write a short note on the following
   - Residual Confounding
   - Confounding by Indicator
   - Reversal Causality

7.0 REFERENCES/FURTHER READING

Marsden-Haug, N., Foster, V.B., Gould, P.L., Elbert, E., Wang, H., & Pavlin, J.A. (2017). Code-based syndromic surveillance for influenza like illness by International Classification of Diseases, ninth revision. *Emerg Infect Dis*,13(2), 207-216.

Scallan, E., Hoekstra, R.M., &Angulo, F.J. (2011). Foodborne Illness Acquired in the United States - Major Pathogens. *Emerging Infectious Diseases*,17(1),7-15.

Trifonov, V., Khiabanian, H., & Rabadan, R. (2009). Geographic Dependence, Surveillance, and Origins of the 2009 Influenza A (H1N1) Virus. **Perspective article in:** *N. Engl. J. Med*, 361(2), 115-119.

UNIT 3 MEASURE OF STANDARDISATION IN EPIDEMIOLOGY

CONTENTS

1.0 INTRODUCTION

A principal role in epidemiology is to compare the incidence of disease or mortality between two or more populations. However, the comparison of crude mortality or morbidity rates is often misleading because the populations being compared may differ significantly with respect to certain underlying characteristics, such as age or sex that will affect the overall rate of morbidity or mortality. For example, age is an important determinant of mortality. An older population will have a higher overall mortality rate than a younger population. As a result, variations in age will complicate any comparison between two or more populations that have different age structures. To understand how a comparison of crude rates can be affected by differing population distributions, it should be recognised that a crude overall rate is simply a weighted average of the individual category specific rates, with the weights being the proportion of the population in each category.

As evident, crude rates are indifferent to differences that usually exist by age, sex, race, or some category of pre-existing disease conditions (confounding factors/variable). Crude rates from different populations cannot be meaningfully compared. Adjustment for structural differences is necessary. But adjusted rates must not be mistaken as real. They are fictional figures used to make valid summary comparisons between two or more groups possessing dissimilar age or other structural distribution and exposure characteristics such as age, sex, race, income, smoking status, diet, and indeed exposure to various risk factors of diseases.

Adjusted rates are standardised summary figures for a defined population by which statistical procedures are carried out to remove the effect of differences in composition of various populations; thus permitting unbiased comparison. Adjustments of rates are carried out to control or neutralise the influence of socio-demographic, socio-economic, and other exposure or susceptibility variables that are known to be strongly associated with the risk of diseases or other health status outcome. Such rates are relative and their absolute magnitude depends on standard population chosen.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- Identify direct standardisation
- describe indirect standardisation
- issues in the use of standardisation

3.0 MAIN CONTENT

3.1 Definition of Standardisation

This unit introduces you to standardisation, a method for overcoming the effect of confounding variables in epidemiological research. Most commonly, standardisation is used to control for age. There are two methods of standardisation, direct and indirect, and both are explained here. The resource in the unit covers methods of standardisation and issue in their use.

3.2 Presentation of Category Specific Rates

One method of overcoming the effects of confounding variables such as age is to simply present and compare the age specific rates. While this allows for a more comprehensive comparison of mortality or morbidity rates between two or more populations, as the number of stratum specific rates being compared increases, the volume of data being examined may become unmanageable. It is, therefore, more useful to combine category specific rates into a single summary rate that has been adjusted to take into account its age structure or other confounding factor. This is achieved by using the methods of standardisation

3.3 Methods of Standardisation

There are two methods of standardisation commonly used in epidemiological studies, and these are characterised by whether the standard used is a population distribution **(direct method)** or a set of

specific rates **(indirect method).** Both direct and indirect standardisation involves the calculation of numbers of expected events (e.g. deaths), which are compared to the number of observed events. Age is a factor that is frequently adjusted for in epidemiological investigations, particularly in comparative mortality studies, since the age structure of a population will greatly affect the population's overall mortality. To illustrate the methods of both direct and indirect standardisation, the age specific mortality rates for two hypothetical populations are compared below.

3.3.1 Direct Method of Standardisation

Table 1 presents crude mortality data for two hypothetical populations (countries A and B). The overall crude mortality rate is higher for country A (10.5 deaths per 1,000 person years) compared with country B (7 deaths per 1,000 person years), despite the age-specific mortality rates being higher among all age-groups in country B.

Table 38: Crude mortality rates stratified by age for two hypothetical populations

				Country B			
Country A							
Age Group	**No of Death**	**Population**	**Rate per 1,000**	**Age Group**	**No of Death**	**Population**	**Rate per 1,000**
0-29	7,000	6,000,000	1.5	0-29	6,300	120,000	4.2
30-59	20,,000	2,500,000	3.9	30-59	3,,000	550,000	5.5
60+	120,00	5,500,000	48	60+	6,000	1,500,000	50
Total	147,000	14,000,000	10.5	Total	12,300	5,170,000	7.0

Table 39 shows the direct method of standardisation - calculation of the number of expected deaths for countries A and B applied to a standard population. (Here the rate is divided back by 1000 to give the basic rate; e.g. 1.2 becomes 0.0012 for the purposes of the formula).

Table 39: Direct method of standardisation

	Country A Expected deaths	Country B Expected deaths
0-29	0,0012x100,000=120	0.0042x100,000=420
30-59	0.0036x65,000=234	0.0055x65,000=357.5
60+	0.048 x20,000=960	0.05x20,000=1,000

Total Expected deaths	1.314	1.777.5
Age Adjusted rate	1.314/185,000=7.1 per 1,000pyrs	1.777.5/185,000=9.6 per 1,000pyrs

Age Standard rate ration (B:A)= 9.6/7.1-1.35

The reason for the difference between the crude mortality rates between country A and country B is that these two populations have markedly different age-structures. Country A has a much older population than country B. For example, 18% of the population in country A are aged over 60 years compared with 6% in country B.

Table 40: Standard population

0-29	100,000
30-59	65,000
60+	20,000
Total	185,000

In the direct method of standardisation, 'age adjusted rates' are derived by applying the category specific mortality rates of each population to a single standard population (Table 40). This produces age standardised mortality rates that these countries would have if they had the same age distribution as the standard population. Note that the 'standard population' used may be the distribution of one of the populations being compared or may be an outside standard population such as the 'European' or 'World' standard population.

The weighted average of the category-specific rates (with the weights taken from the standard population) provides for each population a single summary rate that reflects the numbers of events that would have been expected if the populations being compared had the same age distribution 1. The ratio of the directly standardised rates can then be calculated to provide a single summary measure of the difference in mortality between the two populations. The ratio of the standardised rates is called the Comparative Mortality Ratio (CMR) and is calculated by dividing the overall age adjusted rate in country B by the rate in country A.

For example: **Comparative Mortality Ratio = 9.6/7.1 = 1.35**

This CMR is interpreted as: after controlling for the confounding affects of age, the mortality in Country B is 35% higher than in country A.

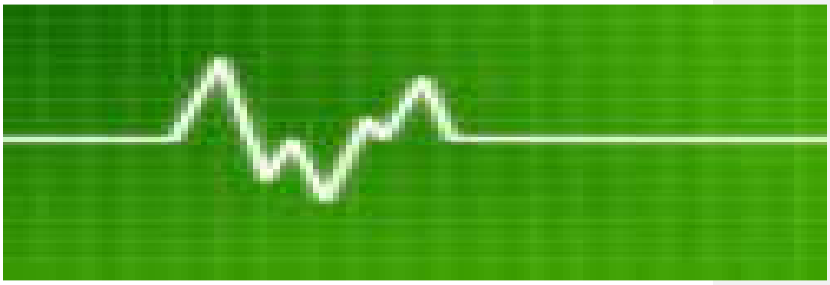


Figure 29: Interrupted Comparative Mortality Ratio (Source: Jones, 2004)

Note that while the crude rates presented in Table 38 represent the actual mortality experience of countries A and B, it is not possible to use these crude rates to make a valid comparison between the two countries because they have very different age distributions. However, by using the direct method of standardisation (while the values of the adjusted rate do not reflect the 'true' mortality experience of countries A and B), it enables us to calculate 'hypothetical' age adjusted rates that can be used to make a valid comparison of overall mortality between the two countries.

3.3.2 Indirect Method of Standardisation

The indirect method of standardisation is commonly used when age-specific rates are unavailable. For example, if we did not know the age specific mortality rates for Community B.

In this method, instead of taking one population structure as standard and applying sets of rates to it to estimate expected events, a set of rates from a standard population (country A) is applied to each of the populations being compared to calculate standardised morbidity/mortality ratios.
Table 41 shows the indirect method of standardisation: Number of expected deaths if the population had the same age-specific mortality rates as Community A.

Table 41: Indirect method of standardisation

	Community A Expected Deaths	**Community B Expected Deaths**
0-29	0.0012 x6,000,000=7,200	0,0012x 1,500,0000=1,800
30-59	0.0036x5,500 x 5,500,000=19,800	0.0036 x 550,000=1,9800
60+	0,048 x2,500,000 =120,000	0.048 x 120,000=5,760
Total Expected deaths (F)	147,000	9,540

Total Observed deaths (O)	147,000	15,300
Standardization Mortality Ratio O/E x 100	100	160

In table above, the indirect method of standardisation is used to calculate how many deaths would be expected in Country B if it had the same age-specific mortality rates as Country A. The expected deaths in Country B are calculated by multiplying the age specific rate for Country A by the population of Country B in the corresponding age group. The sum of the age categories gives the total number of deaths that would be experienced in country B if it had the same mortality experience as country A.
An overall summary measure can then be calculated, that is, the standardised mortality ratio (SMR), which is the ratio of the observed number of deaths to the expected number of deaths.

$$SMR = \frac{\text{Observed number of deaths (O) X 100\%}}{\text{Expected number of deaths (E)}}$$

$$SMR = \frac{160}{100} = 1.6 \text{ X } 100 = 160$$

From Table 41 the SMR is calculated as 160, which means that the number of observed deaths in Country B is 60% higher than the number we would expect if Country B had the same mortality experience as Country A.

3.3 Issues in the Use of Standardisation

i. Standardised rates are used for the comparison of two or more populations; they represent a weighted average of the age specific rates taken from a 'standard population' and are not actual rates.
ii. The direct method of standardisation requires that the age-specific rates for all populations being studied are available and that a standard population is defined.

iii. The indirect method of standardisation requires the total number of cases
iv. The ratio of two directly standardised rates is called the Comparative Incidence Ratio or Comparative Mortality Ratio.
v. The ratio of two indirectly standardised rates is called the Standardised Incidence Ratio or the Standardised Mortality Ratio.

vi. Indirect standardisation is more appropriate for use in studies with small numbers or when the rates are unstable.
vii. As the choice of a standard population will affect the comparison between populations, it should always be stated clearly which standard population has been applied.
viii. Standardisation may be used to adjust for the effects of a variety of confounding factors including age, sex, race or socio-economic status.

Standardisation of rates can be difficult to understand and is explained in several different ways depending on the literature source. It is recommended that as well as using this resource text you use one of the sources in the related links, choosing the text which gives you the appropriate amount of detail.

SELF-ASSESSED EXERCISE

i. How will you explain standardisation?
ii. List and explain important uses of standardisation.

4.0 CONCLUSION

In this unit you have learnt the meaning of standardisation, you also have the understanding that standardisation is a method for overcoming the effect of confounding variables in epidemiological research. You learnt method of measure standardisation both direct and in-direct method. Finally, you are exposed to Issues in the use of standardisation.

5.0 SUMMARY

Standardisation, a method for overcoming the effect of confounding variables in epidemiological research. Most commonly, standardisation is used to control for age. There are two methods of standardisation, direct and indirect, and both are explained in detail. The resource in the unit covers methods of standardisation and issue in their use.

6.0 TUTOR-MARKED ASSIGNMENT (TMA)

1. Give detailed account of direct method of standardisation.
2. Describe in detail how you will apply an indirect method of Standardisation.

7.0 REFERENCES/FURTHER READING

Farmer, R. & Lawrenson, R. (2014). Lecture notes in Epidemiology and Public Health Medicine. 67-68. Blackwell Publishing.

Hennekens, C.H. & Buring, J.E. (2007). Epidemiology in Medicine. Lippincott Williams & Wilkins,

Kirkwood, B.R. & Sterne, J.C. (2013). Essential Medical Statistics. 263-270. Blackwell Science.

NHS Public Health Network. (2018). Understanding Public Health Data [online] [accessed 05/06/2020].

Tauxe, R.V., Tormey, M.P., Mascola, L., Hargrett-Bean, N.T. & Blake, P.A. (1987). Salmonellosis outbreak on transatlantic flights; foodborne illness on aircraft. Am J Epidemiol, 125,150–7.

Tobler, W. A. (1997). Computer movie simulating urban growth in the Detroit region. Econ Geogr, 46(Suppl), 234–40.

Vazquez-Prokopec, G.M., Kitron, U., Montgomery, B., Horne, P. & Ritchie, S.A.(2010). Quantifying the spatial dimension of dengue virus epidemic spread within a tropical urban environment. PLoS Negl Trop Dis, 4:920-930.

White, F., Stallones, L. & Last, J. M. (2013). Global public health: Ecological foundations. New York, NY: *Oxford University Press*,23-30.

MODULE 3 VALID AND EFFICIENT EPIDEMIOLOGIC STUDIES THE TYPES, STRENGTH AND LIMITATION AND INTERPRETING RESULTS

Unit 1 Designing valid and efficient epidemiologic studies
Unit 2 Strength and limitation of epidemiological design
Unit 3 Interpret descriptive epidemiologic results in order to develop hypotheses of possible risk factors of a disease

UNIT 1 DESIGNING VALID AND EFFICIENT EPIDEMIOLOGIC STUDIES

1.0 Introduction
2.0 Objectives
3.0 Main Content
- 3.1 Explain the role of descriptive studies for identifying problems and establishing hypotheses.
  - 3.1.1 Specific tasks of descriptive epidemiology are the following:
  - 3.1.2 Hypothesis Formulation – Characteristics of Person, Place, and Time
  - 3.1.3 Cross-Sectional Studies
  - 3.1.4 Ecological Studies
- 3.2 Analytical Epidemiologic Studies
  - 3.2 1 Case-Control Studies
  - 3.2.2 Case-Crossover Studies
  - 3.2.3 Cohort Studies
- 3.3 Experimental Epidemiologic Studies
  - 3.3.1 Randomised Clinical Trials
- 3.4 Function Epidemiology study design
  - 3.4.1 Issues of Concern in Epidemiology study design
  - 3.4.2 Clinical Significance Epidemiology study design
  - 3.4.3 Other Issues in Epidemiology study design

4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment (TMA)
7.0 References/Further Reading

1.0 INTRODUCTION

In epidemiology, researchers are interested in measuring or assessing the relationship of exposure with a disease or an outcome. As a first step, they define the hypothesis based on the research question and then decide

which study design will be best suitable to answer that question. How the investigation is conducted by the researcher is directed by the chosen study design. The study designs can be broadly classified as experimental or observational based on the approach used to assess whether exposure and an outcome are associated. In an experimental study design, researchers assign – subjects to intervention and control/comparison groups in an attempt to isolate the effects of the intervention. Being able to control various aspects of the experimental study design enables the researchers to identify causal links between interventions sand outcomes of interest. In several instances, an experimental study design may not be feasible or suitable; in such situations, observational studies are conducted. Observational studies, as the name indicates, involve merely observing the subjects in a non-controlled environment without actually interfering or manipulating with other aspects of the study and therefore are non-experimental. The observation can be prospective, retrospective or current depending on the subtype of an observational study.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- explain the role of descriptive studies for identifying problems and establishing hypotheses
- identify observational studies design in term of case series, cross sectional, cohort studies and Ecological studies
- identify and describe experimental studies design
- discuss the functions, issues and clinical significance of studies design.

3.0 MAIN CONTENT

3.1 Explain the Role of Descriptive Studies for Identifying Problems and Establishing Hypotheses

Generally, descriptive epidemiology makes it possible to identify trends in health and disease and also provides a means of planning resources for populations. In addition, descriptive epidemiology is important for generating hypotheses (possible explanations) about the determinants of health and disease. By generating hypotheses, descriptive epidemiology also provides the starting point for analytic epidemiology, which formally tests associations between potential determinants and health or disease outcomes.

3.1.1 Specific tasks of descriptive epidemiology are the following

i. Monitoring and reporting on the health status and health related behaviors in populations
ii. Identifying emerging health problems
iii. Alerting us to potential threats from bioterrorism
iv. Establishing public health priorities for a population
v. Evaluating the effectiveness of intervention programs and
vi. Exploring potential associations between "risk factors" and health outcomes in order to generate hypotheses about the determinants of disease.

3.1.2 Hypothesis Formulation – Characteristics of Person, Place, and Time

Descriptive epidemiology searches for patterns by examining characteristics of ***person, place, & time***. These characteristics are carefully considered when a disease outbreak occurs, because they provide important clues regarding the source of the outbreak.
Hypotheses about the determinants of disease arise from considering the characteristics of person, place, and time and looking for differences, similarities, and correlations. Consider the following examples:

i. **Differences:** if the frequency of disease differs in two circumstances, it may be caused by a factor that differs between the two circumstances. For **example**, there was a substantial difference in the incidence of stomach cancer in Japan & the US. There are also substantial differences in genetics and diet. Perhaps these factors are related to stomach cancer.
ii. **Similarities:** if a high frequency of disease is found in several different circumstances & one can identify a common factor, then the common factor may be responsible. **Example**: AIDS in IV drug users, recipients of transfusions, & hemophiliacs suggests the possibility that HIV can be transmitted via blood or blood products.
i. **Correlations:** If the frequency of disease varies in relation to some factor, then that factor may be a cause of the disease.

Example: differences in coronary heart disease vary with cigarettes consumption.

3.1.3 Cross-Sectional Studies

Cross-sectional studies are observational in nature and give a snapshot of the characteristics of study subjects in a single point of time. Unlike

cohort studies, cross-sectional studies do not have a follow-up period and therefore are relatively simple to conduct. As the exposure status and outcome of interest information is collected in a single moment in time often by surveys, cross-sectional study design cannot provide cause-effect relationship and is the weakest of the observational designs. This study design is generally used to assess the prevalence of a disease in a population.

3.1.4 Ecological Studies

Ecological studies are used when data at an individual level is unavailable or when large-scale comparisons are needed to study the population-level effect of exposures on a disease condition. Therefore, ecological study results are applicable only at the population level. The types of measures in ecological studies are aggregates of individual-level data. These studies, therefore, are subject to a type of confounding called ecological fallacy which occurs when relationships identified at group level data are assumed to be true for individuals. Ecological studies are generally used in public health research.

3.2 Analytical Epidemiologic Studies

3.2.1 Case-Control Studies

Case-control studies are used to determine the degree of associations between various risk factors and outcomes. The factors that affect the risk of a disease are called exposures. Case-control studies can help identify beneficial or harmful exposures. In a case-control study, as the name suggests, there are two groups of subjects -cases and controls. Cases are subjects who have a particular disease, condition, or disability. Controls are those subjects that do not have the disease. Typically, researchers identify appropriate representative controls for the cases that they are studying from the general population. Then they retrospectively look in the past for the possible exposures these subjects might have had to a risk factor. Selecting the subjects for the control group is a very critical component of research based on case-control studies. Due to the retrospective nature of the study design, case-control studies are subject to recall bias. Case-control studies are inexpensive, efficient, and often less time consuming to conduct. This study design is especially suitable for rare diseases that have longer latency periods.

3.2.2 Case-Crossover Studies

Case-crossover studies are helpful to study triggers within an individual. When the researcher is studying a transient exposure or risk factor, the case-crossover design is useful. This is a relatively new study design where there is a case and a control component both of which come from the same individual. Each case is self-matched by serving as its own control. Determining the period of the control and case components is a critical and difficult aspect of a case-crossover study.

3.2.3 Cohort Studies

Cohort studies initially classify patients into two groups based on their exposure status. Cohorts are followed over time to see who develops the disease in the exposed and non-exposed groups. Cohort studies can be retrospective or prospective. Incidence can be directly calculated from a cohort study as you begin with exposed and unexposed patients, unlike a case-control study where you start with diseased and non-diseased patients. Relative risk is the measure of effect for a cohort study. Cohort studies are subject to very low recall bias, and multiple outcomes can be studied simultaneously. One of the disadvantages of cohort studies is that they are more prone to selection bias. Studying rare diseases and outcomes that have long follow-up periods can be very expensive and time-consuming using cohort studies.

3.3 Experimental Studies

3.3.1 Randomised Clinical Trials

Randomised clinical trials or randomised control trials (RCTs) are considered the gold standard of study design. In an RCT the researcher randomly assigns the subjects to a control group and an experimental group. Randomisation in RCT avoids confounding and minimizes selection bias. This enables the researcher to have similar experimental and control groups thereby enabling them to isolate the effect of an intervention. The experimental group gets the exposure/treatment which can be an agent involved in causation, prevention or treatment of a disease. The control group receives no treatment, a placebo treatment or another standard of care treatment depending on the objective of the study. The groups are then followed prospectively to see who develops the outcome of interest. RCT’s are expensive, and researchers using this study design often face issues with the integrity of randomisation due to refusals, drops outs, crossovers, and non-compliance.

3.4 Function of Epidemiology Study Design

the key function of an epidemiology study design is to enable the researcher to address the research question with minimal ambiguity logically.

3.4.1 Issues of Concern in Epidemiology Study Design

study design should be well thought of before initiating a research investigation. choosing an inappropriate study design may undermine overall study validity. critical thinking about the possible study design issues beforehand will ensure that the research question is adequately addressed.

3.4.2 clinical significance epidemiology study design

study design plays a major role in determining the scientific value of a research study. understanding the basic study design concepts will aid the clinicians in practicing evidence-based medicine.

3.4.3 other issues in epidemiology study design

Errors in study design are extremely difficult to correct after study completion. thorough planning is required to avoid weak conclusions or unconvincing results.

SELF ASSESSED EXERCISE

i. What is Hypothesis formulation in a study design?
ii. Describe the difference between cross-sectional and cross-over study design.

4.0 CONCLUSION

This unit described detail and valid epidemiological design, it also discussed the formulation of hypothesis and description of other method of epidemiological design. You are exposed different analytical and experimental method of study design. You learnt the function and clinical significance of epidemiological design.

5.0 SUMMARY

The study designs can be broadly classified as experimental or observational based on the approach used to assess whether exposure and an outcome are associated. In an experimental study design, researchers assign subjects to intervention and control/comparison groups in an attempt to isolate the effects of the intervention. Being able to control various aspects of the experimental study design enables the researchers to identify causal links between interventions sand outcomes of interest. In several instances, an experimental study design may not be feasible or suitable; in such situations, observational studies are conducted

6.0 TUTOR-MARKED ASSIGNMENT (TMA)

1. Give detail account of experimental design.
2. What are the differences between analytical and observational design?

7.0 REFERENCES/FURTHER READING

Chatburn, R.L. (2017). Basics of Study Design: Practical Considerations (From the "Biostatistics and Epid. Lecture Series, Part 1"). *Cleve Clin J Med*, 84(9 Suppl 2),10-19.

DiPietro, N.A. (2010). Methods in epidemiology: observational study designs. Pharmacotherapy. 30(10), 973-84.

Hiebert, R. & Nordin, M. (2016). Methodological Aspects of Outcomes Research. *Eur Spine J*. 15 (l1), 4-16.

Noordzij, M., Dekker, F.W., Zoccali, C. & Jager, K.J. (2019). Study designs in clinical research. *Nephron Clin Pract*. 113(3), 218-21.

Röhrig, B., du Prel, J.B., Wachtlin, D. & Blettner, M. (2019). Types of Study In Medical Research: Part 3 of A Series On Evaluation of Scientific Publications. *Dtsch Arztebl Int.* (15), 262-8.

Yang, W., Zilov, A., Soewondo, P., Bech, O.M., Sekkal, F. & Home, P.D. (2010). Observational Studies: Going Beyond the Boundaries of Randomised Controlled Trials. Diabetes Res. Clin. Pract. 88(1), 3-9.

UNIT 2 STRENGTH AND LIMITATION OF EPIDEMIOLOGICAL STUDIES

1.0 INTRODUCTION

All types of epidemiology study design have their own strengths and limitations. It is important for a Researcher to determine at the planning stage of a research the type of study design and the applicability of the validity and efficient in epidemiological study design selected.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- describe rationale for selecting a study design
- analyse the strength and limitation of descriptive studies

- list the strength and limitation of analytical studies
- explain the Strength and Limitation of Experimental Studies.

3.0 MAIN CONTENT

3.1 Overview of Strength and Limitation of Epidemiological Studies

All research designs can be discussed in terms of their relative strengths and limitations. The merits of a particular design are inherently related to the rationale for selecting it as the most appropriate plan for addressing the research problem. One strength of an experimental design, for example, is the predictive nature of the research findings. Because of the tightly controlled conditions, random sampling, and use of statistical probabilities, it is theoretically possible to predict behaviour in similar settings without actually observing that behaviour. Likewise, if a researcher needs information about the characteristics of a given population or area of interest, a descriptive study is in order. Results, however, would be limited to describing the phenomenon rather than predicting future behaviour.

3.2 Strength and Limitation Descriptive / Observational Studies

Descriptive studies are frequently the first step into a new line of enquiry, and as such have an important role in medical research, where their findings can prompt further study. Their function is to describe the "who, what, why, when, where" without regard to hypothesis, highlighting patterns of disease and associated factors. Descriptive studies that examine individuals can take the form of case reports (a report of a single case of an unusual disease or association), case series (a description of several similar cases) and cross-sectional studies (see "*Cross-sectional, analytical and intervention studies*").

3.2.1 Strengths

i. Study participants are questioned or observed in a natural setting (e.g., their homes, child care or educational settings).
ii. Study data can be used to identify the prevalence of particular problems and the need for new or additional services to address these problems.
iii. Descriptive research may identify areas in need of additional research and relationships between variables that require future study. Descriptive research is often referred to as "hypothesis generating research."

iv. Depending on the data collection method used, descriptive studies can generate rich datasets on large and diverse samples.

Descriptive ***(including ecological***) studies are generally relatively quick, easy and cheap to conduct. Particular strengths of ecological studies include:

i. Exposure data often only available at area level.
ii. Differences in exposure between areas may be bigger than at the individual level, and so are more easily examined.
iii. Utilisation of geographical information systems to examine spatial framework of disease and exposure

3.2.2 Limitations

i. Descriptive studies cannot be used to establish cause and effect relationships.
ii. Respondents may not be truthful when answering survey questions or may give socially desirable responses.
iii. The choice and wording of questions on a questionnaire may influence the descriptive findings.
iv. Depending on the type and size of sample, the findings may not be generalisable or produce an accurate description of the population of interest.
v. Measures of exposure are only a proxy based on the average in the population. Caution is needed when applying grouped results to the individual level **(ecological fallacy**).
vi. Potential for systematic differences between areas in recording disease frequency. For example, there may be differences in disease coding and classification, diagnosis and completeness of reporting between different countries.
vii. Potential for systematic differences between areas in the measurement of exposures.
viii. Lack of available data on confounding factors.

3.2.2.1Weakness of Ecological Studies Include

Weaknesses of case reports and case series are that they have no comparison (control) group, they cannot be tested for statistical associations, and they are especially prone to publication bias (especially where case reports/series describe the effectiveness of an intervention).

3.2.3 Advantages of Descriptive Research

1. It is effective in analysing non-quantified topics and issues.
2. The possibility to observe the phenomenon in a completely natural and unchanged natural environment.
3. The opportunity to integrate the qualitative and quantitative methods of data collection.
4. Less time-consuming than quantitative experiments.

3.2.4 Disadvantages of Descriptive Research

1. Descriptive studies cannot test or verify the research problem statistically.
2. Research results may reflect certain level of bias due to the absence of statistical tests.
3. The majority of descriptive studies are not 'repeatable' due to their observational nature.
4. Descriptive studies are not helpful in identifying cause behind described phenomenon.

3.2.5 Application

All forms of descriptive study can be used to generate hypotheses of possible causes or determinants of disease. These hypotheses can then be tested using further observational or interventional studies. Case reports can identify novel associations, such as the development of a rare benign liver cancer in a woman taking oral contraceptives. Case series are useful in identifying epidemics. For example, the presence of AIDS in North America was identified by the report of a cluster of homosexual men in Los Angeles with a similar clinical syndrome.

3.3 The Strength and Limitation of Analytic Study Design

Analytic study can be informed of case control or cohort study, each of which as its characteristic merit and demerit including limitations. In the previous module above, analytic epidemiology – measure effect prospective cohort studies, cross-sectional studies, retrospective case-control studies, ecologic studies and randomised controlled trials.

3.3.1 Brief overview of the case- control and cohort studies strength and Limitations

Generally, Analytic study is quick and cheap (relatively) and so ideal for outbreaks. Best for study rare diseases (or new) and can evaluate multiple exposures (fishing trips). However, some of the demerits are such that it

cannot estimate disease rates, there is Worry about representativeness of controls, the inefficient if exposures are rare. It is prone to Bias such as: Selection bias, Confounding bias and Measurement (especially recall bias).

1. Case - Control Studies

i. Characteristics: two source populations; assumption that non-cases are representative of the source population of cases.
ii. Merits: least expensive; least time-consuming; suitable for study of rare diseases (especially NCDs).
iii. Limitations: not suitable for rare exposures; liable to selection bias and recall bias; not suitable for calculation of frequency measures.
iv. Effect measure: Odds Ratio.

2. Cohort studies

i. Characteristics: follow-up period (prospective; retrospective).
ii. Merits: no temporal ambiguity; several outcomes could be studied at the same time; suitable for incidence estimation.
iii. Limitations (of prospective type): expensive; time-consuming; inefficient for rare diseases; may not be feasible.
iv. Effect measure: Risk Ratio (Relative Risk).

3.3.2 Advantages and Disadvantages of Prospective Studies (Cohort Studies)

Table 42: Advantages and Disadvantages of Prospective Studies

S/N	Advantages	Disadvantages
1.	Provides good assessment of temporal sequence	Selection bias
2.	Evaluate before onset of disease and watch for disease	Loss to follow-up
3.		Expensive
4.	Can establish population-based incidence	Lengthy and expensive
5.	Accurate relative risk (risk ratio) estimation	May require very large samples
6.	Can examine rare exposures (asbestos > lung cancer)	Not suitable for rare diseases
7.	Temporal relationship can be inferred (prospective design)	Not suitable for diseases with long-latency

8.	Time-to-event analysis is possible	Unexpected environmental changes may influence the association
9.	Can be used where randomisation is not possible	Nonresponse, migration and loss-to-follow-up biases
10.	Magnitude of a risk factor's effect can be quantified	Sampling, ascertainment and observer biases are still possible
11.	Selection and information biases are decreased	Lengthy and expensive
12.	Multiple outcomes can be studied (smoking > lung cancer, COPD, larynx cancer)	

3.3.3 Advantages and Disadvantages of Retrospective Studies

Table 43: Advantages and Disadvantages of Retrospective Studies

S/N	Advantages	Disadvantages
1	Less expensive than cohort (retrospective) Studies	Selective Survival
2	Quicker than cohort	Selective recall
3	Can identify more than one exposure	Temporal sequence not as clear
4	Good for rare diseases	Not suited for rare exposures
5	Well design leads to good etiologic investigation	Gives an indirect measure of risk
6		More susceptible to bias
7		Limited to single outcome
8	Cheap, easy and quick studies	Case and control selection troublesome
9	Multiple exposures can be examined	Subject to bias (selection, recall, misclassification)
10	Rare diseases and diseases with long latency can be studied	Direct incidence estimation is not possible
11	Suitable when randomisation is unethical (alcohol and pregnancy outcome)	Temporal relationship is not clear.
		-Multiple outcomes cannot be studied
		-If the incidence of exposure is high, it is difficult to show the

	difference between cases and controls -Not easy to estimate attributable fraction -Reverse causation is a problem in interpretation - especially in molecular epidemiology studies

3.4 Strength and Limitation Experimental Epidemiologic Studies

In an Experimental Studies it uses an intervention in which the investigator manipulates a factor and measures the outcome. Elements of a complete experiment, manipulation of data, use of a control group and ability to randomised subjects to treatment groups.

3.4.1 Strength and Limitation with Advantages and Disadvantages of Experimental Studies

Strength /Advantages	**Limitation/ Disadvantages**
Prospective direction	Contrive situation
Ability to randomise subjects	Impossible to control human behaviour
Temporal sequence of cause and effect	Ethical Constraints
Can control extraneous variables	External validity uncertain
Best evidence of causality	Expensive Disadvantages

3.4.2 General Overviews of the Strength and Weakness of Experimental Studies Design

How do you make sure that a new product, theory, or idea has validity? There are multiple ways to test them, with one of the most common being the use of experimental research. When there is complete control over one variable, the other variables can be manipulated to determine the value or validity that has been proposed.

Then, through a process of monitoring and administration, the true effects of what is being studied can be determined. This creates an accurate outcome so conclusions about the final value potential. It is an efficient process, but one that can also be easily manipulated to meet specific metrics if oversight is not properly performed.

Here are the advantages and disadvantages of experimental research to consider.

3.4.3 What Are the Advantages of Experimental Research?

1. **It provides researchers with a high level of control.** By being able to isolate specific variables, it becomes possible to determine if a potential outcome is viable. Each variable can be controlled on its own or in different combinations to study what possible outcomes are available for a product, theory, or idea as well. This provides a tremendous advantage in an ability to find accurate results.
2. **There is no limit to the subject matter or industry involved.** Experimental research is not limited to a specific industry or type of idea. It can be used in a wide variety of situations. Teachers might use experimental research to determine if a new method of teaching or a new curriculum is better than an older system. Pharmaceutical companies use experimental research to determine the viability of a new product.
3. **Experimental research provides conclusions that are specific.** Because experimental research provides such a high level of control, it can produce results that are specific and relevant with consistency. It is possible to determine success or failure, making it possible to understand the validity of a product, theory, or idea in a much shorter amount of time compared to other verification methods. You know the outcome of the research because you bring the variable to its conclusion.
4. **The results of experimental research can be duplicated.** Experimental research is straightforward, basic form of research that allows for its duplication when the same variables are controlled by others. This helps to promote the validity of a concept for products, ideas, and theories. This allows anyone to be able to check and verify published results, which often allows for better results to be achieved, because the exact steps can produce the exact results.
5. **Natural settings can be replicated with faster speeds.** When conducting research within a laboratory environment, it becomes possible to replicate conditions that could take a long time so that the variables can be tested appropriately. This allows researchers to have a greater control of the extraneous variables which may exist as well, limiting the unpredictability of nature as each variable is being carefully studied.
6. **Experimental research allows cause and effect to be determined.**
   The manipulation of variables allows for researchers to be able to

look at various cause-and-effect relationships that a product, theory, or idea can produce. It is a process which allows researchers to dig deeper into what is possible, showing how the various variable relationships can provide specific benefits. In return, a greater understanding of the specifics within the research can be understood, even if an understanding of why that relationship is present isn't presented to the researcher.

7. **It can be combined with other research methods.**
   This allows experimental research to be able to provide the scientific rigor that may be needed for the results to stand on their own. It provides the possibility of determining what may be best for a specific demographic or population while also offering a better transference than anecdotal research can typically provide.

3.4.4 What Are the Disadvantages of Experimental Research?

1. **Results are highly subjective due to the possibility of human error.**

Because experimental research requires specific levels of variable control, it is at a high risk of experiencing human error at some point during the research. Any error, whether it is systemic or random, can reveal information about the other variables and that would eliminate the validity of the experiment and research being conducted.

2. **Experimental research can create situations that are not realistic.**

The variables of a product, theory, or idea are under such tight controls that the data being produced can be corrupted or inaccurate, but still seem like it is authentic. This can work in two negative ways for the researcher. First, the variables can be controlled in such a way that it skews the data toward a favorable or desired result. Secondly, the data can be corrupted to seem like it is positive, but because the real-life environment is so different from the controlled environment, the positive results could never be achieved outside of the experimental research.

3. **It is a time-consuming process.**

For it to be done properly, experimental research must isolate each variable and conduct testing on it. Then combinations of variables must also be considered. This process can be lengthy and require a large amount of financial and personnel resources. Those costs may never be offset by consumer sales if the product or idea never makes it to market.

If what is being tested is a theory, it can lead to a false sense of validity that may change how others approach their own research.

4. There may be ethical or practical problems with variable control.

It might seem like a good idea to test new pharmaceuticals on animals before humans to see if they will work, but what happens if the animal dies because of the experimental research? Or what about human trials that fail and cause injury or death? Experimental research might be effective, but sometimes the approach has ethical or practical complications that cannot be ignored. Sometimes there are variables that cannot be manipulated as it should be so that results can be obtained.

5. Experimental research does not provide an actual explanation.

Experimental research is an opportunity to answer a Yes or No question. It will either show you that it will work or it will not work as intended. One could argue that partial results could be achieved, but that would still fit into the "No" category because the desired results were not fully achieved. The answer is nice to have, but there is no explanation as to how you got to that answer. Experimental research is unable to answer the question of "Why" when looking at outcomes.

6. Extraneous variables cannot always be controlled.

Although laboratory settings can control extraneous variables, natural environments provide certain challenges. Some studies need to be completed in a natural setting to be accurate. It may not always be possible to control the extraneous variables because of the unpredictability of Mother Nature. Even if the variables are controlled, the outcome may ensure internal validity, but do so at the expense of external validity. Either way, applying the results to the general population can be quite challenging in either scenario.

7. Participants can be influenced by their current situation.

Human error isn't just confined to the researchers. Participants in an experimental research study can also be influenced by extraneous variables. There could be something in the environment, such an allergy that creates a distraction. In a conversation with a researcher, there may be a physical attraction that changes the responses of the participant. Even internal triggers, such as a fear of enclosed spaces, could influence the results that are obtained. It is also very common for participants to "go

along" with what they think a researcher wants to see instead of providing an honest response.

8. Manipulating variables isn't necessarily an objective standpoint.

For research to be effective, it must be objective. Being able to manipulate variables reduces that objectivity. Although there are benefits to observing the consequences of such manipulation, those benefits may not provide realistic results that can be used in the future. Taking a sample is reflective of that sample and the results may not translate over to the general population.

9. Human responses in experimental research can be difficult to measure.

There are many pressures that can be placed on people, from political to personal, and everything in-between. Different life experiences can cause people to react to the same situation in different ways. Not only does this mean that groups may not be comparable in experimental research, but it also makes it difficult to measure the human responses that are obtained or observed.

The advantages and disadvantages of experimental research show that it is a useful system to use, but it must be tightly controlled in order to be beneficial. It produces results that can be replicated, but it can also be easily influenced by internal or external influences that may alter the outcomes being achieved. By taking these key points into account, it will become possible to see if this research process is appropriate for your next product, theory, or idea.

SELF ASSESSED EXERCISE

i. What do you understand by the term descriptive epidemiology study strenght
ii. How will you classify ecology studies?what are the limitation of the study
iii. Give a brief account of experimental study, strength and Limtiation.

4.0 CONCLUSION

In this unit you are exposed to the detail analysis of the strength and limitation of different epidemiology studies. Particularly, you learnt the most important advantages and disadvantages of the descriptive studies. The relationship between analytic case control and cohort studies in term

of advantages/ disadvantages was also explained. Finally, you learn the strength and limitation of experimental studies design

5.0 SUMMARY

Each of the methods used to collect descriptive data have their own strengths and limitations. The following are some of the strengths and limitations of epidemiology study design. Descriptive research studies in general cheap and less time consuming, though it cannot test or verify the research problem statistically. Research results may reflect certain level of bias due to the absence of statistical tests and easy but analytic study is quick and cheap (relatively) and so ideal for outbreaks, best to study rare diseases (or new) and can evaluate multiple exposures (fishing trips). However, some of the demerits are such that it cannot estimate disease rates, there is worry about representativeness of controls, the inefficiency if exposures are rare. It is prone to bias such as: selection bias, confounding bias and measurement (especially recall) bias. While in experimental study uses an intervention in which the investigator manipulates a factor and measures the outcome. There are elements of a complete experiment, manipulation of data, use of a control group and ability to randomise subjects to treatment groups.

6.0 TUTOR-MARKED ASSIGNMENTS (TMAs)

1. In a tabular form, state the advantages and dis-advantages of descriptive epidemiology study.
2. In a tabular form, list the advantages and dis-advantages of analytical epidemiology study.
3. In a tabular form, enumerate the advantages and dis-advantages of experimental epidemiology study.

7.0 REFERENCES/FURTHER READING

Anon. Pneumocystis pneumonia: Los Angeles. MMWR Morb Mortal Wkly Rep 1981; 30: 250–52.

Center for diseases control http://www.cdc.gov/eis/casestudies/casestudies.htm

Grimes, D.A. & Schulz, K.F. (2012). Descriptive Studies: what they can and cannot do. *Lancet*, 359, 145-9.

Hiebert, R. & Nordin, M. (2016). Methodological aspects of outcomes research. *Eur Spine J*. 15(1), 4-16.

Noordzij, M., Dekker, F.W., Zoccali, C., & Jager, K.J. (2019). Study Designs In Clinical Research. *Nephron Clin Pract*. 113(3), 218

21.

Schenken, J.R. (2016). Hepatocellular adenoma: relationship to oral contraceptives? *JAMA*, 236: 559.

UNIT 3 INTERPRET EPIDEMIOLOGIC RESULTS IN ORDER TO DEVELOP HYPOTHESES FOR POSSIBLE RISK FACTORS

CONTENTS

1.0 INTRODUCTION

Disease surveillance systems and health data sources provide the raw information necessary to monitor trends in health and disease. Descriptive epidemiology provides a way of organising and analysing these data in order to understand variations in disease frequency geographically and over time, and how disease (or health) varies among people based on a host of personal characteristics (person, place, and time). This makes it possible to identify trends in health and disease and also provides a means of planning resources for populations. In addition, descriptive epidemiology is important for ***generating hypotheses (possible explanations)*** about the determinants of health and disease. By

generating hypotheses, descriptive epidemiology also provides the starting point for analytic epidemiology, which formally tests associations between potential determinants and health or disease outcomes.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- identify the different classes of variables (discrete (dichotomous, categorical, ordinal), continuous, time to event)
- distinguish when to use mean and standard deviation versus median and interquartile range (IQR) to characterise the centre and variability for continuous variables data
- use R to compute mean, variance, standard deviation, median, and interquartile range (IQR).
- use R to compute the correlation coefficient for an ecological study
- conduct a narrative case series, present in an abstract format and put in an appropriate table for interpretation
- analysing a cross sectional survey
- computing the Correlation Coefficient
- carry out description and analysis of ecological studies
- conduct calculation of correlation and linear regression using appropriate formula.

3.0 MAIN CONTENT

3.1 Overview of Basic Concepts for Analysing and Presenting Data

This unit in this module will introduce basic concepts for analysing and presenting data from exploratory (descriptive) studies that are essential for disease surveillance, for assessing the health and health-related behaviors in a population, or for generating hypotheses about the determinants of health or disease. However, students may want to refer to other learning modules that address these concepts in greater detail. Procedures to summarise data and to perform subsequent analysis differ depending on the type of data (or variables) that are available. As a result, it is important to have a clear understanding of how variables are classified.

3.2 Types of Variables

There are three general classifications of variables:

1. Discrete Variables

Variables that assume only a finite number of values, for example, race categorised as Hausa, Fulani, Ibo and Yoruba, the White or Black and other. Discrete variables focus on the **frequency of observations** and can be presented as the number, the percentage, or the proportion of observations within a given category.

Discrete variables may be further sub-divided into:

i. Dichotomous variables
ii. Categorical variables (or nominal variables)
iii. Ordinal variables

2. Continuous Variables

These are sometimes called quantitative or measurement variables; they can take on any value within a range of plausible values. For example, total serum cholesterol level, height, weight and systolic blood pressure are examples of continuous variables. Continuous variables (measurement variables) are summarised by **finding a central measure, such as a mean or a median, as appropriate, and characterising the variability of spread around the central measure**.

3. Time to Event Variables

These reflect the time to a particular event such as a heart attack, cancer remission or death. This module will focus primarily on summarising and presenting discrete variables and continuous variables; time to event variables.

3.3 Population Parameters Versus Sample Statistics

A descriptive measure for an entire population is a "parameter." There are many population parameters, for example, the population size (N) is one parameter, and the mean diastolic blood pressure or the mean body weight of a population would be other parameters that relate to continuous variables. Other population parameters focus on discrete variables, such as the percentage of current smokers in the population or the percentage of people with type 2 diabetes mellitus. Health-related behaviours can also be thought of this way, such as the percentage of the

population that gets vaccinated against the flu each year or the percentage who routinely wear a seatbelt when driving.

However, it is generally not feasible to directly measure parameters, since it requires collecting information from all members of the population. We, therefore, take samples from the population, and the descriptive measures for a sample are referred to as "sample statistics" or simply "statistics." For example, the mean diastolic blood pressure, the mean body weight, and the percentage of smokers in a sample from the population would be sample statistics. In the image below the true mean diastolic blood pressure for the population of adults in Ilorin Kwara State, Nigeria is 78 millimeters of mercury (mm Hg); this is a population parameter. The image also shows the mean diastolic blood pressure in three separate samples. These means are sample statistics which we might use in order to estimate the parameter for the entire population. However, note that the sample statistics are all a little bit different, and none of them are exactly the sample as the population parameter.

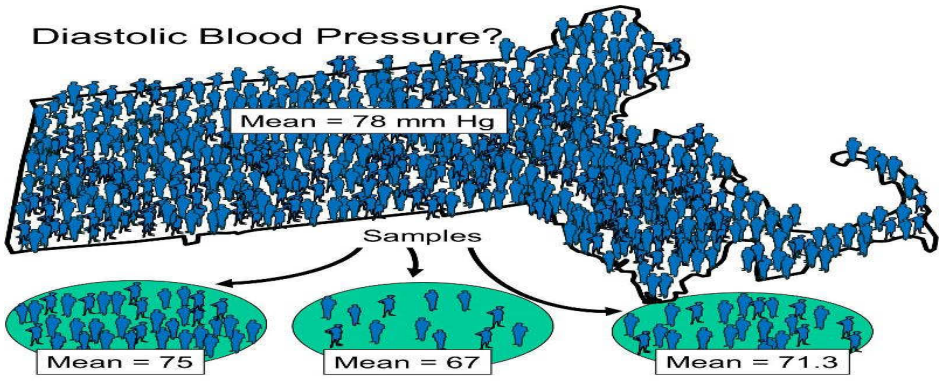


Figure 30: Mean diastolic blood pressure for the population of adults in Ilorin Kwara State

In order to illustrate some fundamentals, let's consider a very small sample with data shown in the table below.

Table 44: Data Values for a Small Sample

Subject ID	Age	Length of Stay in Hospital (days)	Current Smoker	Body Mass Index	Type 2 Diabetes
1	63	2	0	29.6	1
2	74	2	1	26.4	0
3	75	2	1	24.5	0
4	74	2	0	31.9	1

5	70	3	0	22.8	0
6	72	3	0	19.8	0
7	81	3	0	27.6	1
8	68	5	1	26.8	1
9	67	7	0	24.7	1
10	77	9	0	23.0	0

Note that: the data table has continuous variables (age, length of stay in the hospital, body mass index) and discrete variables that are dichotomous (type 2 diabetes and current smoking). Let's focus first on the continuous variables which we will summarise by computing a central measure and an indication of how much spread there is around that central estimate.

3.4 Measures of Central Tendency and Variability

There are three sample statistics that describe the center of the data for a continuous variable. There are:

i. The **Mean**: the average of all the values
ii. The **Median**: The "middle" value, such that half of the observations are below this value, and half are above.
iii. The **Mode**: The most frequently observed value.

The **mean** and the **median** will be most useful to us for analysing and presenting the results of exploratory studies.
One way to summarise age for the small data set above would be to determine the frequency of subjects by age group as show in the table below.

Table 45: Determination of frequency of subjects by age group

Age Group	**Number of Subjects**	**Relative Frequency**
60-64	1	0.1
65-69	2	0.2
70-74	4	0.4
75-79	2	0.2
89-85	1	0.1

This makes it easier to understand the age structure of the group. One could also summarise the age structure by creating a frequency histogram as shown in the Figure 31 below.

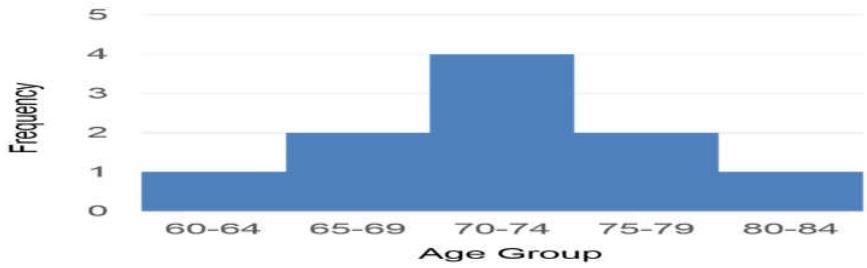


Figure 31: Frequency histogram of age range
If there are no extreme or outlying values of the variable (as in this case), the mean is the most appropriate summary of a typical value.

The Mean

The sample mean is computed by summing all of the values for a particular variable in the sample and dividing by the number of values in the sample.
So, the general formula is $\bar{X} = \frac{\Sigma X}{n}$

The X with the bar over it represents the sample mean, and it is read as "X bar". The Σ indicates summation (i.e., sum of the X's or sum of the ages in this example).
So, in the sample above means is

$$\bar{X} = \frac{63+74+75+70+72+81+68+67+77}{} = \mathbf{72.1}$$

3.5 Sample Variance and Standard Deviation

When the mean is appropriate to characterise the central values, the variability or spread of values around the mean can be characterised with the variance or the standard deviation. If all of the observed values in a sample are close to the sample mean, the standard deviation will be small (i.e., close to zero), and if the observed values vary widely around the sample mean, the standard deviation will be large. If all of the values in the sample are identical, the sample standard deviation will be zero.
To compute the sample standard deviation, we begin by computing the

variance (s^2) as follows:

$$s^2 = \frac{\Sigma(X-\bar{X})^2}{n-1}$$

The variance is essentially the mean of the squared deviations, although we divide by n-1 in order to avoid underestimating the population variance. We can compute this manually by first computing the

deviations from the mean and then squaring them and adding the squared deviations from the mean as shown in the Table 46 below.

Table 46: Computation of Variance for Age

Subject ID	Age	Deviation from The Mean $X-\bar{X}$	Squared Deviation from the Mean $(X-\bar{X})^2$
1	63	-9.1	82.81
2	74	1.9	3.61
3	75	2.9	8.41
4	74	1.9	3.61
5	70	-2.1	4.41
6	72	-0.1	0.01
7	81	8.9	79.21
8	68	-4.1	16.81
9	67	-5.1	26.01
10	77	4.9	24.01
Totals	721	0.0	248.9

Therefore, $s^2 = \frac{248.9}{9} = 27.65556$

However, the more common measure of variability in a sample is the **sample standard deviation (s)**, defined as the square root of the sample variance:

$$\text{Sample standard deviation} = s = \sqrt{s^2} = \sqrt{\frac{\Sigma(X-\bar{X})^2}{n-1}}$$

In this example the standard deviation is: $s = \sqrt{27.65556} = 5.25885$

3.6 Computing Mean, Variance, And Standard Deviation In R

These computations are easy using the R statistical package. First, I will create a data set with the ten observed ages in the example above using the concatenation function in R.

> age data <- c (63, 74, 75, 74, 70, 72, 81, 68, 67, 77)>
To calculate the mean: > mean (age data) = 72.1
To calculate the variance:
> var (age data) = 27.65556
To calculate the standard deviation for age:
> sd (age data) = 5.258855

Next, we will examine length of stay in the hospital (days) which are also a continuous variable. As we did with age, we could summarise hospital length of stay by looking at the frequency, e.g., how many patients stayed 1, 2, 3, 4, etc. days.

Table 47: Frequency of hospital admission (Length of stay)

Days in Hospital	Number of Subjects	Relative Frequency
1	0	0
2	4	0.4
3	3	0.3
4	0	0
5	1	0.1
6	0	0
7	1	0.1
8	0	0
9	1	0.1

And one again, we could also present the same information with a frequency histogram as shown below.

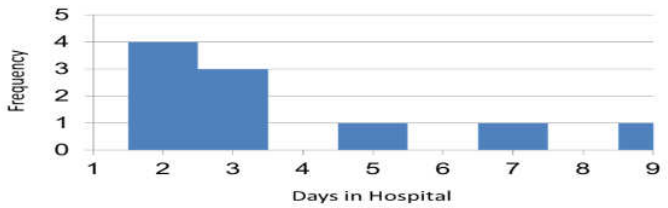


Figure 32: Frequency histogram of hospital admission (Length of stay)

Here, most patients stayed in the hospital for only 2 or 3 days, but there were outliers who stayed 5, 7, and 9 days. This is a skewed distribution, and in this case the mean would be a misleading characterisation of the central value. Rather than compute a mean, it would be more informative to compute the median value, i.e., the "middle" value, such that half of the observations are below this value, and half are above.
To compute the median one would first order the data.

i. If the sample size is an odd number, the mean is the middle value.
ii. If the sample size is an even number, the median is the mean of the two middle values.

However, R is a more convenient way to do this, because it will also enable you to see the **interquartile range (IQR)** which is a useful way of characterising the variability or spread of the data.

3.7 Computing Median and Interquartile Range with R

We can again create a small data set for hospital length of stay using the concatenation function in R:

i. **hospLOS** <- c (2,2,2,2,3,3,3,5,7,9) and we can then compute the median.
ii. **median(hospLOS)**=] 3

However, it is more useful to use the "summary ()" command. > summary (hospLOS)
Min. 1st Qu. Median Mean 3rd Qu. Max.
2.0 2.0 3.0 3.8 4.5 9.0 >

The quartiles divide the data into 4 roughly equal groups as illustrated below.

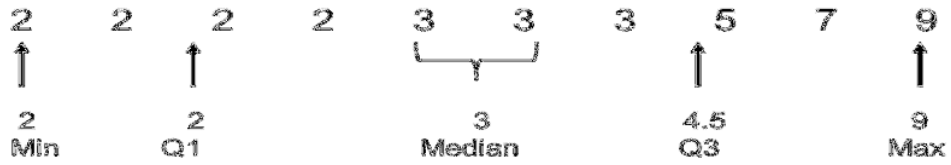


Figure 33: Quartiles division

When a data set has outliers or extreme values, we summarise a typical value using the ***median*** as opposed to the mean. When a data set has outliers, variability is often summarised by a statistic called the ***interquartile range,*** which is the difference between the first and third quartiles. The first quartile, denoted Q_1, is the value in the data set that holds 25% of the values below it. The third quartile, denoted Q_3, is the value in the data set that holds 25% of the values above it.

To summarise:

i. No outliers: sample mean and standard deviation summarise location and variability.
ii. When there are outliers or skewed data, median and interquartile range (IQR) best summarise location and variability, where IQR = Q3-Q1

Box-Whisker Plots
Box-whisker plots are very useful for comparing distributions. A box-whisker plot divides the observations into 4 roughly equal quartiles. The

whiskers represent the minimum and maximum observed values. The right side of the box indicates Q1, below which are the lowest 25% of observations, and the left side of the box is Q3, above which are the highest 25% of observations. The lowest 25% of observations are below Q1 and the highest 25% are above Q3. The median value is shown within the box.

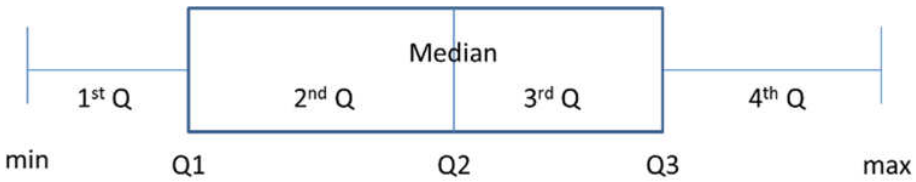


Figure 34: The median value in the box-whisker plots

3.8 Case Series – Analysis of Findings and Presentation of Case Series, Correlation Coefficient and Linear Regression

In order to be useful, the data must be organised and analysed in a thoughtful, structured way, and the results must be communicated in a clear, effective way to both the public health workforce and the community at large. Some simple standards are useful to promote clear presentation. Compiled data are commonly summarised in tables, graphs, or some combination.

3.8.1 Case Series: Presentation and Analysis with Interpretation of Case Series Findings

This is a small, but important case series reported in 2009. Shown below is the Summary and slightly modified versions of the two tables presented in the report.

Summary of the Case Series

"We performed a retrospective case-series study of patients with influenza A (H5N1) admitted to the National Institute of Infectious and Tropical Diseases in Hanoi, Vietnam, from January 2004 through July 2005 with symptoms of acute respiratory tract infection, a history of high-risk exposure or chest radiographic findings such as pneumonia, and positive findings for A/H5 viral RNA by reverse transcription–PCR. We investigated data from 29 patients (mean age 35.1 years) of whom 7 (24.1%) had died. Mortality rates were 20% (5/25) and 50% (2/4) among patients treated with or without oseltamivir ($p = 0.24$), respectively, and were 33.3% (5/15) and 14.2% (2/14) among patients treated with and without methylprednisolone ($p = 0.39$), respectively.

After exact logistic regression analysis was adjusted for variation in severity, no significant effectiveness for survival was observed among patients treated with oseltamivir or methylprednisolone."

Note that both continuous and discrete variables are reported, and note that the authors used the mean and standard deviation for variables like age, but they used median and IQR for many other variables because their distributions were skewed. Note also that discrete variables and continuous variables can be presented in the same table, but it is essential to specify how each characteristic is being presented.

Table 48: Characteristics of 29 patients infected with highly pathogenic avian influenza virus (H5N1), northern Vietnam, 2004–2005*

Characteristic	Value
Age, y, mean ± SD	35.1 ± 14.4
M:F sex (%)	15:14 (52:48)
High-risk exposure, no. (%)†	
Poultry	19 (65.5)
Sick poultry	12 (41.4)
Family infected with H5N1 virus subtype	6 (20.7)
Sick poultry or person	15 (51.7)
Hospitalisation after disease onset, median, d (IQR)	6 (4–8)
Hospital stay, median, d (IQR)	14 (9–17)
Treated with oseltamivir, no. (%)	25 (86.2)
Began treatment with oseltamivir after disease onset, median, d (IQR)	7 (5–10)
Treated with methylprednisolone, no. (%)	15 (51.7)
Died, no. (%)	7(24.1)

Table 49 below shows selected laboratory findings among survivors versus patients who died. Leukocytes are white blood cells, and neutrophils are a specific type of white blood cell; the lower numbers of these two counts in those who died suggests that the immune system was overwhelmed. Hemoglobin is a measure of red blood cells and oxygen carrying capacity. Platelets are essential elements for blood clotting. Albumin is the most abundant protein in blood. AST is an abbreviation for aspartate aminotransferase, an enzyme that is abundant in the liver; high levels of AST in the blood frequently indicate liver damage. Urea nitrogen is a measure of kidney function; high levels of urea nitrogen suggest compromised kidney function but could also be indicative of dehydration.

Table 49: Initial laboratory results for 29 patients infected with highly pathogenic avian influenza virus.

Characteristic	Survived Median IQR)	Died Median (IQR)	p-value
Leukocytes, $x10^3/\mu L$	7.8 (7.1-12.0)	3.4 (1.7-5.6)†	0.0093
Neutrophils, $x10^3/\mu L$	6.8 (4.8-9.9)	2.3 (1.1-3.8)†	0.0101
Hemoglobin, grams/L	130 (107-137)	121 (103-138)	0.6102
Platelets, $x10^3/\mu L$	214 (181-284)	86 (38-139)†	0.0101
Albumin, grams/L	34.5 (31.2-35.1)	21.7 (10.4-29.4)†	0.0265
AST, U/L	45 (28-69)	327 (77-352)	0.0077
Total bilirubin, µmol/L	10.3 (7.6-16.8)	11.4 (7.0-27.1)	0.7921
Urea nitrogen, mmol/L	4.5 (3.4-5.5)	9(3.4-14.3)	0.0462

†$p<0.05$, by Wilcoxon test or Fisher exact test.

Student should find out about the p-values and statistical tests like the Wilcoxon test and the Fisher exact test.

3.8.2 Interpretation of the Case Series on 95% confidence interval

Our best estimate of the case-fatality rate from bird flu is 24%. With 95% confidence the true case-fatality rate is likely to be between 8.5% to 39.1%.

Note that this 95% confidence interval is quite broad because of the small sample size (n=29).

Answer to 95% Confidence Interval for the Case-Fatality Rate from Bird Flu

The point estimate is

$$\hat{p} = \frac{x}{n}$$

$$\hat{p} = \frac{7}{29} = 24.1\%$$

There are 7 persons who died and 22 who did not, so we can use the following formula:

$$\text{Confidence Interval} = Z \cdot SE(\hat{p}) = Z \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Substituting:

$$95\%\ \text{confidence interval} = 0.241 \pm 1.96 \cdot \sqrt{\frac{0.241(1-0.241)}{29}} = 0.241 \pm 0.150$$

So, the 95% confidence interval is 0.085, 0.391.

3.8.3 Computing the Correlation Coefficient and Linear Regression

The module/unit on Descriptive Studies showed an ecologic study correlating per capita meat consumption and incidence of colon cancer in women from 22 countries. Investigators used commerce data to compute the overall consumption of meat by various nations. They then calculated the average (per capita) meat consumption per person by dividing total national meat consumption by the number of people in a given country. There is a clear linear trend; countries with the lowest meat consumption have the lowest rates of colon cancer, and the colon cancer rate among these countries progressively increases as meat consumption increases.

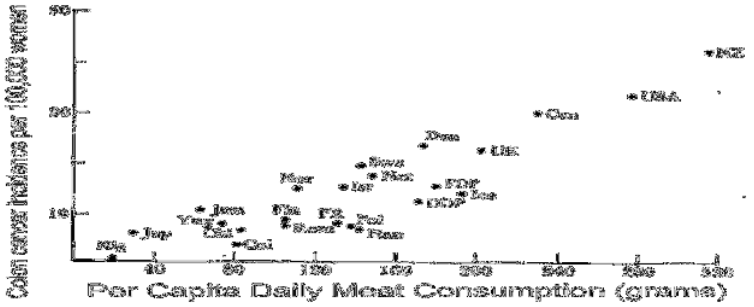


Figure 50: Correlation between colon cancer and meat consumption by Chao 2005

Note that in reality, people's meat consumption probably varied widely within nations, and the exposure that was calculated was an average that assumes that everyone ate the average amount of meat. This average exposure was then correlated with the overall disease frequency in each country. The example here suggests that the frequency of colon cancer increases as meat consumption increases.

3.8.4 Analysis of Correlation Coefficient and Linear Regression Information

As noted in the module/unit on Descriptive Studies, ecologic studies invite us to assess the association between the independent variable (in this case, per capita meat consumption) and the dependent variable (in this case, the outcome, incidence of colon cancer in women) by computing the correlation coefficient ("r"). This section will provide a

brief outline of correlation analysis and demonstrate how to use the R statistical package to compute correlation coefficients. Correlation analysis and simple linear regression are described in a later module for this course.

The most commonly used type of correlation is Pearson Correlation, named after Karl Pearson, introduced this statistic around the turn of the 20th century. Pearson's r measures the ***linear*** relationship between two variables, say X and Y. A correlation of 1 indicates the data points perfectly lie on a line for which Y increases as X increases. A value of -1 also implies the data points lie on a line; however, Y decreases as X increases. The formula for r is:

$$r = \frac{Cov(x,y)}{\sqrt{s_x^2 s_y^2}}$$

where Cov (x,y) is the covariance of x and y defined as

$Cov(x,y) = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{n-1}$ and s_x^2 and s_y^2 are the sample variances of x and y, defined as follows:

$s_x^2 = \frac{\Sigma(X-\bar{X})^2}{n-1}$ and $s_y^2 = \frac{\Sigma(Y-\bar{Y})^2}{n-1}$

The variances of x and y measure the variability of the x scores and y scores around their respective sample means of X and Y considered separately. The covariance measures the variability of the (x, y) pairs around the mean of x and mean of y, considered simultaneously.
We can combine all of this into the following equation:

$$r = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\sqrt{\Sigma(X-\bar{X})^2 \Sigma(Y-\bar{Y})^2}}$$

While this looks quite tedious, one can use R Studio to compute the correlation coefficient quite easily.

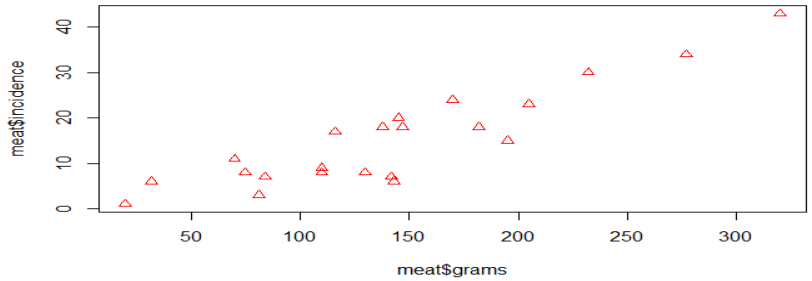


Figure 51: Visual inspection of the plot suggests a linear relationship with a strong positive correlation, and the correlation coefficient r=0.90 confirms this.

3.8.5 Simple Guidelines on Data Presentation

There are two fundamental methods for presenting summary information: tables and graphs.

i. Tables are generally best if you want to be able to look up specific information or if the values must be reported precisely.
ii. Graphics are best for illustrating trends and making comparisons

For examples of how to create effective tables and graphs and how to avoid pitfalls in data presentation:

In order to be useful, the data must be organised and analysed in a thoughtful, structured way, and the results must be communicated in a clear, effective way to both the public health workforce and the community at large. Some simple standards are useful to promote clear presentation. Compiled data are commonly summarised in tables, graphs, or some combination.

Simple guidelines for tables

1. Provide a concise descriptive title.
2. Label the rows and columns.
3. Provide the units in the column headers.
4. Provide the column total, if appropriate.
5. If necessary, additional explanatory information may be provided in a footnoted legend immediately beneath the title.

Table 50: Treatment with Anti-hypertensive Medication in Men and Women

Sex	Number on Treatment / n	Relative Frequency, %
Male	611/1,622	37.7
Female	608/1,910	31.8
Total	1,219/3,532	34.5

Simple guidelines for figures:

1. Include a concise descriptive title.
2. Label the axes clearly showing units where appropriate.
3. Use appropriate scales for the vertical and horizontal axes that display the results without exaggerating them with ranges that are either too expansive or too restrictive.
4. For line graphs with multiple groups include a simple legend if necessary.

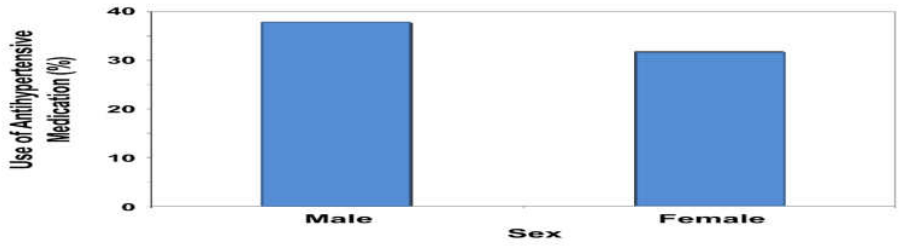


Figure 52: Relative Frequency of Anti-hypertensive Medication Use in Men and Women

Question:
The Framingham Heart Study reported that in a sample of 3,326 subjects the mean body mass index was 28.15, and the standard deviation was 5.32. What was the 95% confidence interval for the population's mean body mass index?

Answer:
We can use $X \pm z \frac{s}{\sqrt{n}}$

with z=1.96 for 95% confidence. So the 95% confidence interval is 28.15 ± 1.96. (5.32√3326) = 28.15. 1.96 (0.0922) = 28.15 ± 0.18 = (27.97.28.33)

Interpretation:
Our estimate of the mean BMI in the population is 28.15. With 95% confidence the true mean is likely to be between 27.97 and 28.33.

SELF ASSESSED EXERCEISES

i. How can you summarise data?
ii. How do you produce basic figures and tables?
iii. How can you analyse the correlation between two continuous variables?
iv. How can you apply this to the analysis and description of an ecologic study?
v. How can you use R to do descriptive analyses?

4.0 CONCLUSION

In this Unit you must have learnt how to interpret descriptive epidemiologic results in order to develop hypotheses of possible risk factors. You were exposed to case series and how to go about in the presentation, analysis and interpretation of such data. You were also exposed to scenario of an ecological case in relation to correlation coefficient and liner regression on analysis and presentation. Finally, you leant about simple guidelines on data presentation.

5.0 SUMMARY

Descriptive epidemiology has Specific tasks of on monitoring and reporting on the health status and health related behaviours in populations, identifying emerging health problems, alerting us to potential threats from bioterrorism, establishing public health priorities for a population, evaluating the effectiveness of intervention programmes and exploring potential associations between "risk factors" and health outcomes in order to generate hypotheses about the determinants of disease. In achieving all this function of descriptive epidemiology is important for ***generating hypotheses (possible explanations)*** about the determinants of health and disease. By generating hypotheses, descriptive epidemiology also provides the starting point for analytic epidemiology, which formally tests associations between potential determinants and health or disease outcomes.

6.0 TUTOR-MARKED ASSIGNMENT (TMA)

1. The Framingham Heart Study reported that in a sample of 3,326 subjects the mean body mass index was 28.15, and the standard

deviation was 5.32. What was the 95% confidence interval for the population's mean body mass index?

2. What are the simple guidelines in data presentation?
3. Describe formula for correlation coefficient and liner regression.
4. What are the differences between correlation coefficient and liner regression?

7.0 REFERENCE/FURTHER READING

Abu-Sbeih, A. & Fontaine, R.E. (2000). Secondary water-borne hepatitis E outbreak from water storage in a Jordanian town (abstract LB01). *Presented at the TEPHINET First International Conference Ottawa, Canada*, 17–21.

Biggs, H.M., Behravesh, C.B., & Bradley, K.K. (2016). Diagnosis and management of tickborne rickettsial diseases: Rocky Mountain spotted fever and other spotted fever group rickettsioses, ehrlichioses, and anaplasmosis—United States. MMWR Recomm Rep. 65(2), 1–44.

Centres for Disease Control and Prevention. (1999). National Centre for Health Statistics. National Vital Statistics System. Death rates for selected causes by 10-year age groups, race, and sex: death registration states, 19(33), 82-98. https://www.cdc.gov/nchs/nvss/mortality/hist290.htm

Centres for Disease Control and Prevention. Injuries and deaths associated with use of snowmobiles—Maine, 1991–1996. MMWR. 46, 1–4.

Centers for Disease Control and Prevention. (1999). Intussusception among recipients of rotavirus vaccine—United States. MMWR. 48, 577–81.

Dou, F., Sun, H. & Wang, Z.J. (2017). Dengue outbreak at a fishing port: Guangdong Province, China, 2007. Presented at the International Conference on Emerging Infectious Diseases. 16–19.

Ehrenberg, A.C. (2018). The problem of numeracy. Am Stat. 35, 67–71.

Fleiss, J.C. (2018). Statistical methods for rates and proportions. New York: John Wiley & Sons; 79-81.

Fontaine, R.E., Van Severin, M. & Houng, A. (2018). The stratification of malaria in El Salvador using available malaria surveillance

data (Abstract 184). Presented at the XI International Congress for Tropical Medicine and Malaria Calgary, Alberta, Canada. 16–22, 1984,

Hawley, B., Casey, M.L, Cox-Ganser, J.M., Edwards, N., Fedan, K.B. & Cummings, K.J. (2016). Notes from the field: respiratory symptoms and skin irritation among hospital workers using a new disinfection product—Pennsylvania. MMWR. 65,400–1.

Galvao, R.M, Ribeiro, C.M., Johnson, W.D. & Riley, L.W. (1999). Urban epidemic of severe leptospirosis in Brazil. Salvador Leptospirosis Study Group. Lancet. 354, 20–5.

Nguyen, D.H. & Nguyen, H.H. (2009). Human infection with highly pathogenic avian influenza virus (H5N1) in Northern Vietnam. *Emerg Infect Dis*. 15(1),19–23.

MODULE 4 OBSERVATIONAL EPIDEMIOLOGY

UNIT 1 DESCRIPTIVE EPIDEMIOLOGY

CONTENTS

1.0 INTRODUCTION TO EPIDEMIOLOGY

Epidemiology is considered the fundamental science of Public Health, and with good reason. Epidemiology is: a) a quantitative basic science built on a working knowledge of probability, statistics, and sound research methods; b) a method of causal reasoning based on developing and testing hypotheses pertaining to occurrence and prevention of morbidity and mortality; and c) a tool for Public Health action to promote and protect the public's health based on science, causal reasoning, and a dose of practical common sense. As a Public Health discipline, epidemiology is instilled with the spirit that epidemiologic information should be used to promote and protect the Public's Health.

Hence, epidemiology involves both science and art of Public Health practice. The term **applied epidemiology** is sometimes used to describe the application or practice of epidemiology to address Public Health issues.

Examples of applied epidemiology include the following:

- the monitoring of reports of communicable diseases in the community
- the study of whether a particular dietary component influences the risk of developing cancer.

- evaluation of the effectiveness and impact of a cholesterol awareness program
- analysis of historical trends and current data to project future public health resources.

The word **epidemiology** comes from the Greek words **epi**, meaning "on or upon," **demos**, meaning "people," and **logos**, meaning "the study of." Many definitions have been proposed, but the following definition captures the underlying principles and the public health spirit 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 prevention and control of health problems." This definition of epidemiology includes several terms which reflect some of the important principles of the discipline. As you study this definition, refer to the description of these terms below.

Study: Epidemiology is a scientific discipline, sometimes called "the basic science of public health." It has, at its foundation, sound methods of scientific inquiry.

Distribution: Epidemiology is concerned with the frequency and pattern of health events in a population. Frequency includes not only the number of such events in a population, but also the rate or risk of disease in the population. The rate (number of events divided by size of the population) is critical to epidemiologists because it allows valid comparisons across different populations.

- Pattern refers to the occurrence of health-related events by time, place, and personal characteristics.
- Time characteristics include annual occurrence, seasonal occurrence, and daily or even hourly occurrence during an epidemic. It also includes such temporal factors about disease occurrence like the natural history of the disease, time-incidence curve (epicurve) duration of disease, etc.
- Place characteristics include environment context, geographic variation, urban-rural differences, and location of worksites or schools.
- Personal characteristics include demographic factors such as age, race, sex, marital status, and socioeconomic status, as well as behaviours and environmental exposures.

This characterisation of the distribution of health-related states or events is one broad aspect of epidemiology called **descriptive epidemiology**. Descriptive epidemiology provides the *What*, *Who*, *When*, and *Where* of health-related events.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- define epidemiology
- describe the three main characteristics of descriptive epidemiology.

3.0 MAIN CONTENT

3.1 Descriptive Epidemiology

Determinants: Epidemiology is also used to search for causes and other factors that influence the occurrence of health-related events. **Analytic epidemiology** attempts to provide the *Why* and *How* of such events by comparing groups with different rates of disease occurrence and with differences in demographic characteristics, genetic or immunologic make-up, behaviours, environmental exposures, and other so-called potential risk factors. Under ideal circumstances, epidemiologic findings provide sufficient evidence to direct swift and effective public health control and prevention measures.

Health-related states or events: Originally, epidemiology was concerned with epidemics of communicable diseases. Then epidemiology was extended to endemic communicable diseases and non-communicable infectious diseases. More recently, epidemiologic methods have been applied to chronic diseases, injuries, birth defects, maternal-child health, occupational health, and environmental health. Now, even behaviours related to health and well-being (amount of exercise, seat-belt use, etc.) are recognised as valid subjects for applying epidemiologic methods. In these lessons, the term “disease” to refer to the range of health-related states or events.

Specified populations: Although epidemiologists and physicians in clinical practice are both concerned with disease and the control of disease, they differ greatly in how they view “the patient.” Clinicians are concerned with the health of an individual; epidemiologists are concerned with the collective health of the people in a community or other area. When faced with a patient with diarrheal disease, for example, the clinician and the epidemiologist have different responsibilities. Although both are interested in establishing the correct diagnosis, the clinician usually focuses on treating and caring for the individual. The epidemiologist focuses on the exposure (action or source that caused the illness), the number of other persons who may have been similarly exposed, the potential for further spread in the community, and interventions to prevent additional cases or recurrences.

Application: Epidemiology is more than "the study of." As a discipline within public health, epidemiology provides data for directing public health action. However, using epidemiologic data is an art as well as a science. Consider again the medical model used above: To treat a patient, a clinician must call upon experience and creativity as well as scientific knowledge. Similarly, an epidemiologist uses the scientific methods of descriptive and analytic epidemiology in "diagnosing" the health of a community, but also must call upon experience and creativity when planning how to control and prevent disease in the community.

Descriptive epidemiology is the study of the amount and distribution of disease within a population by person, place and time. In other words, descriptive studies are carried out in order to determine the frequency of a disease, the kind of people suffering from it, where it occurs and when it occurs. Descriptive studies identify non-random variations in the distribution of disease. At the end of the study, the investigator is able to generate putative hypothesis, which is a testable preposition regarding the etiology or cause of (risk factors for) the disease in question. This hypothesis can either be accepted or rejected after some further studies, using appropriate analytic epidemiological study designs.

Descriptive studies are difficult to carry out in developing countries due to lack of basic census data on the numbers, characteristics and distribution of people. A descriptive study may comprise observations made at one point in time, known as cross-sectional study. It may also be made of observations repeated in the same community over a prolonged period of time, known as longitudinal study. A group of individuals may be followed up for many years and observed for patterns of development of illness among them.

Descriptive studies yield information that may be used for planning, implementation and evaluation of health services. The target people with the prevailing health problem, their location and time of occurrence of the problem can be identified and appropriate remedies and resources channeled to them. In etiological enquires, descriptive studies can lead to a specific hypothesis relating a suspected etiological factor to a disease entity. Examples are the suspected relationship between cigarette smoking and lung cancer, and that between Burkitti's lymphoma and malaria.

Descriptive studies also identify problems that require further studies using other epidemiological methods. Descriptive epidemiological studies form an essential part of information needed for understanding the inter-relationship between the environment, the disease agent and the human host (Epidemiological Triad) in many infectious and parasitic

diseases. For example, *Vibrio cholera* is the causative agent of cholera, but outbreak of cholera epidemic is determined by various environmental factors such as water supply, methods of refuse and sewage disposal, prevalence of cholera carriers, personal hygiene, food hygiene, attitude to and utilisation of medical services.

Descriptive studies usually make use of routinely collected data (archival data) such as hospital records, birth and death registration records. However, in some cases, the data required for the purpose of describing disease distribution in a population and related variables are not readily available. In these circumstances, it becomes pertinent that special surveys should be conducted in order to provide the materials for a descriptive study. These are known as cross-sectional surveys, which involve the collection of data in a planned manner for a specific purpose.

Cross-sectional studies are also known as point prevalence surveys. In descriptive studies, some broad questions must be answered at the end of the process:

- Who are affected (i.e. the person)?
- Where do the cases occur (i.e. the place)?
- When do the cases occur (i.e. the time)?

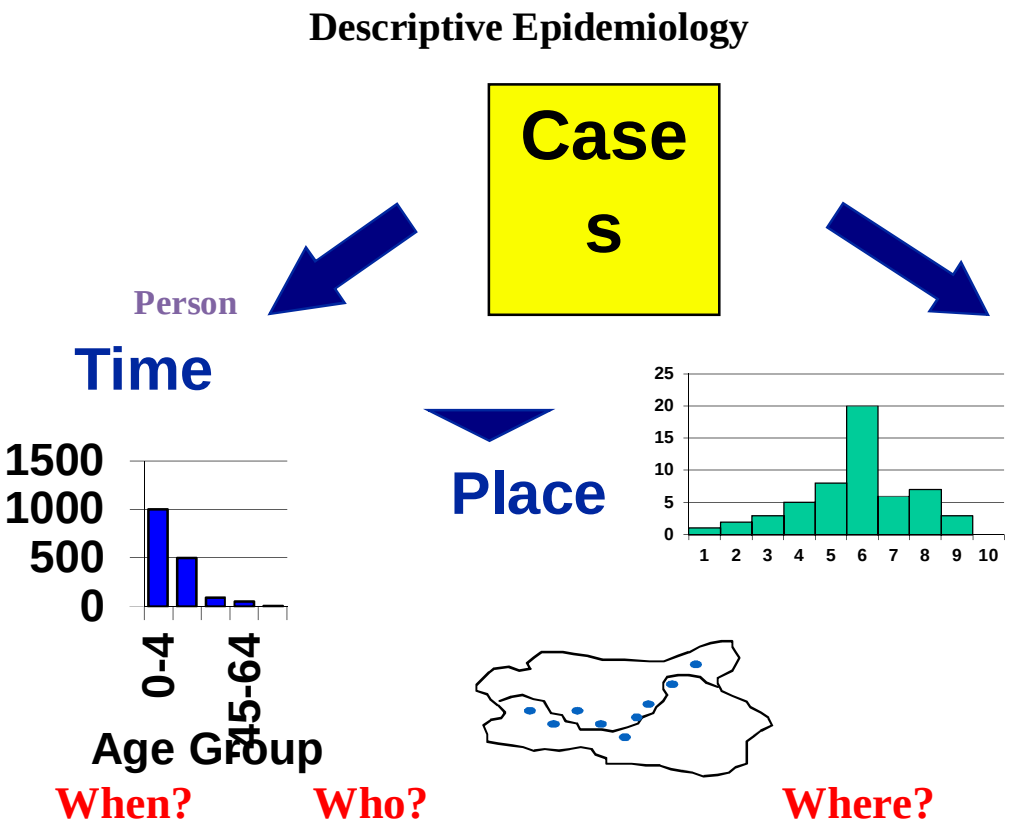


Figure 1.2

Graphical Pictorial Diagrams of Distribution of Person, Place and Time in Descriptive Epidemiology. *Ekanem (2015). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Nigeria.*

3.2 Characteristics of PERSON

Information about patients or study subjects is analysed to show the distribution attributes or variables as the case may be. An attribute is a quality or characteristic of a person e.g. sex, a variable is a quantity that may vary in value e.g. age, parity, etc. The three characteristics of a person that are almost always specified in any epidemiological study are age, sex and race or ethnic group. These are known as demographic characteristics. Other less frequently used characteristics are social class, occupation, marital status and place of residence. Descriptive studies also involve information on family and personal characteristics. The commonly used demographic characteristics are as follows:

3.2.1 Age

Mortality and all morbidity rates of almost all health conditions show some relationship to age. Mortality rates from all causes are higher at both extremes of age for both males and females of all races. Mortality rates are fairly high in infancy especially after the age of 6 months when maternal immunity decreases. After this period, the death rate decreases markedly and reaches its lowest point between the ages of 5 and 14 years. The rate then climbs gradually until the age of 40 years when it then increases exponentially, almost doubling with each decade. Age is equally related to patterns of morbidity. Morbidity rates for chronic conditions tend to increase with age e.g. arthritis, dental problems etc. For acute conditions, the relationship is less consistent. Young children readily acquire respiratory infections after the age of 6 months when passive immunity derived from mother at birth (maternal/puerperal immunity) wanes. Some occupational diseases are related to the ages of the workforce. These diseases are caused by occupational exposure or are somehow else occupationally related. The longer people are exposed to hazards of the workplace, the more likely they are to suffer from these occupational diseases. Most occupational induced diseases are therefore found in older workforce.

3.2.2 Sex

Mortality rates from all causes (with the exception of gynecological and obstetrical matters) are higher in males than in females of all ages and in all races. Fetal and neonatal death rates are also higher in males. The higher the death rates for males may be due to sex-linked inheritance or to differences in hormonal balance, environment or habit patterns.

However, the sex difference in mortality rates varies greatly for different disease entities. For homicide, the male-to-female ratio is almost four but for suicide and chronic respiratory disease, it is approximately three. There is an almost equal sex differential for ischemic heart disease, congenital anomalies and diabetes mellitus. Conversely, females usually have increased morbidity than males for a number of different conditions and in almost all age groups, especially the reproductive age group and after the age of 45 years. This is reflected in the fact that they report more for illness and attend hospitals and health institutions more. Women also seek medical help at an earlier stage of the illness. The same disease also tends to have a less lethal course in women than men. These differences are from patterns of sick role and illness behaviour, largely sociological, cultural and psychological.

3.2.3 Ethnic Group and Race

Many diseases differ markedly in frequency, severity or both in different racial groups. Statistics by race also help in identifying health problems. Some observed differences are due to differences in socio-economic status e.g. the non-white population United States of America have higher mortality rates than the whites, but in recent years the gap has continued to narrow. Blacks have higher rates of hypertensive heart disease, cardiovascular accidents, tuberculosis, syphilis, homicide and accidental deaths than the whites. Cancer of the cervix is higher in black females, while cancer of the breast is higher in white females. Sickle cell disease is virtually restricted to blacks. Alcoholism is rare in Jews because of their religion, which discourages it.

3.2.4 Marital Status

Marital status has been found to be associated with the level of mortality and morbidity of both sexes. Death rate from most specific diseases and from all causes combined, have been found to vary from lowest to highest in the following order: married, single, widowed and divorced. Better health of the married in most cases may be attributed to the psychologic and physical support provided by the spouse. Marital status in women may affect health as a result of differences in sexual exposure, pregnancy, childbearing and lactation. Cancer of the cervix is more prevalent in married women and may also be affected by early sexual exposure and multiple partners. In contrast, cancer of the breast is more common in single women, but its occurrence may also be influenced by hormonal balance. Sexual activity is important in morbidity and mortality rates because it affects the risk of pregnancy and sexually transmitted diseases. Pregnancy and childbearing entail special risks to the women as well as possible effects on subsequent development of cancer. Many problems are associated with pregnancy, delivery and puerperium. In early

pregnancy, there may be the problem of either spontaneous or induced abortion with the associated complications. At this period, there may also be the problem of a ruptured ectopic pregnancy. Towards the end of pregnancy, there may be the problem of pre-eclampsia, which may lead to death. Other demographic variables are: Social Class, Occupation and Family Variables.

3.3 Characteristics of PLACE

Diseases differ in frequency in terms of geographic location or natural barriers such as mountain ranges, rivers or deserts, or by political boundaries. These variations in frequency often provide clues to the possible causes of these diseases, and subsequently the possible ways of controlling them. Natural boundaries are more useful than political boundaries in eliciting the cause of disease. However, political boundaries provide denominators for rate of diseases from census data. They also collect information on cases (numerators), which are used for planning and provision of services. For example, a local government area health department may need to know the number of persons with newly diagnosed cases of tuberculosis residing in its area of jurisdiction, the number of persons living with acquired immune deficiency syndrome, or the number of handicapped children unable to attend school. Political boundaries are often arbitrary and may either bisect homogenous areas or join disparate ones. An example of the latter situation is a typical metropolitan city in Nigeria with variations in health and socio-economic conditions.

An area defined by natural boundaries may have particular environmental or climatic conditions such as temperatures, humidity, rainfall, altitude, mineral content of soil or water supply. The nature of the terrain also affects economic activities and patterns of transportation, including access to medical care facilities. Characteristics of physical and biological environment can cause certain diseases to be more prevalent in certain places compared with others. Diseases that occur in specific environmental condition are known as *place* diseases. Examples are parasitic and infectious diseases, which occur more frequently in tropical areas.

3.3.1 Urban-Rural Variations in Health and Disease

Globally, extensive migration has been taking place over the years from the rural to the urban areas. The causes have been the availability of jobs and other attractions in the cities, coupled with the mechanisation of farm work and consequent decrease in the number of jobs in the farms. In Nigeria, the urban dwellers constitute 35 percent of the total population, while the rural dwellers make up 65 percent. A large proportion of the

rural dwellers are disadvantaged through illiteracy, lack of job opportunities, malnutrition, disease, and a shortage of medical personnel and facilities. A good number of people living in the rural areas are farmers, and so suffer from occupational health problems of agricultural workers. These are farm accidents which result from the use of mechanised equipment without the benefit of proper training, supervision and regulation. Other problems are skin cancers, as a result of repeated exposure to pesticides and other chemicals used in farming. Agricultural workers also suffer from exposure to a variety of micro-organisms such as tetanus, anthrax among others.

Besides, city dwellers also suffer from a variety of health hazards as a result of their peculiar location. An important problem is that of air pollution, chiefly due to a concentration of industries and automobiles that emit some noxious gases into the atmosphere. Some extensive changes are necessary in order to reverse the critical environmental deterioration in the form of air pollution in the urban areas. The big cities are also home to variety of social vices such as homicide, terrorism and other acts of violence. The habits of the urban dwellers also favour the spread of sexually transmitted infections, including HIV/AIDS, as well as drug and substance abuse. In as much as these problems are not limited to urban cities, a greater proportion are present in these places.

Similar problems are springing up in the rural areas because of easier communication from improved means of transportation.

3.4 Characteristics of TIME

Disease occurrence in populations is often related to time. This is often expressed on a monthly or annual basis. When occurrence of disease is expressed every 10 years, it is known as decennial basis, e.g. 1993, 2003. In places where population counts are regular, census counts are used to calculate rates of disease occurrence rather than estimated population. If a particular disease is not commonly present in an area, several years may be combined to provide the rate occurrence. The disease in question must have stable rates. The three major types of change in disease occurrence with time are as follows.

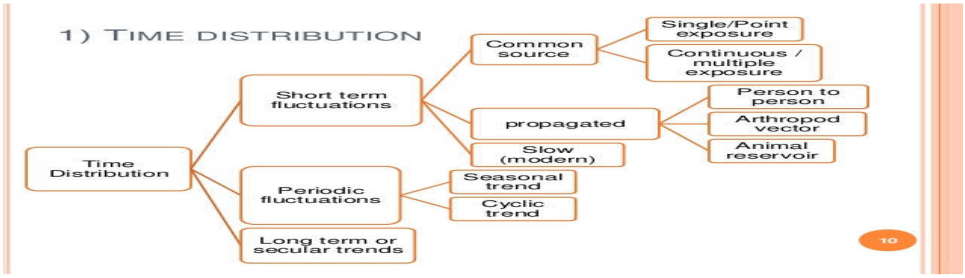


Figure 1.2

Graphical Tree Diagrams of Time Distribution in Descriptive Epidemiology. ***Ekanem EE (2015). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Nigeria.***

3.4.1 Secular Trends

This term refers to changes over a long period of time, years or decade i.e. long-term fluctuations. It occurs in both infectious and non-infectious conditions. Secular trend has been noted in some cancers e.g. mortality rates from cancers of the lung and pancreas have been noted to increase in the United States of America. This may be due to better diagnostic procedures. Secular trends may also be observed with reference to deaths from diseases. It is worthwhile to consider how they are related to changes in both incidence and survival from the disease. Death rates are similar to incidence rates if the disease is fatal and death occurs shortly after diagnosis. An example is lung cancer, with a high and early fatality rate. Apparent changes in secular trends may have been due to changes in doctors' index of suspicion, in diagnostic methods, and changes in rules for reporting and coding causes of death on death certificates of some countries.

3.4.2 Cyclic Changes

Cyclic changes are recurrent alterations in the frequency of disease or death. The cycles may be annual i.e. seasonal, or may have other periodicity. Seasonal changes in frequency of disease or death are observed in many conditions, both infectious and non-infectious. Examples are measles and influenza epidemic, which used to occur every two to three years in some areas. Leptospirosis is more prevalent during the holiday season, when a lot people indulge in swimming and fishing and therefore get exposed to Leptospira, a type of spirochete presents in water that is contaminated with the urine of infected animals. In temperate countries, overall death rates from all causes fluctuate markedly by season, becoming higher in winter than in summer. Injuries in children also show some periodicity, with increased rates during the holiday period. No seasonal pattern has been found in the onset of cancer, except in the case of malignant melanoma of the upper extremities, which has been noted to increase during summer months as a result of increased ultraviolet radiations from the sunlight. However, other factors such as age at exposure and outdoor activities, especially in the youths are likely to have some effects on the onset of melanoma. The study of seasonal changes has been used for evaluating the role of insect vectors in disease transmission, since these vectors thrive more in certain conditions of temperature and humidity.

In addition to seasonal patterns, diseases show long-term trends that are cyclic in nature; hence it is referred to as secular cyclicity in disease occurrence. Secular Cyclicity, therefore, is the reoccurring (cyclic) changes in the pattern of occurrence of disease over long periods of time, often encompassing several decades. With infectious diseases, this phenomenon is often determined by changes in the pathogenicity or the virulence of the etiologic agent of the disease as may be caused by genetic changes, as well as long term changes in the susceptibility of the host.

This phenomenon of periodicity in the occurrence of a disease in a population in which changes occur gradually over long periods of time is important consideration in disease control programming. It is a particularly important consideration in the frequency of disease encompassing several decades. Major factors influencing secular cyclicity are changes in the genetic make-up (gene pool) of the host population and in that of the agent of the disease as may result from nutrition.

3.4.3 Short-term Fluctuations

An epidemic is the best known short-term fluctuation. It is define as "The occurrence in a community or region of cases of an illness or other health related events clearly in excess of normal expectancy". The level of normal expectancy in the occurrence of disease in a population is often referred to as its **endemicity**.

There are three types, namely: Common source epidemics, Propagated epidemics and Slow (modern) epidemics. Common source epidemics can be divided into Single and Continuous exposures. In single exposure, it can occur due to an infectious agent or as a result of contamination of the environment and develops within one incubation period. This epidemic curve rises and falls rapidly, usually has one peak.

Epidemiologists call this a **holomintic curve**, characteristic of mass exposure of susceptible persons (virgin population) to an infectious agent of disease over short time interval. It tends to be explosive (i.e. clustering of cases within a short-time). In continuous exposure, this occurs when the exposure from the same source is prolonged and the epidemic continues more than one incubation period. The epidemic reaches a sharp peak, but tails off gradually over a long period of time, for example, a well of contaminated water or nationally distributed vaccine (polio vaccine) or food; water-borne cholera. Epidemiologists refer to this pattern of time-incidence curve as a **prosodemic curve**. Typically, the curve tappers as the rate of transmission to susceptible persons decline.

4.0 CONCLUSION

Descriptive epidemiology provides data regarding the magnitude of the disease load and types of disease problems. In terms of morbidity and mortality rates and ratios, it provides clue as factors that are associated with disease occurrence and help in the formulation of an etiological hypothesis. It provides background data for planning, organising and evaluating preventive service. Contribute to research by describing variations in disease occurrence by person, place and time.

5.0 SUMMARY

You should now be able to describe key aspects of the epidemiological approach to health events in terms of the event's distribution in person, place and time. You should be able to understand epidemiological models of causation of disease. You should be able to understand that results of epidemiological investigations are required to provide information about the natural history of disease and prognosis, and to help identify appropriate interventions and measures of control in public health.

6.0 TUTOR- MARKED ASSIGNMENT (TMA)

1. Define the term epidemiology
2. Discuss the characteristics of PERSON, PLACE and TIME in the concept of descriptive epidemiology.

7.0 REFERENCES/FURTHER READING

Chika Onwasigwe, (2004). Principles and Methods of Epidemiology. Institute for Development Studies, University of Nigeria, Enugu.

Ekanem EE, (2015). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Akinsola OJ, (2019). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Lucas AO & Gilles HM (2003). Short Textbook of Public Health Medicine for the Tropics, (4th ed.) .

Rothman KJ, Greenland S. & Lash T. (2011). Modern Epidemiology, (3rd ed.). An Introduction to Applied Epidemiology and Biostatistics, (1992). (2nd ed.). Centre for Disease Control.

UNIT 2 ANALYTIC EPIDEMIOLOGY I: CASE-CONTROL STUDY

CONTENTS

1.0 INTRODUCTION

The last module dealt with descriptive epidemiology, which is used to identify groups with high or low rates of a specific disease or health condition. This information is used to plan and implement health programmes that are aimed at ameliorating the problems. Analytic studies are then used to determine ***why*** the rate is high or low in a particular group. Analytic epidemiology studies are hypothesis-testing studies that are used to verify the hypotheses that were generated after the descriptive studies. The purpose of hypothesis testing is to verify the association between a suspected risk factor and the disease in the general population.

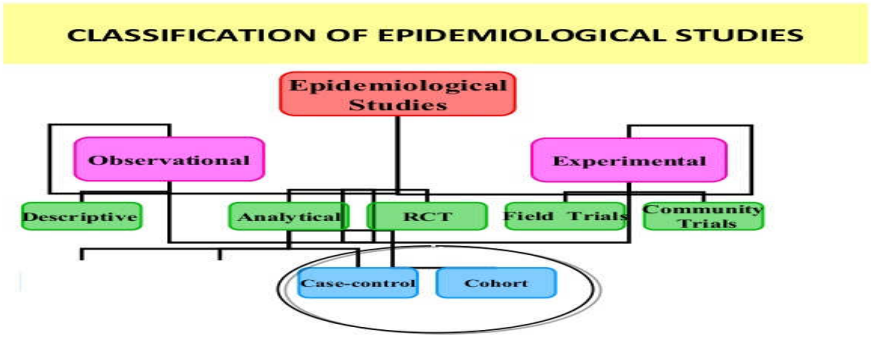


Figure 2.1: Graphical Classification of Epidemiological Studies. ***An Introduction to Applied Epidemiology and Biostatistics, Second Edition. Center for Disease Control, USA (1992).***

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- define case control study
- describe the procedure for conducting case control study

3.0 MAIN CONTENT

3.1 Analytical Studies

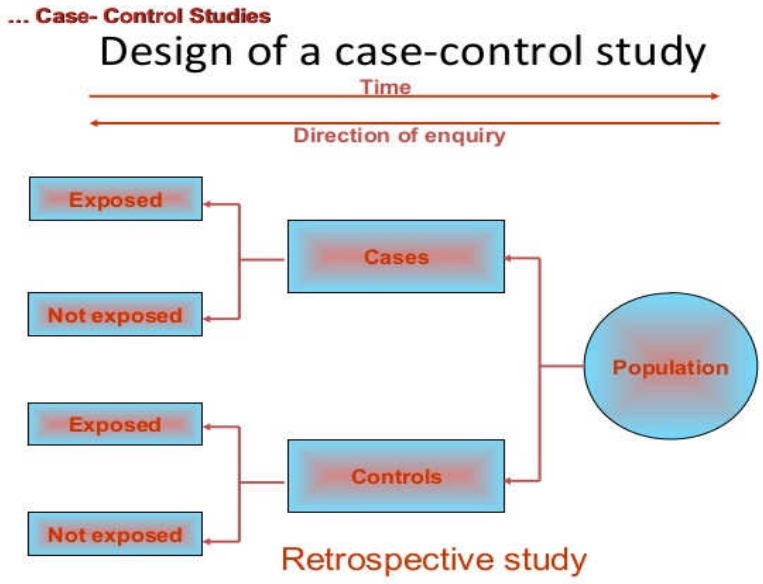
Analytic studies are then undertaken to test specific hypotheses. Samples of subjects are identified and information about exposure status and outcome is collected. The essence of an analytic study is that groups of subjects are compared in order to estimate the magnitude of association between exposures and outcomes. Analytical studies are done in order to find out if an outcome is related to exposure. They can take the form of observational (we don't influence the population), as well as interventional/experimental (we influence the population, e.g.: clinical trials; we give drugs). There are four main types of analytical studies: Ecological, Cross-sectional, Case-control and Cohort.

3.2 Case-Control Study

Case control studies are also known as retrospective studies. They are the first approach to estimate the cause-effect relationship between a suspected risk factor and a disease. The case-control study design has three distinct characteristics:

1. Exposure to risk factor and disease must have occurred before the start of the study.
2. The study goes backward in time from effect to cause
3. A control or comparison group is used to support or refute inference.

In a retrospective study, people diagnosed as having a disease (cases) are compared with persons that do not have disease (controls). The purpose is to find out if the two groups differ in the proportion of persons that have been exposed to a specific factor or factors in the past. This type of study is a backward-looking one. It is called a retrospective study because it compares the cases and controls with regards to the presence of a suspected etiological factor or factors in their past experience.



3.3 Steps in Conducting A Case-Control Study

Four steps are involved in conducting a case control study.

1. Selection of cases and control
2. Matching
3. Measurement of exposure
4. Analysis and interpretation

3.3.1 Selection of Cases and Controls

- *Selection of Cases*

Certain guidelines are followed when selecting cases. A case is first of all defined my means of diagnostic criteria and the stage of disease if any, e.g. cancer of the cervix stage I. The case must be specified before a study is started. All the cases to be studied must belong to the same clinical or histological group. Eligibility criteria must be adhered to. An example is the requirement in a particular study that only newly diagnosed cases (Incident cases) within a specified period of time are to be studied. The old cases or those in advanced stages of the disease (prevalent cases) are not included in the study.

Cases for a study should be all newly diagnosed patients (incident cases) with specified characteristics during a specified period of time in a defined population. Incident cases are preferable to prevalent cases made up of people that are currently sick (both old and new cases) as a result of the disease under study. Prevalent cases do not include patients that have had a short course of the disease, having been rapidly cured or died.

This could introduce bias into the study because different risk factors may induce or maintain a disease. The bias may be removed by including deceased cases and those that are alive into the study. Deceased cases may also be selected for a study especially if the disease that led to the death is a rare one. These cases are obtained from hospital records, death certificates, employers' records, and the general population, etc. Information about cases may be gathered by the cases themselves, their relatives and friends or from records. It is more convenient to select cases from hospital, either randomly or a series of patients that

presented at the same time (case series). In a population-based study, cases of a disease that occur within a defined geographic area are selected. This is done through a health survey, a disease registry or hospital network. The case series or a random sample that represents the cases in the community may be selected.

- *Selection of Controls*

This is more difficult than selection of cases. The controls, which must be free of the disease under study, should resemble the cases in other characteristics. The controls must not be exposed to the suspected causative factor of the disease under study. Selection of an appropriate control group is necessary in order to make good comparisons, draw inferences and make judgments about the investigation. When studying large number of cases, an investigator may use one control for each case. If the total number of cases is small (less than 50), two to four controls may be used for each case. The control should be selected from more than one source in order to prevent selection bias. However, the controls may be sourced from hospitals, relatives and neighbors of cases, or from the general population.

3.3.2 Matching

Matching is the process of selecting controls in such a way that they are similar to cases with regard to certain pertinent variables, e.g. age, sex, race, occupation, socio-economic status, educational level, which are known to influence the outcome of disease. If the controls are not adequately matched with the cases, comparability between the two groups would not be ensured and the results could be distorted or confounded.

- *Confounding Variables*

Confounding variables are factors that are known to be associated with both the exposure of interest and causally with the disease under study. In other words, confounding variables are also risk factors for the disease

under study or could modify the risk factors. It is important to remove the influence of these in order to remove bias that may arise. Examples are age, educational level and socio-economic status, which are common confounding variables that are often associated with exposure and disease under study. Another example of confounding variable is the study of the association between lung cancer and cigarette smoking, where cases of chronic pulmonary disease were used as controls. Smoking affects chronic pulmonary disease as well as lung cancer.

When designing a study, bias should be removed by matching cases and controls with respect to the confounding factor. A 25 year-old married female graduate of Ibo origin (case), may be matched with another 25±2 year-old married female graduate of Ibo origin (control). This is referred to as pair-wise matching technique and it is common in twin studies. Group matching may also be done, so that the matching factor is divided into strata and the control group chosen so that their distribution is similar to that of the cases. This is called frequency matching technique. Matching is only done on the most important confounding variables; otherwise it can lead to matching of the risk factor under study. Couple with that, the cost of matching may be too high, especially if there are several matching variables and if suitable matches are difficult to find. These problems may inadvertently prolong the duration of the study. Errors of confounding variables may further be removed by statistical methods of stratification and regression techniques.

3.3.3 Measurement of Exposure

Criteria and definitions of exposure or variables, which may of etiological importance, are specified at the beginning of the study. Information about exposure is obtained in a similar way among the cases and controls. This may be done by means of interviews, questionnaires, examinations, focus group discussions, use of check lists, studying past records such as hospital records or employment records, etc. Bias should be avoided while measuring the exposure by blinding the investigator.

3.3.4 Analysis and Interpretation

This involves two steps:

1. Exposure rates among cases and controls to suspected etiological factor
2. Estimation of disease risk associated with exposure (odds ratio).

	Cases With lung cancer)	Controls Without lung cancer)
ers	A	B
smokers	C	D
	(a+c)	(b+d)

- **Exposure rates:**

Cases = a/(a+c), Control = b/(b+d)

- **Estimation of Risk (Odds Ratio):**

Odds ratio is also known as cross-product ratio. It is a key parameter in the analysis of case control study. Odds ratio is a measure of the strength of association between risk factor (exposure or susceptibility factors) and health outcome, measured as dichotomous variables. It is closely related to relative risk, another measure of the strength of association between exposure and outcome when both variables are expressed as a contingency cross-classification. Odds ratio is applicable in the following conditions:

1. The disease being studied must be relatively rare.
2. The cases must be representative of those with the disease.
3. The controls must be representative of those without the disease.

3.3.5 Advantages of Case-Control Studies

1. Case control studies are relatively easy to conduct
2. Case control studies are comparatively rapid and inexpensive to carry out.
3. They used to investigate rare diseases or diseases about which little is known.
4. Case control studies constitute no particular risk to subjects.
5. Results can be obtained relatively quickly because the diseases under study are already present, or have occurred in the past.
6. Retrospective studies can identify more than risk factor in the same set of data.
7. Rational preventive and control measures can be instituted after identifying the risk factors.
8. There are no attrition problems because case control studies do not require follow-up of individuals into the future.
9. Ethical problems are minimal in case control studies.

3.3.6 Disadvantages of Case-Control Studies

1. High chances for bias.
2. Validation of information obtained is difficult or sometimes impossible.
3. Selection of an appropriate control group may be difficult.
4. We cannot measure incidence, and can only estimate the odds ratio.
5. Not suited to the evaluation of therapy or prophylaxis of a disease.

3.4 Sources of Bias in Case-Control Studies

Bias is recurrent or systematic error in the determination of association between the exposure or risk factor and disease. Bias affects the relative risk estimate by increasing or decreasing it. The different types of bias in epidemiological studies are as follows:

3.4.1 Bias due to confounding

A bias may arise when a confounding factor that affects both exposure and the disease under study is not taken into account when selecting the controls and cases. This type of bias can be removed by adequately matching the controls to the cases.

3.4.2 Memory or Recall Bias

Memory or recall bias occurs when cases and controls are asked questions about their past medical history, and cases have a different recall compared with the controls. The cases are more likely to remember the existence of certain events, habits or factors than the controls who are healthy persons.

3.4.3 Selection Bias

This type of bias occurs when the cases and controls are not representative of the cases and controls in the general population. There may be systematic differences in the characteristics of cases and controls.

3.4.4 Interviewer's Bias

Bias may occur when the interviewers knows the hypothesis that is being tested by the study, and is also aware of who the cases are. He may therefore question the cases more thoroughly than the controls regarding a positive history of the suspected etiological factor. Allotting equal time to the questioning of cases and control can check this type of

bias. Double blinding, where the interviewer and the study subjects do not know whether they are cases or controls, can also eliminate interviewer's bias.

4.0 CONCLUSION

The case-control study, the other type of observational study is more common than the cohort study. In a case-control study, we enrol a group of people with disease ("cases") and a group without disease ("controls") and compare their patterns of previous exposures. The key in a case-control study is to identify an appropriate control, or comparison, group, because it provides our measure of the expected amount of exposure.

5.0 SUMMARY

In this unit, you have learned about the key features of case-control studies. You should now be able to describe the features of a case-control study, understand the importance of the selection of cases and control, discuss the potential sources of bias, and understand the strengths and limitations of this sort of study.

6.0 TUTOR MARKED ASSIGNMENT (TMAs)

There have been recent media reports of a suspected association between the consumption of an Eastern Nigerian delicacy popularly known as "*Nkwobi*" and the development of breast cancer among Nigerian women. Design a case-control study to confirm or refute these reports.

7.0 REFERENCES/FURTHER READING

Chika Onwasigwe, (2004). Principles and Methods of Epidemiology. Institute for Development Studies, University of Nigeria, Enugu.

Ekanem EE, (2015). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Akinsola OJ, (2019). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Lucas AO and Gilles HM (2003). Short Textbook of Public Health Medicine for the Tropics, (4th ed.).

Rothman KJ, Greenland S. and Lash T (2011). Modern Epidemiology, (3rd ed.).

An Introduction to Applied Epidemiology and Biostatistics, (1992). (2nd ed). Center for Disease Control.

UNIT 3 ANALYTIC EPIDEMIOLOGY II: COHORT STUDY

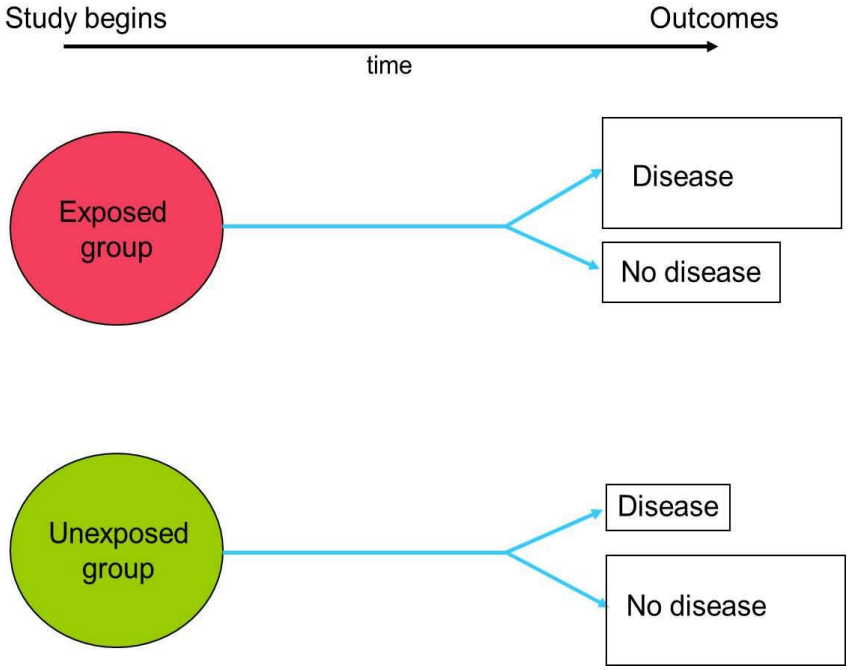
CONTENTS

1.0 INTRODUCTION

Cohort study is known by a variety of other names: prospective study, longitudinal study, incidence study, and forward-looking study. It is a type of observational study in analytic epidemiology, which aims to obtain additional evidence to support or reject the existence of an association between a suspected risk factor or cause and a disease or health status outcome. In this type of study, a group of people known as cohort, who are considered to be free from of a given disease, but who vary in exposure to a supposed etiological factor, are followed up in time to see who will develop/developed the disease under study. A cohort is defined as a group of persons who share a common characteristic or experience within a define time period (e.g. age, occupation, students in class, exposure to a drug, exposure to particular vaccine, pregnancy, etc.). Recruitment into a cohort is on the basis of exposure.

A birth cohort is a group of people born on the same day or in the period of time. A birth cohort of 2003 comprises all the babies born in that year, in the defined population. A marriage cohort is a group of men and women that are married on the same day or at the same period of time. An occupational cohort is a group of people that are engaged in the same occupation. A group of persons that are exposed to the same infection, drug or vaccine within a specified period of time constitute an exposure cohort. A cohort might be a group of persons that survived a particular type of infection or health problem. These persons are classified with regards to exposure to hypothesised risk factor. The individuals represent the total population with respect to exposure and non-exposure to the risk factor. The main features of a cohort study are as follows:

- The cohorts are identified before the appearance (observation) of the disease under investigation.
- The study groups, so defined are observed for a period of time to determine the frequency of disease among them.
- The study proceeds forward from cause to effect.



2.0 OBJECTIVES

By the end of this unit, you will be able to:

- define cohort study
- describe the procedures for conducting cohort study.

3.0 MAIN CONTENT

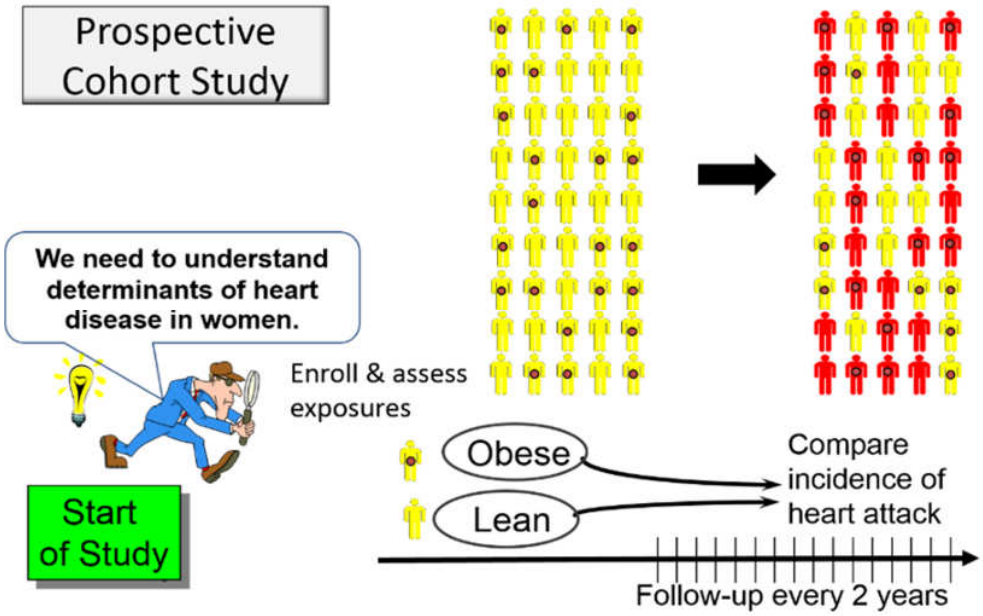
3.1 Cohort Study

One method of establishing an association between exposure and disease is to follow a population over time. In a cohort study, participants are followed over time to see whether they develop the disease of interest. Cohort studies have long been used as forms of natural experiment, since defined groups are followed as they would be in an intervention trial, although the investigator's job is purely to observe and not to intervene.

3.2 Types of Cohort Studies

3.2.1 Prospective Cohort Studies

A prospective cohort study is one in which the disease or other outcome has not occurred at the beginning of the investigation. It starts in the present and continues to the future. An example is the United States of Public Health Services Framingham Heart Study that tested the association between cigarette smoking and lung cancer. Cohorts of smokers (exposed) and non-smokers (non-exposed) are followed up in time to see who develops lung cancer. The principal finding was that smokers had more cases of lung cancer compared with non-smokers. Since the disease has not yet occurred when the study was undertaken, this was a prospective cohort design. A similar study design was the evaluation of the long-term effects of exposure of uranium and occurrence of lung cancer. A cohort of uranium miners and another comparison cohort of non-miners were observed for the subsequent development of lung cancer. The uranium miners had an excess frequency of lung cancer compared with the non-miners.

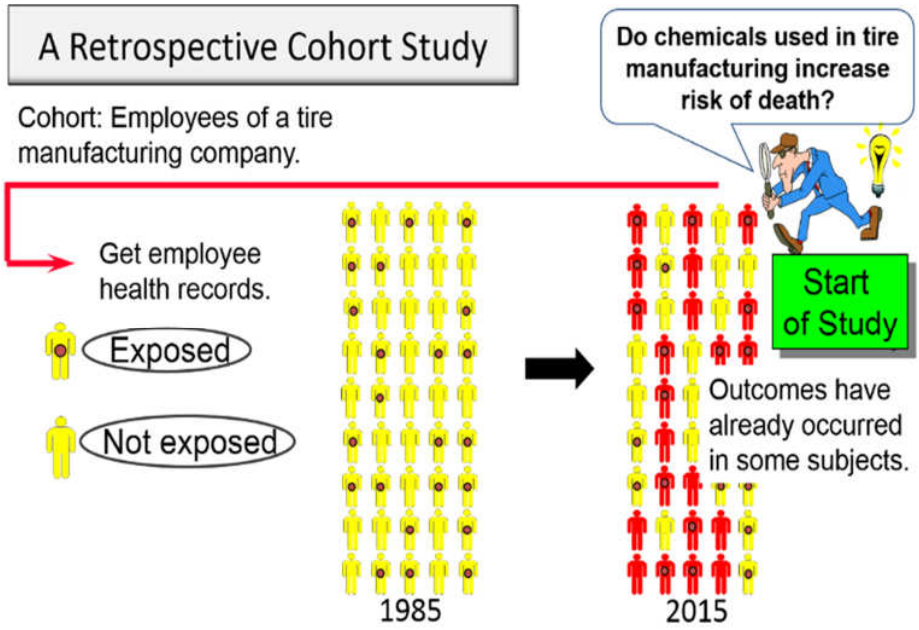


3.2.2 Retrospective Cohort Studies

Other names for the retrospective cohort study are the historical cohort study, prospective study in retrospect and non-concurrent prospective study. The outcomes of the study have already occurred before the start of the investigation. The study cohorts are selected from existing records of past events, such as medical, school and employment records. The

events are then followed forward in time from a past date fixed on the records, which may be many years previously.

Examples are studies of occupational exposures such as the role of arsenic in human carcinogenesis, the study of lung cancer in uranium miners, and the study of mortality of physician in relation to their probable exposure to radiation. A more recent study was that of a rare disease, angiosarcoma of the liver in relation with poly-vinyl chloride. Retrospective cohort studies are generally more economical and quicker to produce results than prospective cohort studies.



3.2.3 Combination of Retrospective and Prospective Cohort Studies

In this type of study, the cohort is selected from past records and assessed during a particular date that has passed for outcome. The same cohort is then followed up prospectively into future for further assessment of outcome. This design of study was employed in the study of the effects of radiation in the etiology of leukaemia and aplastic anaemia, following radiotherapy for ankylosing spondylitis. A cohort was assembled in 1955 consisting 13,352 patients suffering from ankylosing spondylitis who had received large doses of radiation therapy during a 21-year period spanning 1934 t0 1954. Mortality from leukemia and aplastic anaemia was found to be higher in the study cohort compared with the general population. A prospective component was added to the study in 1955 by identifying deaths that occurred in subsequent years.

3.3 Procedures for Conducting Cohort Study

The procedure for conducting cohort include

1. Selection of study subjects
2. Collection of data on exposure
3. Selection of subjects for comparison
4. Follow-up
5. Analysis

3.3.1 Selection of study subjects

The study subjects are either selected from the general population or are made up of special groups of people that can be readily studied and have different degrees of exposure to the etiological factor. The general population is studied when the exposure to the risk factor for a health problem is fairly frequent. The study subjects should reside in well-defined geographical, political and administrative areas. A sample that is representative of the general population may be used if the population is very large. The special groups are either made up of select groups or exposure groups. Selected groups are homogenous in constitution and may be professional groups (e.g. doctors, nurses, teachers, engineers, civil servants, traders, farmers), pregnant women, undergraduates, war veterans, volunteers, etc. These groups are readily accessible for prolonged follow-up. Another type of special groups are exposure groups made up of persons that have been exposed to the suspected causal factor of a disease or health problem. If the exposure is rare, it is more economical to study the exposure cohort. Readily accessible exposure cohorts are workers in industries and those in high-risk situations like radiologists that are exposed to harmful rays. These cohorts are classified according to the degree or duration of exposure to the suspected factor, for subsequent analytical study.

3.3.2 Collection of Data on Exposure

Information may be obtained directly from the cohort members themselves by means of interviews or questionnaires. In advanced countries with large literate population, the questionnaires may be mailed, thereby offering a simple and more economic study. Information may be available from past records, especially those that are specific in nature and cannot be easily given by lay people. Examples are details of medical treatment, types of surgery performed, types and doses of radiotherapy administered, etc. Information may also be obtained from medical examination or special laboratory tests such as measurements of weight, blood pressure, serum cholesterol, ECG, CAT scan, etc. Environmental surveys may be used to obtain information on exposure levels of the

suspected causal factor in the environment where the study subjects lived and worked. Basic information on demographic variables are also obtained from the cohort members, as these may have some influence on the disease under study.

3.3.3 Selection of Subjects for Comparison

In some cohort studies, the study subjects are classified into several comparison groups based on degrees or levels of exposure to risk factor before the disease develops, e.g. smoking, blood pressure, serum cholesterol, etc. These are known as internal comparisons. These groups are compared in terms of their subsequent morbidity and mortality rates. External comparisons can be carried out with outside comparison groups that serve as the control groups. This is necessary when information on the degree of exposure of the study groups is not available. Cohort studies can be carried out between smokers and non-smokers; radiologists and Public Health physicians, etc. However, the study and control cohorts should be similar in demographic and other important variables, but not the ones under study.

Comparison can be made with the general population rates in the same geographic area as the exposed people. An example is the comparison of frequency of lung cancer among uranium miners with lung cancer mortality in the general population where the miners resided. A similar study design is the comparison of frequency of cancer among asbestos workers with the frequency in the general population in the same geographic area. Rates of disease in the study and control subjects are considered in terms of age, sex and other variables considered being important for the study. Rates in the control cohort are applied to those in the study cohort to determine the “expected” values in the absence of exposure. The effect of the risk factor under study can be estimated by the ratio of the “observed” and “expected” rates in the control group.

3.3.4 Follow-up

The procedures for regular follow-up visits of participants in cohort studies are as follows

- Periodic medical examination of each member of the cohort
- Review of physician and hospital records
- Routine surveillance of death records
- Mailed questionnaire, telephone calls and periodic home visits

The greatest amount of information is obtained from periodic medical examination of each member of the cohort. Loss to follow-up may result from death, change of residence, migration or withdrawal of occupation.

As these losses may introduce bias to the results, it is necessary to obtain basic information on outcome for those who cannot be followed up in detail for the whole duration of the study. However, it is recommended that researchers should achieve as close to a 95 percent follow-up as possible.

3.3.5 Analysis

At the end of the data collection, analysis is done to determine the incident rates among exposed and non-exposed cohorts, and also to estimate the risk of outcome in the two groups.

Hypothetical example of cigarette smoking and lung cancer.

	Developed Lung Cancer	**Did not develop Lung Cancer**
Smokers	A	B
Non-smokers	C	D
Total	(a+c)	(b+d)

3.4 Estimation of Risk

The estimation of the risk of disease or death among the exposed and non-exposed cohorts can be carried out by means of two indices: (1) Relative Risk (2) Attributable Risk.

3.4.1 Relative Risk

Relative Risk (RR) is the ratio of the incidence of the disease (or death) among exposed and the incidence among the non-exposed. It is also known as the "risk ratio".

$$\textit{Relative Risk (RR)} = \frac{\textit{Incidence of disease (or death) among exposed}}{\textit{Incidence of disease (or death) among non-exposed}}$$

In etiological studies where the units of observation are aggregate of persons rather than individuals, relative risks form the direct measure (or index) of the "strength" of association between suspected causal factor and disease. It does not necessarily imply causal association. In the interpretation of the calculated vales of the Relative Risk, quotient of 1.0 signifies that the incidence rate is the same among exposed and non-exposed subjects and indicates a lack of association between exposure and disease. A Relative Risk above 1.0 suggests that exposed people are at higher risk of disease than non-exposed persons.

3.4.2 Attributable Risk

Attributable Risk (AR), also known as **Risk difference**, is the difference in incidence rates of disease (or death) between an exposed group and non-exposed group. It is a measure of the residual risk of a disease due to exposure to a specific agent or risk factor. It is determined by subtracting the incidence rate of the disease in persons not exposed to the agent or risk factor(s) from the corresponding rate among exposed persons. The difference is the rate (risk) of disease occurrence attributable to exposure to the specific agent or risk factor. The applicable formula for calculating attributable risk is:

$$\text{Attributable Risk (A.R.) among persons} = \left[\text{Rate of Disease among Exposed Persons}\right] - \left[\text{Rate of Disease Non-Exposed}\right]$$

= number of cases of disease per "k" population in time interval "t" that is due to exposure to specified risk factor(s). Where "k" and "t" are as defines in the rates used to express risk in the formula.

Attributable risk percent is another indicator derivable from cohort studies data that is readily applied in interpretations of the risk of a disease due to exposure to a given risk factor. It is defines as the proportion of the overall risk of a disease in a population, expressed in percent, which can be reduced by eliminating specific exposure to specified risk factors, or by mitigating specific susceptibility factors, as the case may be. It is derived from the concept of population attributable risk. Indeed, it is more appropriately termed population attributable risk percent. It is given by the formula:

$$\text{Population Attributable Risk Percent} = \frac{\left[\text{Incidence Rate of the Disease in Total Population}\right] - \left[\text{Incidence Rate of the Disease in Non-exposed/ Non-susceptible persons}\right]}{\text{Incidence Rate in Total Population}} \times 100$$

= proportion of risk of occurrence of given disease in the population in the specified time interval that is due to exposure to the specified risk factor.

There are two ways of interpreting population attributable risk percent indicator. The first is as the proportion of risk of occurrence of given disease in the population, in the specified time interval, that is due to exposure to the specified risk factor. Thus, it is a good way of expressing the impact of magnitude of specified risk factors. The second way is as

the proportional amount, in percentage, by which a disease in the referent population can be reduced by controlling or eliminating exposure to the risk factor in question. Thus, it is a good indicator of the effectiveness of prophylactic intervention regimen, when such interventions are implemented as measures designed to reduce exposure or susceptibility to a disease.

3.4.3 Advantages of Cohort Studies

1. Incidence rate can be calculated among those exposed and non-exposed
2. Prospective studies permit observation of many outcomes
3. Cohort studies provide a direct estimate of relative risk
4. Dose-response ratios can be calculated from cohort studies
5. Certain forms of bias arising from misclassification of study subjects into exposed and exposed groups can be prevented

3.4.4 Disadvantages of Cohort Studies

1. They are usually expensive and take a lot of time
2. Cohort studies are generally unsuitable for investigating rare diseases or diseases with low incidence
3. There may be problem of selection of comparison groups that are representative of the exposed and unexposed segments of the population.
4. There may be loss of study subjects to follow-up as a result of lack of interest, migration or death from other causes.
5. There may a change of status of subjects with respect to variable of interest.
6. There may be changes in diagnostic criteria and standard methods over time, affecting the classification of individuals as diseased or not diseased.
7. There may be administrative problems as a result of loss of experienced staff, loss of funding and the high cost of extensive record keeping required.

4.0 CONCLUSION

Cohort studies are analytic epidemiological research studies in which a group of subjects (persons) who share a common experience within a define time period (a cohort) and who are considered to be free of a given disease but who vary in exposure to supposed risk factor are followed up over time in order to determine differences in the rate at which disease develops in relation to exposure to the factor. The unit of observation and analysis is individuals (not group). The basic feature of cohort studies is an Intuitive approach to studying disease incidence and risk factors which

starts with a population at risk referred to as a cohort, then measures their relevant characteristics at baseline. The study group is subsequently followed up over time with either surveillance of routinely-detected cases (people get symptoms and go to their health care providers, who diagnose them or, sometimes, people get screened and then referred for diagnosis) or re-examination of people periodically or at the end of the follow-up.

Finally, event (disease/health status outcome) rates in people with and without characteristics (exposure/susceptibility factors of risk) of interest are compared. Several indictors and outcome measures of great utility in Public Health Practice re applicable to the cohort study technique. They include:

- Incidence rate (risk) of health status outcome of interest in the exposed persons
- Incidence rate (risk) in the unexposed persons;
- Relative risk or Risk ratio;
- Attributable risk (Risk difference);
- Population attributable risk;
- Attributable risk percent;
- Population attributable risk percent;
- Cumulative incidence

For emphasis, an important distinction between the two analytic study designs presented here is that in an observational cohort study, subjects first are enrolled on the basis of their exposure, and then are followed to document occurrence of disease or health status outcome of choice. In an observational case-control study, subjects are first enrolled according to whether they have the disease or not, then are questioned or tested to determine their prior exposure.

5.0 SUMMARY

In this unit, you have learned about the key features of cohort studies which involve study design where one or more samples (called cohorts) are followed prospectively and subsequent status evaluations with respect to a disease or outcome. In addition, you should be able to discuss the potential sources of bias, and understand the strengths and limitations.

6.0 TUTOR- MARKED ASSIGNMENT (TMA)

Oral contraceptives (OCs) have been consistently associated with a reduced ovarian cancer risk; however, previous studies included women in older birth cohorts [born before 1964] using traditional OC formulations. Recently, there has been interest in the conduct of a cohort study to assess the relationship between OC use and ovarian cancer

among a cohort of younger women [born after 1964] and using a specific widely-used and more recent OC formulation.

a). Briefly describe the steps the authors would take to conduct this study.
b). Describe in detail the differences between a prospective and retrospective cohort study, giving examples and stating the merits and demerits of each.

7.0 REFERENCES/FURTHER READING

Chika Onwasigwe, (2004). Principles and Methods of Epidemiology. Institute for Development Studies, University of Nigeria, Enugu.

Ekanem EE, (2015). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Akinsola OJ, (2019). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Lucas AO and Gilles HM (2003). Short Textbook of Public Health Medicine for the Tropics, (4th ed.).

Rothman KJ, Greenland S. & Lash T. (2011). Modern Epidemiology,(3rd ed.).

An Introduction to Applied Epidemiology and Biostatistics, (1992). (2nd ed). Centre for Disease Control.

MODULE 5 EXPERIMENTAL, SCREENING AND INVESTIGATIVE EPIDEMIOLOGY

UNIT 1 EXPERIMENTAL EPIDEMIOLOGY

CONTENTS

1.0 INTRODUCTION

Experimental studies constitute a major area of epidemiological inquiry by which we demonstrate cause and effect relationships, otherwise known as causal inference, Public Health. They are also known as intervention studies and are carried out under the direct control of the investigator. This is contrary to what obtains in observational studies, e.g. descriptive, case control and cohort studies, where the investigator takes no action but only observes the natural course of events of outcome. Experimental studies involve some deliberate action, intervention or manipulation of natural order such as application or withdrawal of the suspected cause of a health problem. Changes may be made in causative chain by withdrawing one variable in the experimental group while retaining the same variable in the control group. The outcomes in the two groups are later compared.

Two types of experimental studies are prophylactic and therapeutic trials. Prophylactic trials are experiments conducted in human or animal subjects aimed at preventing diseases. Therapeutic trials, on the other hand, are experiments aimed at treating established disease processes. These can involve selected groups of individuals as units of observation and analysis, as in clinical trials, or whole communities or aggregates of population as units of observation and analysis, as in community trials.

The group that is studied is known as the experimental population, while the group of ultimate interest is called the reference population.

An experimental study must begin with a clearly formulated hypothesis – a supposition on conjecture put forth to account for known facts, which serves as a starting point for further investigation by which it may be proved or disproved and the true theory arrive at. The aims of experimental studies are twofold. The first is to provide "scientific proof" of etiological (risk) factors, which may permit the modification or control of disease; and the second is to provide a method of measuring the effectiveness and efficiency of health services for the prevention, control and treatment of disease and thereby improve the health of the community.

In experimental studies, the researcher controls the assignment of subjects to the treatment groups (those exposed to the risk factors under study) and the control group (those not exposed). These studies usually build in rigorous procedures designed to control several sources of bias such as randomisation, matching, blinding, etc., and also allows the researcher to introduce intervention in a controlled manner, including manipulation of dose of intervention (hypothesised risk factors), timing of the administration of intervention and the mode thereof.

In experimental studies, allocation of the study group to the different treatment/interventions/exposures under investigation is done randomly. The research protocol focuses primarily on how to measure the effect of an exposure (causal factor) on an outcome (disease/health status) with consideration of the effects of other factors (potential confounders as well as factors related to the efficacy of the delivery of the intervention). The designs of these studies also prioritise the control of biases that may threaten internal validity of findings. There is also requirement that the experimental research design builds in procedures that promote external validity in order to assure generalisability of findings to the reference population. These are similar to prospective cohort studies, in addition to sharing the usual advantages and disadvantages. However, experimental studies also have additional problems of cost, ethics and feasibility. They may be conducted in animals or human beings.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- define experimental study
- describe the procedures for experimental study.

3.0 MAIN CONTENT

3.1 Experimental Studies

Experimental studies are ones where researchers introduce an intervention and study the effects. Experimental studies are usually randomised, meaning the subjects are grouped by chance. The researchers then study what happens to people in each group. Any difference in outcomes can then be linked to the intervention. Experimental studies are less susceptible to confounding because the investigator determines who is exposed and who is unexposed. In particular, if exposure is allocated randomly and the number of groups or individuals randomised is large then even unrecognised confounding effects become statistically unlikely.

There are, of course, ethical constraints on experimental research in humans, and it is not acceptable to expose subjects deliberately to potentially serious hazards. This limits the application of experimental methods in the investigation of disease aetiology, although it may be possible to evaluate preventive strategies experimentally. For example, factories participating in a coronary heart disease prevention project were assigned to two groups, one receiving a programme of screening for coronary risk factors and health education, and the other being left alone. Subsequent disease incidence was then compared between the two groups. The main application of experimental studies, however, is in evaluating therapeutic interventions by randomised controlled trials.

3.2 Animal Experiments

Animal studies have contributed immensely to man's knowledge about himself and his environment. These are used for the following purposes:

- To reproduce human disease in animals under laboratory conditions, in order to confirm etiological hypothesis and to study pathogenetic or mechanism phenomena or mechanism.
- To test the efficacy of prophylactic and therapeutic measures such as drugs and vaccines.
- To complete the natural history of disease.

The advantages of animal experiment are that the experimental animals can be bred in laboratories and easily manipulated according to the wishes of the investigator. These animals can reproduce more rapidly than human beings and thereby enable the investigator to carry out certain experiments, (e.g. genetic experiments) that will involve several generations. Animal experiments are of growing relevance in the rapidly emerging area of molecular epidemiology. An important limitation of

animal experiments is that not all human diseases can be reproduced in animals. In addition to that, not all conclusions arising from animal experiments can be strictly applied to human beings. Certain difficulties may be encountered in extrapolating findings from animal experiments in man.

3.3 Human Experiments

Human experiments are used to investigate disease etiology and to evaluate the preventive and therapeutic measures. These studies are necessary especially in diseases that cannot be reproduced in animals. A classic example of a human experiment is the clinical trial performed by James Lind in 1747 among twelve soldiers, who were suffering from scurvy. He divided the patients into six pairs and administered different dietary supplements to their usual food. These supplements included cider, vinegar, nutmeg, garlic, lemon, orange, etc. After six days, the group with lemon and orange recovered fully and were fit for work. Another example of human experiment is Goldberger's classical experiment in 1915, investigating the etiology of pellagra. He was able induce pellagra in human beings by administering diets that are deficient in nicotinic acid. He therefore proved pellagra to be nutritional deficiency disease as was then supposed.

Human beings have since participated in experimental studies of diverse disease entities. However, ethical and logistic considerations have limited human involvement in certain experimental studies. The benefits of the experiment have to be weighed against the risks involved. The volunteers should also be made fully aware of all possible consequences of the experiment. When an illness is life threatening, it would be unethical to suspend a known procedure or drug in order to test the efficacy of a new therapeutic regimen. It is also ethically unacceptable if a drug is brought into general use without first establishing its effectiveness by controlled trials.

There are two major types of experimental study designs:

- Randomised controlled trials. These involve a process of random/allocation assignment of subjects to experimental (intervention) and control groups.
- Non-randomised or "non-experimental" trials. These do not follow strict randomisation but have sound theoretical basis for conclusions.

3.4 Randomised Controlled Trials

Randomised controlled trials are scientific techniques that are used to evaluate methods of treatment and treatment. They are epidemiologic experimental studies to assess the validity widely used or recent preventive and therapeutic procedures. Randomised controlled trials are used to evaluate new programmes or new therapies. The steps in conducting a randomised controlled trial are as follows:

3.4.1 Drawing a protocol

This specifies the aims and objectives of the study, including the research questions, criteria for the selection of study and control groups and treatment to be applied. Standardisation of working procedures and schedules, as well as modalities for evaluation of outcome of the study is specified. The protocol is aimed to prevent bias and reduce the sources of error in the study. Pilot studies may be carried out before the protocol is completed in order to assess the feasibility or operational efficiency of certain procedures, their unknown effects, or the acceptability of certain policies.

3.4.2 Selecting Reference and Experimental Populations

The reference population is also known as the target population. This is the population that is expected to benefit from the findings of the trial if successful, e.g. vaccine or drug trial. It may comprise the population of a wide geographic area or that of a smaller group such as school children, pregnant women attending antenatal clinic, etc.

Experimental or study population is derived from the reference population, where it is randomly chosen. It is beneficial to select a stable population that will be cooperative in the study and also make follow-up easier. The participants or volunteers must give “informed consent”. This means they must agree to participate in the trial after receiving full information on the aims or purpose, procedures and possible dangers of the trial. Most funding organisations e.g. Tropical Diseases Research (TDR) maintains that an ethical clearance by a disinterested committee in the investigator’s institution must be obtained before the programme can be funded. This protects the rights of study subjects. The participants must be representative of the reference population. They should be qualified or eligible for the trial. This means that they should fall into the inclusion criteria for the trial and should be fully susceptible to the disease under study. The volunteers to a study are likely to be different from those who refused to participate; and this may affect the outcome under investigation in many ways.

3.4.3 Randomisation

Randomisation or random allocation is the process of allocating participants into "study" and "control" groups; to receive or not to receive an experimental preventive and therapeutic manoeuvre or intervention or intervention. The purpose of the procedure is to eliminate "bias" and allow for comparability. Randomisation performs three functions:

- It eliminates selection bias on the part of the participants and investigators.
- It creates groups that are comparable in all factors that influence prognosis.
- It gives validity in the statistical treatment of data.

In randomisation, the investigator has no control over allocation of participants into study and control groups, thereby eliminating "selection bias". By random allocation, every individual has equal chance of being placed into either the study or control group. The two groups formed are:

- The experimental or study group, which receives a new drug, vaccine or other procedure, and
- The control group, which receives no treatment, a placebo (dummy) procedure, or a standard (old) form of therapy.

The unit of randomisation may be an individual, a family unit, a hospital, or even a community. The study and control groups should be similar in certain variables or characteristics that affect the outcome of the experiment, e.g. age and sex. Individuals are stratified into subgroups in terms of the variable, out which the study and control groups are then selected. Randomisation is best achieved by using a Table of Random Numbers. In analytical studies, comparability is usually achieved by matching the diseased individuals to non-diseased ones, or exposed to non-exposed. There is no randomisation in this case because the participants have already been matched.

3.4.4 Manipulation or Intervention

Manipulation or intervention is carried out by deliberately applying, withdrawing or reducing the suspected causal factor, e.g. a drug, vaccine, dietary component, a habit etc. Manipulation creates an independent variable (e.g. drug, vaccine, a new procedure), whose effect is then determined by measurement of the final outcome i.e. the dependent variable (e.g. incidence of disease, survival time, recovery period).

3.4.5 Follow-up

This involves the examination of the study and control groups at define time intervals, using a standard procedure, with equal intensity, under the same circumstances until final assessment of outcome. The duration of trials depends on the time when appreciable difference (e.g. mortality) is demonstrated. Follow-up may therefore be short or long depending on the study undertaken. Loss to follow-up may be death, migration or loss of interest. This is known as attrition. Every effort should be made to eliminate the losses to follow-up.

3.4.6 Assessment

Finally, the outcome of the trial is assessed in terms of positive and negative result. Positive results are beneficial results such as reduced incidence or severity of disease, reduced cost of health service, etc. Negative results are severity and frequency of side effects and complications, death, etc. The incidences of positive and negative results are compared in experimental and control groups, and the difference tested for statistical significance. The data may be analysed, as they are collected, using appropriate techniques (sequential analysis). However, it is better to analyse the results at the end of the trial.

3.4.7 Blinding

Blinding is also known as masking. It is a technique that is adopted to reduce these problems. Blinding removes bias in assessment of outcomes from the expectations of the participant, the investigator or the person analysing the results of the study. It ensures that the outcome of the therapeutic is assessed objectively. The types of blinding are as follows:

- *Single blind trial:* The participant is not aware whether he belongs to the study group or the control group. Only the investigator knows the group to which the subject is assigned.
- *Double blind trial:* A further protection against bias is the double blind study. Neither the investigator nor the participant knows which group the participant was placed in nor the type of treatment he received.
- *Triple blind study:* The participant, the investigator ad the person analysing the data is all unaware whether the participant was placed in the study group or the control group. The code is broken at the end of the therapeutic trial.

3.4.8 Non-Randomised Trial

It is always feasible to conduct experimental studies on human beings as a result of ethical, administrative and other reasons:

- Induction of diseases for experimentation has not been practiced on human subjects.
- Certain preventive measures can only be applied to groups or communities, without restricting them to randomly selected individuals.
- When disease frequency is low and the natural history long (e.g. cancer of the cervix), randomised controlled trials may require follow-up of thousands of people for 10 or more years. The logistics for this is not often possible.

In such cases, non-randomised or non-experimental trials are conducted. However, these studies are not accurate and valid in their comparability as the randomised controlled trials. As clearly evident, randomisation is a very important procedure in experimental studies. It involves the random assignment implies that individuals or communities are allocated randomly to each study group and that allocation of subjects to a group is independent of the allocation of other subjects, the purpose of which is to ensure that differences between treatment and control groups or towns/populations in potential confounders and levels of other important variables arise by chance alone. The random allocation of subjects to groups also ensures that neither the observers nor the individual participating in the study can influence, by way of personal judgment or prejudice, who is allocated to receive which treatment. In a community trial, randomisation occurs at the level of the community, subjects within a community are not randomly assigned to treatment or control group. However, for practical reasons the two largest community trials in the US did not fully randomly allocate towns, and this may undermine the confidence with which the results of these studies are judged.

Establishing causal inference is also important in experimental studies. Once the issues of bias and confounding factors have been mitigated through adequate design of the study and the role of chance bordering on how likely it is that what was found is a true finding, epidemiologists apply certain theories of causation to arrive at conclusions as to causal relationships. The Braxton-Hill's criteria are one such formwork applied. The postulates are as follows:

Strength of Association – *S*trong associations are more likely to be causal than weak associations because, if they could be explained by some other factor, the effect of that factor would have to be even stronger than the observed association and therefore would have become evident. Weak

associations, on the other hand, are more likely to be explained by undetected biases.

Biologic Credibility – This criterion refers to the biological plausibility of the causal hypothesis, an important concern but one that is far from objective or absolute. It is based on the assumption that an implausible explanation is likely to be incorrect explanation to causal relationship. Biological plausibility is too often not based on logic or data, but only on prior beliefs. This makes it prone to bias.

Specificity - The criterion of specificity requires that a cause leads to a single effect, not multiple effects. Specificity can be used to distinguish some causal hypotheses from non-causal hypotheses, when the causal hypothesis predicts a relation with one outcome but no relation with another outcome. Thus, specificity can come into play when it can be logically deduced from the causal hypothesis in question

Consistency with other associations-refers to the repeated observation of an association in different populations under different circumstances. Lack of consistency, however, does not rule out a causal association, because some effects are produced by their causes only under unusual circumstances. More precisely, the effect of a causal agent cannot occur unless the complementary component causes act, or have already acted, to complete a sufficient cause. Consistency serves only to rule out hypotheses that the association is attributable to some factor that varies across studies.

Time Sequence – The temporality criterion refers to the necessity for a cause to precede an effect in time. This criterion is inarguable, insofar as any claimed observation of causation must involve the putative cause C (risk factor) preceding the putative effect D (disease/health status outcome).

Dose-Response Relationship -. this criterion, often interpreted in terms of "Biological gradient", refers to the presence of a unidirectional dose–response curve. We often expect such a monotonic relation to exist. For example, more smoking means more carcinogen exposure and more tissue damage, hence more opportunity for carcinogenesis. Some causal associations, however, show a single jump (threshold) rather than a monotonic trend; an example is the association between DES and adenocarcinoma of the vagina. A possible explanation is that the doses of DES that were administered were all sufficiently great to produce the maximum effect from DES. Associations that do show a monotonic trend in disease frequency with increasing levels of exposure are not necessarily causal; confounding can result in a monotonic relation between a non-

causal risk factor and disease if the confounding factor itself demonstrates a biological gradient in its relation with disease

Analogy – The analogy criterion provides a source of more elaborate hypotheses about the associations under study; absence of such analogies only reflects lack of imagination or experience, not falsity of the hypothesis.

Experimental evidence - From Hill's examples, it seems that what he had in mind for experimental evidence was the result of removal of some harmful exposure in an intervention or prevention program, rather than the results of laboratory experiments. The lack of availability of such evidence would at least be a pragmatic difficulty in making this a criterion for inference. Logically, however, experimental evidence is not a criterion but a test of the causal hypothesis, a test that is simply unavailable in most circumstances. Although experimental tests can be much stronger than other tests, they are often not as decisive as thought, because of difficulties in interpretation.

Coherence - the term coherence implies that a cause-and-effect interpretation for an association does not conflict with what is known of the natural history and biology of the disease. On the other hand, presence of conflicting information may indeed refute a hypothesis, but one must always remember that the conflicting information may be mistaken or misinterpreted.

4.0 CONCLUSION

In conclusion, you have learnt that the purpose of an epidemiologic study is to quantify the relationship between an exposure and a health outcome. The hallmark of an epidemiologic study is the presence of at least two groups, one of which serves as a comparison group. In an experimental study, the investigator determines the exposure for the study subjects; in an observational study, the subjects determine their own exposure. Eexperimental study designs in epidemiology, especially the randomised control designs experimental study experimental study are ideal for establishing causal inference which are employed to determine the effectiveness and efficacy of health interventions. With their ability to rule out biases and confounding factors, the application of basic theory of causation, as postulated in Bradford-Hill's Criteria, to the findings of such studies help in concluding causal relationships.

5.0 SUMMARY

In this unit, you have learnt about the key features of experimental/intervention studies. You should now be able to describe

the features of an intervention study, understand the strengths and limitations, and be able to discuss the associated ethical and policy issues.

6.0 TUTOR- MARKED ASSIGNMENT (TMA)

Discuss issues of concern that may arise in the design, conduct, analysis and interpretation of a randomized control trial (RCT) to evaluate the effectiveness of Truvada (A pill that combines two anti-HIV drugs to prevent HIV infection through sexual intercourse.

7.0 REFERENCES/FURTHER READING

Chika Onwasigwe, (2004). Principles and Methods of Epidemiology. Institute for Development Studies, University of Nigeria, Enugu.

Ekanem EE, (2015). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Akinsola OJ, (2019). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Lucas AO & Gilles HM (2003). Short Textbook of Public Health Medicine for the Tropics, (4th ed.).

Rothman KJ, Greenland S. and Lash T (2011). Modern Epidemiology, (3rd ed.).

An Introduction to Applied Epidemiology and Biostatistics, (1992). (2nd ed.). Centre for Disease Control.

UNIT 2 SCREENING IN EPIDEMIOLOGY

CONTENTS

1.0 INTRODUCTION

Screening is defined as the presumptive identification of unrecognised disease in an apparently healthy, asymptomatic population by means of tests, examinations or other procedures that can be applied rapidly and easily to the target population. A screening programme must include all the core components in the screening process from inviting the target population to accessing effective treatment for individuals diagnosed with disease. Screening is a process – one that begins with invitation to participate and ends with treatment for appropriately identified individuals. An effective screening programme should meet the following criteria:

- Mechanisms for systematic invitation and follow-up for individuals identified by the screening test as having an abnormal finding (call and recall mechanisms);
- Participation of over 70% of the target population to be screened;
- Necessary infrastructure and resources to offer the test periodically and to adequately diagnose and treat those found to have cancer or a precancerous lesion, and;
- Robust monitoring and evaluation framework to assure quality.

In advocating screening programmes, it is important to avoid imposing models from high-resource settings with advanced health systems on countries that lack the infrastructure and resources to achieve adequate coverage of the population. Screening programmes require significant health resources, infrastructure and functional health systems to be effective. Policies on cancer screening differ markedly between countries and health system capacity. There is no single approach that fits all

situations thus necessary adaptations are needed depending on the local context.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- define screening and process for initiating the programme
- distinguish between screening and diagnostic tests.

3.0 MAIN CONTENT

3.1 Process of Screening

Screening for disease is an important function of Public Health practice. Screening can be defined as the search for unrecognised disease or defect by means of rapidly applied tests, examinations or other procedures in apparently healthy individuals. The process of screening sorts out apparently well persons who probably have a disease from those who probably do not.

3.2 Aims of Screening

The general aim of screening is to sort out from a large group of apparently health persons those that are likely to have the disease, or are at an increased risk of having the disease under study. The persons that are "apparently abnormal" should then be brought under medical supervision and treatment. Detection of disease before symptoms develop alters the natural history of the disease in a favorable manner and therefore improves the chances of preventing death and disability.

3.2.1 Uses of Screening

Screening is used for the following purposes:

- ***Case Detection***

Screening is used for the presumptive identification of unrecognised disease, which does not arise from a patient's request, e.g. compulsory antenatal screening of mothers for HIV in some health facilities. These persons are screened primarily for their health benefit. Other diseases that can be screened for are breast cancer, cervical cancer, diabetes mellitus, pulmonary tuberculosis, haemolytic disease of the newborn, etc.

- ***Control of Disease***

This is also called "prospective screening". In this type of screening, people are examined for the benefit of others. Examples are screening of immigrants for the detection of diseases like tuberculosis, syphilis and HIV, in order to protect the home population. The screening programme leads to early diagnosis of these problems and prompt treatment, thereby limiting the spread of infection to others and/or resultant death from the disease.

- ***Research***

Screening may be used for research purposes because it aids in obtaining the basic knowledge of the natural history of certain diseases such as cancer and hypertension. The initial screening gives the prevalence estimate, while subsequent ones gives the incident rate.

3.2.2 Types of Screening

There are 3 types screening

- ***Mass Screening***

Mass screening is the screening of the whole population or a population subgroup. An example is the screening of all adults in a community for pulmonary tuberculosis. This type of screening is offered to all persons in the population under study, irrespective of the particular risks individual may have for the disease under study. There may be little justification for mass screening in many instances.

3.3 Difference Between Screening and Diagnostic Tests

A screening is test not intended to be diagnostic. Persons with positive or suspicious findings must be referred to physicians for diagnosis and necessary treatment. A screening test is not a basis for treatment. Screening differs from diagnostic medical examination in that it is done on apparently healthy people. The initiative for screening comes from an investigator interested in a particular disease entity or the agency that provides care or funds for a particular research. Diagnostic tests are usually carried out on individuals with indications or those that are sick and have presented themselves for medical advice test and management. Screening is capable of wide application and usually one or a few diseases are considered. This is contrary to what obtains in diagnostic tests, where single patients are considered and all diseases are tested for.

Screening test are based on one criterion and have only one cut-off point, e.g. an individual is considered to suffer from hypertension if his systolic blood pressure is 140 mmHg and above or diastolic blood pressure 90 mmHg and above. This is unlike what obtains in diagnostic tests, where evaluation is based on a number of symptoms, signs and laboratory findings. A screening test is therefore less accurate and less expensive than a diagnostic test. It is also less time wasting for the physician, who is not often required to administer the test but only to interpret the results.

- ***High-Risk or Selective Screening***

This type of screening is applied selectively to the high-risk groups for the particular disease under study. The screening is more effective and economical. Family members can be selectively screened for diseases that are familial in origin e.g. hypertension, diabetes mellitus, breast cancer, etc. Risk factors can also be screened for because they antedate the actual disease in question, e.g. elevated serum cholesterol can lead to coronary heart disease. Preventive measures can then be applied on time before the disease develops.

- ***Multiphasic Screening***

Multiphasic screening means the application of two or more screening tests to a large number of people at one time, rather than carrying out separate screening tests for single diseases. The procedure may include administration of questionnaire, clinical examinations and a variety of measurements and investigations. All these can be performed rapidly with the appropriate staff and equipment. However, most multiphasic screening has been wasteful of resources, thereby casting doubts on their overall usefulness.

3.3.1 Evaluation of Screening Test

The following measures, calculated as percentages are used in evaluating a screening test:

(Please refer to the 2x2 contingency cross-classification table below)

Contingency Cross-classification of Screening Test Results in Relation to Diagnosis

Screening Test Results	**Diagnosis**		**Total**
	Diseased	Not Diseased	
Positive	a (True Positives)	b (False Positives)	a + b
Negative	c (False Negatives)	d (Negatives)	c + d
Total	(a+c)	(b+d)	a + b + c + d

- Sensitivity = a/(a+c) x 100
- Specificity = d/(b+d) x 100
- Predictive value of a positive test = a/(a+b) x 100
- Predictive value of a negative test = d/(c+d) x 100
- Percentage of false negatives = c/(a+c) x 100
- Percentage of false positives = b/(b+d) x 100

The following definitions apply:

- ***Sensitivity***

Sensitivity is the ability of attest to identify correctly all those who have the disease, i.e. the "true positives". An 80% sensitivity means that 80% of the diseased persons that have been screened by the test will give a "true positive" result, while the remaining 20% will give a "false negative" result.

- ***Specificity***

Specificity is the ability of attest to identify correctly those who do not have the disease, i.e. the "true negative". An 80% specificity means that 80% of the non-diseased people will give true negative results, while the reaming 20% of the non-diseased people will be wrongly classified as diseased. An ideal screening test should be 100% sensitive and 100% specific. In practice, this does not occur; sensitivity and specificity are inversely related. As one increases the other decreases and vice versa.

- ***Predictive value of a Test***

The predictive value of a screening test reflects its diagnostic power. The predictive value of a positive test indicates the probability that a patient with appositive test result has the disease in question. The more prevalence a disease is in a community, the more accurate will be the predictive value of a positive screening test.

- ***False Negatives***

False negatives concern the physician more than the epidemiologist, who is more concerned with the sensitivity and the specificity of a screening test. False negatives are patients who actually have the disease but are told that they do not have it. This may lead to a false sense of well-being and postponement of necessary treatment. A sensitive screening test has few false negatives.

- ***False Positives***

False positives are persons who do not have the disease in question but are told they have it. These normal and healthy people are therefore subjected to further, often inconvenient, uncomfortable and expensive diagnostic test until their freedom from disease is established. A screening test with a high specificity will have few false positives.

- ***Yield***

Yield is the amount of previously unrecognised disease that is diagnosed as a result of the screening programme. It depends on the sensitivity and specificity of the test, prevalence of the disease, and participation of individuals in the screening exercise.

a) Sensitivity of the test. If a test has low sensitivity, it can identify only a fraction of the diseased individuals, thereby leading to a poor yield.

b) Incidence and prevalence of unrecognised disease. The incidence of a disease influences yield. The higher the incidence of a disease, the higher is the yield of a screening test. Prevalence of a disease also affect yield. The higher the prevalence, the higher is the yield of a screening test. Therefore, screening should be aimed at a population with high prevalence of a disease. High-risk individuals are usually selected for screening, thereby increasing the yield.

4.0 CONCLUSION

Screening is the presumptive identification of unrecognised disease or defects by the application of tests, examinations, or other procedures that can be applied rapidly. Differs from diagnosis, which is the process of confirming an actual case of a disease. The purpose is to classify individuals as to whether they are likely to have disease or be disease free.

5.0 SUMMARY

You should now be familiar with the statistical methods for evaluating the validity and reliability of screening and diagnostic tests. You should be able to describe and calculate the measures of validity of a test. You should also be able to explain the relationship between prevalence and predictive values. Finally, you should also know the WHO guidelines for initiating a screening programme and be able to review epidemiological data to make a decision as to the efficacy of screening for a particular disease.

6.0 TUTOR- MARKED ASSIGNMENT (TMA)

1. a). "Screening is an epidemiological tool" Discuss.
   b). The table below shows the results of screening for prostate cancer with PS (Prostate-specific antigen).

True Disease Status

PSA Results	Biopsy-proven Cancer	No Cancer
PSA test positive	900	1142
PSA Screen negative	140	668

From these results, determine and interpret the following measures:

i Sensitivity
ii Specificity
iii Positive predictive value
iv Negative predictive vale
v Accuracy of the test
vi Prevalence

2. A screening test with 95% sensitivity and 80% specificity screened 4000 persons for a disease that has a prevalence of 4%.
   a. How many individuals screened will test be positive?

b. Out of those yielding positive results, how many will be true positives?
c. Determine the positive predictive values

7.0 REFERENCES/FURTHER READING

Chika Onwasigwe, (2004). Principles and Methods of Epidemiology. Institute for Development Studies, University of Nigeria, Enugu.

Ekanem EE, (2015). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Akinsola OJ, (2019). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Lucas AO & Gilles HM (2003). Short Textbook of Public Health Medicine for the Tropics, (4th ed.).

Rothman KJ, Greenland S. & Lash T (2011). Modern Epidemiology, (3rd ed.).

An Introduction to Applied Epidemiology and Biostatistics, (1992). (2nd ed.). Centre for Disease Control.

UNIT 3 INVESTIGATION OF DISEASE OUTBREAKS

CONTENTS

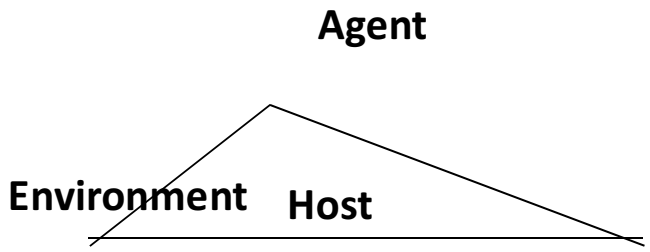
1.0 INTRODUCTION

One of the most exciting and challenging tasks facing an epidemiologist working in a public health department is investigating an outbreak. Frequently, the cause and source of the outbreak are unknown. Sometimes large numbers of people are affected. Often, the people in the community are concerned because they fear more people, including themselves, may be stricken unless the cause is found soon. There may be hostilities and defensiveness if an individual, product, or company has been accused of being the cause. Into this pressure-packed situation comes the epidemiologist, sometimes from the local health department, more often from "the outside." In this setting the epidemiologist must remain calm, professional, and scientifically objective. Fortunately, epidemiology provides the scientific basis, the systematic approach, and the population and prevention orientations that are needed.

Disease that is easily transmitted from one person to another through direct (person-to-person) transmission or indirect (vehicles, fomites, vectors) transmission.

- Direct person-to-person may be through:
- Physical contact, proximal contact as in coughing and sneezing

Epidemiologic triangle



Examples of communicable diseases:

1. Epidemic-prone:

Meningococcal meningitis, cholera, dysentery, plague, measles, viral haemorrhagic fever like yellow fever, etc.

2. Diseases targeted for eradication/ elimination:
Polio (AFP), Dracuntiasis, Measles.

3. Diseases of public health importance:
Malaria, TB, AIDS, Corona Virus (COVID-19) etc.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- define the terms epidemic, threshold, and measures of morbidity explain mortality associated with outbreak investigation
- describe the steps involved in an investigation of outbreaks
- analyse sources of data for recognition of an outbreak.

3.0 MAIN CONTENT

3.1 Reasons for Investigating Communicable Disease Outbreaks

- Define/Know what you are dealing with.
- Assess magnitude (severity of illness, potential for further spread).
- Guidance on control measures needed (to prevent further spread and minimise
- deaths).
- Mandate of MOH
- Political pressure/ legal obligation, public concern.
- Research opportunity (unknown illness, known illness but need to understand better).

3.2 Steps for Investigating a Communicable Disease Outbreak

1. Prepare for field work
2. Verify the diagnosis
3. Establish existence of an epidemic
4. Identify and count cases
5. Data analysis: time, place and person
6. Formulate and test hypothesis
7. Assess the local response capacity
8. Set up immediate control measures
9. Address the resource gaps
10. Report writing
11. Dissemination of findings
12. Intensify surveillance

3.2.1 Prepare for field work

Assemble a team (Rapid Response Team?). Assemble relevant supplies and equipment (transport media, specimen bottles, IEC, treatment guidelines & medical supplies, transport, communication means, investigation and surveillance forms, funds, fuel, etc.). Alert district authorities. Consult colleagues (microbiologist, vet, entomologist….). Review literature. Decide who will lead the team. Identify who provides support in the head office and back home. One page summary of planned activities before leaving (objectives). Arrange initial meeting(s) for your arrival.

3.2.2 Verify the diagnosis

Review clinical findings. Visit patients (interview and examine for symptoms and signs).

Laboratory diagnosis. Choose a working case definition: who is a case and who is not (by person, place, time). Should be highly sensitive. Establish index case.

3.2.3 Establish existence of an epidemic

Compare observed incidence with expected:

- No seasonality: compare with incidence from previous weeks/months,
- Seasonality: compare incidence from similar periods of earlier years.

Use action threshold.

What is a threshold?

- A marker that alerts public health officials to take action
- It uses past data to decide if an event is abnormal
- It helps to identify possible outbreaks with surveillance data

Thresholds are disease specific

Characteristics of disease

- Pathogenicity/severity
- Freq of disease within the country

Control or disease elimination program

- Internationally accepted thresholds
- External factors
- Politics/media

Money/funding
Low Frequency and Severe Disease

For low frequency and severe diseases, the threshold is often just one case
One case is enough to justify a call for action
e.g. Smallpox
- Would want to investigate after one case has been suspected

High Frequency and Less Severe Disease

- Expected baseline number of cases
- A full investigation for each case may not be necessary
- A threshold determines if action needs to be taken

If the number of cases exceeds the threshold, further action may be necessary

Examples of thresholds: CSM in the "Meningitis belt"
ALERT THRESHOLD:

- Population ≥30 000, 5 cases per 100 000 inhabitants per week
- Population <30 000, 2 cases in 1 week or an increase in the number compared to the same time in previous years

RESPONSE:

- Intensify surveillance in the area
- Prepare to conduct a mass vaccination campaign

Example: CSM in Meningitis belt

ACTION (EPIDEMIC) THRESHOLD:

- Population ≥30 000, 15 cases per 100 000 inhabitants per week (10 cases if no epidemic in 3 years and coverage <80%)
- Population <30 000, 5 cases in 1 week or an doubling in the number cases in 3 week period

RESPONSE:

- Begin mass vaccination campaign within 10 days of detection
- Inform the public

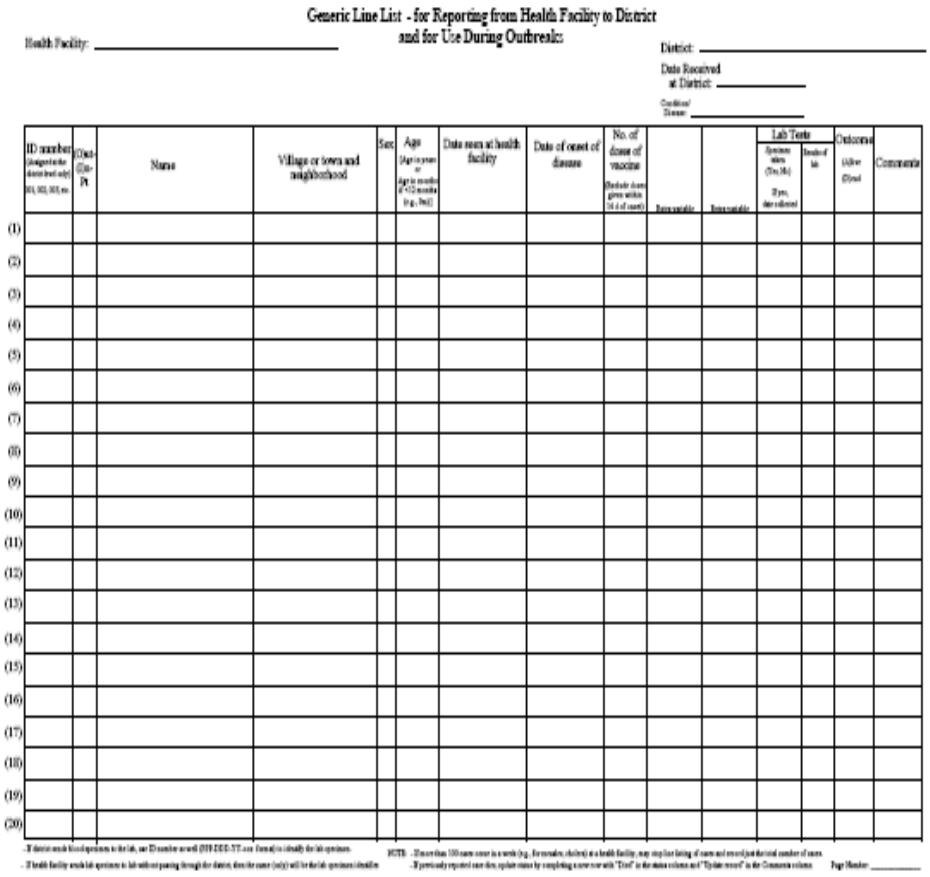
3.2.4 Identify and count cases

Using the working case definition
Collect information on cases (deaths) and line-list, e.g. identifying information: name, address.

- Demographic: age, sex, tribe.
- Clinical: symptoms and signs, date of onset, lab results, treatment, outcome of treatment.
- Exposure and risk factor information.

3.2.5 Data analysis: time, place and person

- To describe the outbreak by person (tables, bar charts, pie charts), place (spot maps) and time (histograms, graphs).
- Person: who is the population at risk (age, sex, race, occupation, medical status, etc.).
- Exposure: occupation, environment, cultural practices, socio-economic factors, etc.
- Get the population size at risk. Calculate Attack Rate, Case Fatality Rate (assess quality of case management).



Generic Line List - for Reporting from Health Facility to District and for Use During Outbreaks

Health Facility: _____ District: _____

Date Received at District: _____

Condition/ Disease: _____

	ID number	Check (Out-Pt	Name	Village or town and neighborhood	Sex	Age	Date seen at health facility	Date of onset of disease	No. of dose of vaccine			Lab Tests		Outcome	Comments
(1)															
(2)															
(3)															
(4)															
(5)															
(6)															
(7)															
(8)															
(9)															
(10)															
(11)															
(12)															
(13)															
(14)															
(15)															
(16)															
(17)															
(18)															
(19)															
(20)															

3.2.6 Formulate and test hypothesis

Hypothesis (or hypotheses) should address:

- Source of the agent.
- Mode of transmission.
- Exposures (risk factors).

Compare cases with controls in respect to exposure. Calculate OR, Chi-Square test, look up p-value. If sure of the cause, then may need only to study the cases.

3.2.7 Assess the local response capacity

- Ascertain number and type of staff available locally?
- Which drugs/ medical supplies/ guidelines are available to treat the cases?
- What has been done in terms of epidemic response?
- What steps have been taken to interrupt transmission?
- Has any health education been conducted?

3.2.8 Set up immediate control measures

Be guided by the Epidemiological triangle:

- Agent.
- Host.
- Reservouir.

Deal with the reservouir (if any).
Interrupt transmission.
Reduce susceptibility of the host (vaccination, chemoprophylaxis, improve nutrition, etc.). Treat cases.

3.2.9 Address the resource gaps

Laboratory support
Environmental support
Public information
Specific disease control needs in terms of:

- Personnel
- Drugs, vaccines and equipment
- Transport, communication and logistics

3.2.10 Report writing

- Describe the situation using the steps outlined above.
- Describe the need for outside assistance based on the gap in resources.
- Make conclusions on the outbreak you are dealing with.
- Give recommendations on priority activities (short term, long term) based on findings and conclusions.

3.2.11 Dissemination of findings

- Convey the report to Ministry of Health (relevant division/ program, senior/ top management).
- If epidemic has been confirmed, convey report to WHO through top management.
- Disseminate report to the LGA, state.

3.2.12 Intensify surveillance

- Maintain contact with the LGA for daily updates (cases, deaths, number admitted, number discharged, areas affected, etc.) until end of the epidemic.

- Three consecutive “nil” reports

4.0 CONCLUSION

Many health departments routinely offer a variety of programmes to control and prevent illnesses such as tuberculosis, vaccine-preventable diseases, and sexually transmitted diseases. An outbreak of a disease targeted by a public health programme may reveal a weakness in that programme and an opportunity to change or strengthen the programme’s efforts. Investigating the causes of an outbreak may identify populations which have been overlooked, failures in the intervention strategy, changes in the agent, or events beyond the scope of the programme. By using an outbreak to evaluate the programme’s effectiveness, program directors can improve the programme’s future directions and strategies.

5.0 SUMMARY

An outbreak is essentially the same thing as an epidemic, i.e., an increased frequency of a disease above the usual rate (endemic rate) in a given population or geographic area. Pandemic refers to simultaneous epidemics occurring in multiple locations across the globe. Traditionally, these terms referred to infectious diseases, but they can also be used to describe non-infectious diseases and chronic conditions, such as lung cancer and obesity. In addition, the principles of investigation are similar for all of these. This module provides a practical introduction to the steps involved in outbreak investigations, and it provides some useful tools.

6.0 TUTOR- MARKED ASSIGNMENT (TMA)

There has been Yellow fever outbreak in WAZOBIA state in the last one year. As the state epidemiologist, discuss how you would investigate the outbreak. Discuss the measures you would put in place for the effective control of the epidemic.

7.0 REFERENCE/FURTHER READING

Chika Onwasigwe, (2004). Principles and Methods of Epidemiology. Institute for Development Studies, University of Nigeria, Enugu.

Ekanem EE, (2015). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Akinsola OJ, (2019). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Lucas AO & Gilles HM (2003). Short Textbook of Public Health Medicine for the Tropics, (4th ed.).

Rothman KJ, Greenland S. & Lash T. (2011). Modern Epidemiology, (3rd ed.).

An Introduction to Applied Epidemiology and Biostatistics, (1992). (2nd ed.). Centre for Disease Control.

MODULE 6 BASIC DEMOGRAPHIC METHODS

UNIT 1 INTRODUCTION TO DEMOGRAPHY

CONTENTS

1.0 INTRODUCTION

The growing awareness of the need for data on human population and activities has continued to increase the popularity of the subject; "Demography" starting from total population counts, data requirement for development plans and other purposes have grown to more detailed information on the socio-economic, cultural, ethnic and political characteristics of a population. This increase in details of data requirements has also led to increase in the scope of Demography as a subject by increasing the complexity of the methods of collecting the data on human population, analysing and interpreting them as well as disseminating the results. The increasing interest arose partly because of the link between failed development plans and inadequate data base on the people for whom the plans were made and the dynamics of such data base.

In terms of formal definition, demography is the study of the size, composition, dynamics and distribution of human population in society. It is concerned with such direct health-related issues as sex ratio, get structure – (population pyramid), socio-economic structure, social stratification, urban rural distribution, residential patterns, marital status, etc. Such study is important because the size of the aggregate of persons that constitute the population, its standard of living, its dynamics, its social system and culture are all interrelated and have serious implications for the health and wellbeing of members of that population.

Note that Public Health has to do with health status of population or aggregates of persons. Epidemiology which is the fundamental science of Public Health is the study of "things" that befall populations. Both Demography and Epidemiology share their origin from studying mass phenomenon in population - "demos". Such study is important because demographic patterns influence the patterns of disease occurrence, the effectiveness and efficiency in the provision and utilisation of health services, the demand pressure on the physical environment and related amenities/facilities, as well as the culture and lifestyle of the aggregate of persons that constitute the population. Although Demography as a scientific study is much more encompassing, we shall concern ourselves here with the following areas:

- Examination of the size:
  - No. of persons in the population
  - Distribution:
- The arrangement of the population in space at a given time
  - Structure
- The distribution according to age and sex
  - Growth or decline:
- Change in population size over time
- Extends beyond formal statistical counting or measurement of the components of change.
- Includes consideration of the
  - Social
  - Economic
  - Historical
  - Political
- Characteristics of the population as related to the demographic process
- Extensive and accurate Statistics
  - Expensive to produce
  - But costs justifiable in economic terms:
- Essential for
  - Administrative
  - Social and Economic Planning

1.1 Sources of Demographic Data

- *Traditional Sources*
  - Periodic censuses
  - Sample surveys
  - Vital registration systems
  - Population registers

- *Other Sources*
  - Population & Socio-economic features
  - Parish registers
  - School registers
  - Direct taxation registers
- *Fertility*
  - Maternity clinics
  - Maternity and child welfare services
  - Social security
- *Mortality*
  - Hospitals and Clinics
  - Records of insurance companies and social security claims
  - Burial society registers
- *International Migration*
  - Frontier customs posts
  - Airports and sea ports

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- define the terms demography, population census and vital statistics
- enumerate and explain components of morbidity indicators.

3.0 MAIN CONTENT

3.1 Population Census

This is defined as the total process of collecting, compiling and publishing demographic, economic and social data pertaining at a specified time or times to all persons in a country or delineated area. (United Nations, 1965). This is the main source of demographic data in many countries. It requires huge undertaking and careful advance planning. The census records size, distribution and other characteristics of the population at fixed intervals. All persons living in the area together with some selected characteristics are physically enumerated and recorded.

3.11 *De facto* and *De jure* Enumerations

Total population can be enumerated under two different schemes. Namely: *De jure* count and De facto count. ***De facto*** - meaning enumeration of individuals as of where they are found in the census, regardless of where they normally reside. ***De jure*** - meaning enumeration of individuals as of where they usually reside, regardless of where they are on census day. A "*de jure*" census tallies people according to their

regular or legal residence, whereas a "*de facto*" census allocates them to the place where enumerated -normally where they spend the night of the day enumerated,

3.12 Advantages and Disadvantages of *De facto* and *De jure* Counts

In *De facto* count, it is simple and unambiguous and record only those present on census night but less suitable for administrative and planning processes such as: Taxation, Housing and Education. Besides, enumeration of the floating population is also problematic in *De facto* enumeration. However, in *De jure* count it's more suitable for administrative and planning purposes while normal residence may be difficult to define in *De jure* enumeration.

3.13 Sample Surveys

A sample is a part of a population. Therefore, sample survey is the enumeration of a part of a population. That is, information is collected only from a fraction of a population. It is one of the most important sources of demographic data, which has helped to improve the demographic data situation, especially in most developing countries of the world. This is an important source of demographic data for developing countries. Data collection is from a sample selected from the population. It is employed to arrive at estimates of demographic characteristics of the population by Size, Distribution, Mortality, Fertility and Migration.

3.13 Types of Demographic Sample Surveys

Demographic sample surveys may be conducted in a single round or multi-rounds.

(a) Single-Round Sample Survey:

In a single round sample survey, enumerators collect information from members of the selected households in one visit. The main disadvantage is that the error rate is relatively high especially the non-sampling errors, non-response error and errors of under-reporting of vital events (births, deaths, migration etc.). In spite of this drawback, single round sample survey is the most widely adopted method. This is because it is simple, flexible and easy to administer and involves relatively low cost and time.

(b) The Multi-Rounds Survey:

The multi-round survey, also known as follow-up survey, requires two or more visits in order to collect data. It is designed to address some of the short-comings of the single-round survey. Lists of persons and households compiled during the earlier visits are used in the subsequent rounds of visits at intervals to collect data on changes which may have occurred since the previous visits.

3.14 Advantages of multi-round survey

i. Multi-round sample surveys afford the enumerators opportunities to correct mistakes in the previous visits.
ii. Error rates, especially non-sampling errors, errors of omission and dating are relatively low.
iii. The difficulties in administering questionnaires may be reduced with increasing number of repetitions.

3.15 Disadvantages of multi-round survey

i. Records of vital events may be lost when persons associated with them leave their locations.
ii. The vital events (birth, death etc.) can only be collected retrospectively.
iii. The errors of omission are not completely eliminated
iv. The cost can still be high when enumerators are paid for every visit.
v. There may be an additional problem of matching current results for a particular person or household with previous ones.
vi. Dubious enumerators may report fictitious results without actually going to the field.

3.16 Advantages of Sample Survey

i. The costs involved in sample surveys are relatively low when compared with that involved in censuses.
ii. Sample surveys can be taken more frequently by Government, non-governmental organisations, research institutions, universities and even private individuals.
iii. The time taken to publish results is relatively short.
iv. It gives more accurate results when few well-trained enumerators are used.
v. It provides data simultaneously on both demographic events and the population at risk of such events
vi. Few enumerators can be motivated and closely supervised to ensure good returns.

3.17 Disadvantages of Sample Survey

i. Sampling and non-sampling errors may be high.
ii. It may not provide enough data to permit tabulation for localities.

3.2 Vital Statistics

Abanobi (2010) defines vital statistics as parameters used to assess the health status of communities and population groups. They involve the collection, analysis, and interpretation of various numerical data connected with the life and health of man within the community. Such numerical data include information concerning the occurrence of illness, disability, birth, and death in a population.

Vital statistics are systematically collected data on information including, among other things, details of live births, deaths, fetal deaths, marriages and divorces. The most common way of collecting information on these events is through civil registration, an administrative system used by governments to record vital events which occur in their populations. The information on vital events includes: Information derived from analysis of vital events, Epidemiology concentration on mortality and morbidity, Birth (Fertility & Natality), Death (Mortality), Marriage (Nuptiality), Sickness (morbidity), Migration and Divorce. Vital Registration System provides data about births, deaths, marriages and usually a legal requirement in the developed countries.

3.21 Reasons for Collection

Vital registration records are required to determine legal rights to inheritance, for issuing of birth certificates, international passports. Besides, it is used for admission to schools, employment/recruitment into public services and social security. In addition, vital registration is also used for health planning, as in determining cause of death, place of death. Finally, it is vital used for calculating or estimating rates of vital events and in population projection.

3.22 Special Analytical Techniques

There are advanced techniques for calculating events. These include: Standardisation, Life tables and Cohort analyses which will be discussed later.

3.23 RATE

Rate is the most important tool for measuring disease or death i.e. morbidity or mortality. Useful observations can be interpreted when they are related to a denominator in terms of a rate. It is helpful to say for instance that School A has more children with helminthiasis than School B without relating the occurrence to the number of children in each school. Also, it is not right to say that more cases of tuberculosis in a particular community are seen in males than in females without knowing the total number of each sex in that locality. Rate is used to measure events that are related to the population or subgroup of it, in which they occur, or special events that are related to the total events. The event may be death, birth, occurrence of disease, immunisation coverage, admission to hospital etc. Rate is usually expressed as so many events per standard population size,

e.g. 1.5 per 1000 or 15 per 10,000.

In measurement of morbidity, rate indicates the probability or risk of disease in a defined population over a specified period of time.

$$\text{Rate} = \left[\frac{a}{b}\right] K = \left[\frac{\textit{Number of events in a specified period}}{\textit{Population at risk in a specified period}}\right] X \quad K$$

Where,
a = the frequency with which an event has occurred during some specified period of time
a + b = the number of persons exposed to the risk of the event during the same period of time. K is the multiplier. Some number such as 10, 100, 1000, 10000, 100000.

3.24 RATIO

A ratio is a fraction of the form: $\left[\frac{c}{d}\right] K$

where k is some base as already defined and both ***c*** and ***d*** refer to the frequency of occurrence of some event or item. In the case of a ratio, as opposed to a rate, the numerator is not a component part of the denominator. We can speak, for example, of the person-doctor ratio or the person-hospital-bed ratio of a certain geographic area. The values of ***k*** most frequently used in ratios are 1 and 100.

3.3 Morbidity Indicators

- Incidence

- Prevalence
- Attendance rate at out-patient departments, health centres etc.
- Admission, readmission and discharge rates
- Duration of stay in the hospital
- Spell of sickness or absence from work or school

3.31 Incidence Rate

Incidence rate measures the probability that healthy people will develop disease or health-related event during a specified period of time. It indicates the rate at which new disease occurs in a defined, previously disease-free population.

$$\textit{Incidence rate} = \left(\frac{\textit{Number of new cases of a disease over a period of time}}{\textit{Population at risk in time period (Number of persons susceptible and exposed to risk of the disease)}} \right) x\ K$$

This rate which measures the degree to which new cases are occurring in the community is useful in helping determine the need for initiation of preventive measures. It is used for acute diseases.

3.32 Prevalence Rate

The prevalence rate measures the numerator of people in a population that have a disease at a given time. There are two types of prevalence rate, the point prevalence and the period prevalence. *Point prevalence* measures the probability of people having a disease at one particular point in time, whereas *Period prevalence* measures the number of people that have a disease within a given period of time. The prevalence rate of a disease in a particular locality includes the incidence rate and the average duration before it is terminated either by recovery or death. Prevalence rate can be altered when people with a disease immigrate into or emigrate from a population. Among workers, for example, any serious illness is likely to lead to absence from workplace, and consequently the prevalence rate is less than that predicted from the incidence rate and duration. Prevalence depends on the number of people that have been ill in the past and the duration of their illness.

Prevalence = Incidence X Average duration

For an acute epidemic disease like CHOLERA, EBOLA or CORONA VIRUS (COVID-19), the prevalence rate is low because the duration of illness is short as a result of either quick recovery or death. This is contrary to the situation in a chronic disease like DIABETES MELLITUS or TUBERCULOSIS.

Where the prevalence rate is high as a result of the long duration of the illness.

$$\left[\frac{\text{Total no of cases, new or old, existing at a point in time}}{\text{Total population at that point in time}}\right] \times K$$

4.0 CONCLUSION

Demography is concerned with how large (or how small) are the populations; how the populations are composed according to age, sex, race, marital status, and other characteristics; and how the populations are distributed in physical space. Demography is also interested in the changes over time in the size, composition, and distribution of human populations, and how these result from the processes of fertility, mortality, and migration. This unit discusses these topics in much more depth and detail and will provide you with a thorough introduction to demography.

5.0 SUMMARY

In summary, demography is the study of human populations with respect to their size, structure, and dynamics. The structure or composition of a population refers to the distribution of its members by age, sex, and other characteristics, such as place of residence and marital or health status.

6.0 TUTOR- MARKED ASSIGNMENT (TMAs)

1. Distinguish between population census and vital statistics
2. Briefly explain the following in demography analysis:
   a) Crude Rates
   b) Sex Ratio
   c) Child Woman Ratio
   d) Specific Rates
   e) Incidence Rate
   f) Prevalence Rate

7.0 REFERENCE/FURTHER READING

Chika Onwasigwe, (2004). Principles and Methods of Epidemiology. Institute for Development Studies, University of Nigeria, Enugu.

Ekanem EE, (2015). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Akinsola OJ, (2019). Lecture Notes on Epidemiology. College of Medicine, University of Lagos, Lagos, Nigeria.

Lucas AO & Gilles HM (2003). Short Textbook of Public Health Medicine for the Tropics, (4th ed.).

Rothman KJ, Greenland S. and Lash T (2011). Modern Epidemiology, (3rd ed.).

An Introduction to Applied Epidemiology and Biostatistics, (1992). (2nd ed.). Centre for Disease Control.

Nwogu, E.C. & Iwueze, I.S. (2009). Textbook on Introduction to Demography

UNIT 2 MEASURES OF FERTILITY AND MORTALITY

CONTENTS

1.0 INTRODUCTION

Fertility and mortality data typically derive from more than one data source. When numbers of births and deaths are derived from civil registration, for example, corresponding numbers of persons required for the calculation of rates and summary measures are usually estimated from population census data. When population censuses are used to collect data on numbers of births and deaths, they are often supplemented by surveys of various kinds, which may provide more detailed and timely data. Whatever sources of data are used, evaluation of data from each source usually involves comparisons with data from other sources.

Planning the collection of fertility and mortality data therefore involves two distinct stages. The first stage identifies what fertility and mortality data will be obtained from which sources and addresses issues of coordination between different sources. The second stage feeds this information into the planning for the civil registration system, the next population census, an upcoming household survey or whatever data collection operations are involved.

Fertility and mortality data are generated by fieldwork, during which members of the general public supply information about themselves, their families and the households in which they live, to fieldworkers representing the data-collection organisation. The completeness and accuracy of data collected by any method depend on the quality of fieldwork. Fieldworkers are in this sense the most important people in every data-collection operation. Their position at the bottom of the organisational hierarchy should not be allowed to obscure this fundamental fact.

Fundamental indicators of the level of mortality are the infant mortality rate and life expectancy at birth. The infant mortality rate indicates what proportion of infants born may be expected to die before reaching their first birthday. Life expectancy at birth indicates how long a child just born may be expected to live if this child experiences the age-specific death rates observed during a given year or other time period. Life expectancy is one of many summary measures that may be derived from a life table calculated from age-specific death rates.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- define the term fertility and components of fertility rate
- discuss the term mortality and components of mortality rate.

3.0 MAIN CONTENT

3.1 Basic Fertility Rates

A population changes through the interaction of fertility, mortality and migration. Among these three components, fertility remains the ultimate determinant of population change especially in most developing countries. This accounts for the emphasis being placed on fertility analysis. Some of the basic concepts often encountered in the discussion of fertility are reproduction are:

- **Natality:** This is a term used to describe the role of birth in population change and human reproduction.
- **Fertility:** This refers to the actual birth performance, frequency of child bearing or actual number of live births achieved by a woman or group of women.
- **Fecundity:** This refers the potential level of fertility performance, physical capacity for bearing children of a population or physiological capacity to produce.
- **Fecundability:** This refers to the probability of conceiving, measured on a monthly basis, among basis, among cohabiting women who are not pregnant, sterile or temporary or temporary infecundable.

- **Live birth:** "Live birth is the complete explusion or extraction from its mother of a product of conception, irrespective of duration of pregnancy, which after such separation, breathes or shows any other evidence of live such as beating of the heart, pulsation of

umbilical cord or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached each product of such a birth is considered live born, otherwise it is still birth". This is to be distinguished from foetal deaths comprising still births, miscarriages and abortions. Foetal deaths are of products of conception prior to birth or complete separation from mothers.

- **Reproduction**

This refers to the balance of births and deaths or the ability of a population to grow and replace itself through the interaction of birth and death. Measures, of reproducibility (also known as population replacement) are expressed in terms of a generation rather than a year or other brief periods. On the other hand, fertility measures refer to brief periods, usually one year.

3.2 Fertility Measures

The basic fertility measures commonly in use can be classified into board categories. These are the current and lifetime measures. Current fertility measures are those based on data on births in 12 months or one year. They include Crude Birth Rate (CBR), General Fertility Rate (GFR), Age-Specific Fertility Rate (ASFR), Total Fertility Rate (TFR), Gross Reproduction Rate (GRR) and Net Reproduction Rate (NRR). Current measures defined for the ever-married include Age-specific Marital Rate (ASMFR) and Duration-Specific Marital Fertility Rate (DSMFR). Life time fertility, on the other hand, refers to the total number of children a woman has ever had in her whole life up to the time of investigation. These measures include the Child Woman ratio (CWR). Mean Number of Children Born (MNCB) and completed family size.

- **Crude Birth Rate (CBR):**

This is defined as the number of live births during the year per 1000 population of the specified area. The numerator is derived from all reported live births during the period. It is therefore derivable from archival records and from census reports. The denominator is expressed as the average population of the area during the referent period; determined by calculating the mid-term population estimate. When the referent duration is one calendar year, the mid-year population estimated is average of the January 1st and the December 31st population estimated, which is interpreted to mean the estimated size of the population on July 1st of that year.

The general formula for determining the crude birth rate is expressed thus:

$$\text{Crude Birth Rate in time, t (CBR)} = \left(\frac{\text{Number of live births in t}}{\text{Average size of the referent population in time, t.}} \times 1000 \right)$$

It should be noted that this rate is crude because it is based on the total population of the area as opposed to being specific to the population at direct risk of childbirth. The denominator clearly includes people that ordinarily are not at risk of childbirth namely males, pre-puberty girls, and post-menopausal women. It is not sensitive as a measure of fertility. One would need to further refine the denominator before this rate can be interpreted as a measure of the risk of childbirth. In that case, it would be further narrowed to include only those women that can bear children who are sexually exposed to men as in marriage and living together during the woman's fertile period. With this refinement, the number of children born live per 1000 population remains what it is, crude birth rate, and only a partial measure of the rate at which a population adds new members to its fold by birth.

- **General Fertility Rate (GFR):**

The **General Fertility Rate** is stated as the ratio of the number of live births in a year for every 1000 women who are in their childbearing years during that year. Yearly, number of live births per 1000 women of reproductive age, usually 15-44yrs or 15-49 yrs.

$$GFR = \left(\frac{B}{{}_{35}W_{15}} \right) \times 1000$$

B = total live births in one year

${}_{35}W_{15}$ = total mid-year population of females of reproductive age 15-49yrs

- **Age Specific Fertility Rates (ASFR):**

The age-specific fertility rate measures the annual number of births to women of a specified age or age group **per 1,000** women in that age group.

$$f_x = \left(\frac{\text{No. of births to women in age group x, x+n}}{\text{mid-year pop of women}} \times 1000 \right)$$

in age group x, x+n

$$= \left[\frac{{}_nB_x}{{}_nW_x} \right] \times k$$

- **Total Fertility Rate (TFR):** Total no. of children a woman would have or bear from age 15 to 49 if she were to bear children according to the present schedule of Age Specific Fertility Rates (asfrs) throughout her productive ages. TFR is obtained by summing over all the Age Specific Fertility Rates (asfrs) for each year of the child bearing span.

$$TFR = \left[\sum_{x=15}^{49} f_x \right] = \left[\sum_{15}^{49} \frac{B_x}{W_x} \right]$$, But if we have n-year groupings, then

$$= n \left[\sum_{x=15}^{49} {}_nf_x \right]$$

$$= n \sum_{49} (B_x/W_x)$$

- **Gross Reproduction Rate (GRR):**

This is the total number of daughters a woman would have or bear if she experiences a given set of ASFRs throughout the reproductive ages with no allowance for mortality over this period. Thus GRR is restricted to female births only, yielding values that are approximately half as large as the TFR.

The Gross Reproduction Rate (GRR) is the average number of daughters a woman would have if she survived all of her childbearing years, which is roughly to the age of 45, subject to the age-specific fertility rate and sex ratio at birth throughout that period. This rate is a measure of replacement fertility if mortality is not in the equation.[1] It is often regarded as the extent to which the generation of daughters replaces the preceding generation of women and so on and so forth. If the value is equal to one that indicates that women will replace themselves. If the value is more than one that indicates that the next generation of women will outnumber the current one. If the value is less than one that indicates that the next generation of women will be less numerous than the current one. The gross reproduction rate is similar to the net reproduction rate (NRR), the average number of daughters a woman would have if she

survived her lifetime subject to the age-specific fertility rate and mortality rate throughout that period.

GRR = TFR x Proportion of births that are daughters. If the sex ratio at birth is m males per female, then;

$$GRR = \left[\frac{TFR}{(1+ m)}\right]$$

- **Net Reproduction Rate (NRR):**

Average number of daughters that a woman will bear if she experiences a given set of ASFRs throughout the reproductive ages with allowance made for mortality of women over reproductive years. If NRR = Ro. Then if fertility and mortality remain constant at current levels, then the population will eventually grow by (R_o-1) x 100 per cent per generation (or decline by (1-R_o) x 100 per cent per generation if $R_o < 1$) i.e. 1000 girl babies will produce in the course of their lives 1000 R_o female babies.

NRR = GRR x P(ā), Where P(a) is the probability of survival from age 0 to the average age of child bearing a. Where ā is calculated from a schedule of ASFRs and P(ā) is obtained from an appropriate life table.

In population ecology and demography, the net reproduction rate, R_0, is the average number of offspring (often specifically daughters) that would be born to a female if she passed through her lifetime conforming to the age-specific fertility and mortality rates of a given year. This rate is similar to the gross reproduction but takes into account that some females will die before completing their childbearing years. An R_0 of one means that each generation of mothers is having exactly enough daughters to replace themselves in the population. If the R_0 is less than one, the reproductive performance of the population is below replacement level. The R_0 is particularly relevant where sex ratios at birth are significantly affected by the use of reproductive technologies, or where life expectancy is low. The current (2015–20) estimate for the R_0 worldwide under the UN's medium variant model is 1.09 daughters per woman.

- **Child-Woman Ratio (CWR):**

This is the number of children per 1000 women of reproductive age. The Child-Woman Ratio (CWR) estimates fertility rates based on a ratio of children to women of child-bearing age. Basically, here's how it works. You take the number of children under the age of five within a population;

this represents the most recent trends in fertility. Then, you divide that by the number of women between the ages of 15 and 49. That's the assumed number of women who are capable of having children. Then, multiply it by 1,000 and turn it into a ratio. Some similarity with GFR. This is a useful index in the absence of vital statistics registration data.

$$CWR = \left[\frac{{}_5P_o}{{}_{35}W_{15}} \right]$$

- **Mean number of Children ever born per woman:**

This is the mean number of children ever born to a group of women of a specific age or age group. It is calculated from census or survey data on the no. of children ever born and the distribution of women. Children ever born (CEB) to women in a particular age group is the mean number of children born alive to women in that age group. The number of children ever born to a particular woman is a measure of her lifetime fertility experience up to the moment at which the data are collected.

In most cases, the mean number of children ever born is computed as the ratio of the number of children born alive to all women in a particular age group to the number of women. In cases where the total number of children born to women in the age group is not provided but tabulation is available on the distribution of women by age-group and number of children ever born, the mean number of children ever born to women in the age-group is obtained as:

$$CEB = \sum jP_j$$

Where j is the number of children and P*j* is the proportion of women in that age-group who have given birth to a total of *j* children.

3.3 Basic Concepts in Mortality Analysis

Mortality is one of the components of population change. However, unlike the other two (fertility and migration), mortality is completely out of human control and affects every segment of a population. Mortality is term used to describe the contribution of death to population change. Some of the concepts commonly found in mortality analysis include:

- **Death**

Death has been defined as “the permanent disappearance of all evidence of life at any time after live-birth has taken place”. Death may be described as “postnatal cessation of vital functions without capability of

resuscitation". In other words, this definition excludes all forms of foetal deaths (which are deaths occurring prior to completion of birth process). Foetal deaths include;

- **Stillbirth or Late Foetal deaths:** which are deaths occurring after 20 or 28 completed weeks of gestation?
- **Miscarriage:** This refers to the spontaneous or accidental termination of foetal life occurring early in pregnancy.
- **Death rate and Mortality rate**

Both death rate and mortality rate are measures of the frequency of deaths in relation to the population exposed to the risk of death. However, while death rate is used to describe the frequency of deaths in relation to the exposed population at the midpoint of an interval (i.e. central rate), mortality rate is used to describe the frequency of death in relation to the exposed population at the beginning of the interval (i.e. probability). However, in most discussions the two concepts are used interchangeably.

- **Crude Death Rate**

Annual crude rate of a given population is defined as the number of deaths per 1000 persons in that population in a calendar year. As with crude birth rate, the denominator is the estimated average population for that year or, as is the practice, the estimated mid-year (July 1) population size. The said mid-year population estimate is obtained by taking the average of the population figure as at January 1st of that year and that as at December 31st of the same year. The resulting quotient is multiplied by the constant or multiplier factor "k" equals 1000.

The general formula for determining the crude death rate is expressed thus:

$$\textit{Crude Death Rate in time, t.} = \left(\frac{\textit{No. of deaths reported in t}}{\textit{Average size of the referent population in time, t.}} \right) \quad X \quad 1000$$

The annual death rate is a generalised indicator of the health status of a population. Because it is based on total population, it is a crude rate. In its use, it is often compared to similar rate for a larger or different population. Thus the crude death rate for Federal Capital Territory, Abuja, Nigerian, can be compared to that for the whole country to see if the general health status of people in FCT, Abuja as indicated by mortality experience is poorer or better than that for the entire country. Similarly, it can be compared to that for any State or geopolitical area to see if their general mortality experiences are similar.

Such comparisons remain crude, however. It is often incautious to make such comparisons, especially when the two populations are known to differ on such important characteristic that are associated with the risk of death as age, gender, race, etc. More appropriate comparisons are made by the use of adjusted death rates. Generally, also, death rates may be specific for age, for sex, for race, or for some particular cause of death.

- **Age Specific Death Rate**

This mortality rate is based on the number of deaths among persons of a given age bracket, say 45-49 years old, in a given year per 1,000 people in the average population (mid-year population estimate) in the specified age group for that year. It is expressed in terms of the number of deaths per 1,000 population of the age group in the specified year. For example, using hypothetical figures, out of the 2,676,000 persons in the age bracket, 20-44 years in Kano State, Nigeria in 1977, there were 5,097 deaths recorded in that same calendar year. From the above information, the age-specific death rate for persons aged 20-44 years in the Kano State, Nigeria, in 1977 is calculated as follows:

$$\text{Age-Specific Death Rate For persons aged 20-44 years in 1977} = \frac{5{,}097}{2{,}676{,}000} \times 1{,}000$$

= 1.9 deaths per 1,000 persons aged 20-44 years in Kano State, Nigeria in 1977.

The rate above can also be expressed as 19. 0 deaths per 10,000 persons aged 20-44 years in Florida in 1977. This is done by simply multiplying the number of deaths per 1,000 by 10 and doing the same thing to the constant or multiplier factor to give 10,000. This does not change the value of the rate. It may, however make it easy to compare the rate figure with other population where the age-specific death rate for equivalent population is less than 1.0.

The formula is also stated as:

$$(\text{asdr or } {}_nM_x) = \left[\frac{{}_nD_x}{{}_nP_x}\right] \times k$$

Where, ${}_nD_x$ = No. of deaths to persons Aged x, x+n

${}_nP_x$ = Mid-year population of persons Age x, x+n, K = 1000

- **Infant Mortality Rate (IMR):** No. of infant deaths below 1yr of age in a year per 1000 live births during the same period.

$$= \frac{\text{No. of infant deaths}}{\text{No. of live births}} \times 1000$$

Infant mortality rate has been considered of great significance in public health and demography. It is a proxy measure of the level of development of maternal and child health services in an area. Because much of the risk of morbidity and mortality for children in the age bracket 0 -1 old derive from their immediate environment, it is a good proxy measure of the level of environmental sanitation in an area as well. This is particularly true of people who live in contexts that are typical of most African and technologically less advanced nations, as well as people living in rural and urban slum areas of some technologically developed countries. A high rate is often taken to indicate significant unmet health needs and unfavorable environmental factors such as economically disadvantaged conditions, poor nutrition, inadequate educational services, inaccessible and ineffective child health care services, and poor sanitation.

Upon closer look at this indicator, one notices that this rate has inherent problems in those populations that are experiencing rapidly changing birth rates. As can be readily appreciated, the numerator includes some infants that died in time "t" but were born in the preceding year; and some infants born in "t" who may die in the succeeding year. In a population with a stable birth rate, such as the technologically advanced countries of the United States of America, and Western Europe, such differences are likely to cancel out; this is not the case in a population that is undergoing radical change in its birth rate.

An important aspect of this rate that deserves emphasis is the concept of ***live births***. It also appears in such other rates as fetal death rate, fetal death ratio, puerperal/maternal mortality rate, perinatal mortality rate, neonatal mortality rate, and post-neonatal mortality rate. The concept of live birth is operationally defined as the complete expulsion or surgical extraction from its mother of the product of conception (foetus), irrespective of the duration of the pregnancy, which after separation from mother, breathes or shows evidence of life. Relevant evidence of life include the beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached.

- **Neonatal Mortality Rate:** No. of deaths under 4 weeks (under 28 days) divided by the number of live births.

$$\left[= \frac{\text{Infant deaths under 28days}}{\text{Live Births}} \times 1000 \right]$$

Based on the clarification given about the distinction between rates, ratios, and proportions, one would argue that this measure is not a rate but a proportion. In deed this is correct. However, by force of habit and customary practice many Public Health professional continue to

designate this rate. The use of neonatal mortality rate in place of neonatal mortality proportion here is to avoid the confusion that may be attendant on using proportion. The reader is however cautioned to appreciate that the measure is actually a proportion.

Neonatal mortality (proportion) rate is defined as the number of deaths of neonates (infants less than 28 days of age) that occurred in a calendar year per 1,000 live births in that year. Here again, the term live birth is operationally defined as in infant mortality rate above. The formula for calculating this is as shown below:

$$\textit{Neonatal Mortality (Rate) Proportion in year "t"} = \left[\frac{\textit{No. of deaths of Children aged 28 days or less in "t"}}{\textit{Total No. of live births in "t"}}\right] X \ 1{,}000$$

If in a given geographic area in 1988 there were 401,582 live births, and 2,743 deaths of infants aged zero to 28 days old were recorded, then the neonatal mortality rate for 1988 for this population was

$$= \frac{2{,}743}{401{,}581} \times 1{,}000$$

= 6.83 neonatal deaths per thousand live births in 1988.

It should be noted that neonatal deaths are considered as deaths occurring within the first one month of life following successful expulsion, naturally or surgically from the mother. These deaths occur prior to the first official routine visit of the child for post-natal care. In general, neonatal mortality comprises the majority of infant mortality of any population. Quite logically, one would expect such rates to be high in areas where maternal nutrition is grossly inadequate, passive immunity derivable from breast feeding is inadequate, and the home environment to which the child returns following complete delivery is unhealthful, and follow-up care for infants with post-delivery complication ineffective.

- **Early Neonatal Mortality Rate:** No. of deaths aged under 7 days divided by the number of live births

$$= \left[\frac{\text{deaths under 7 days}}{\text{Live births}}\right] \times 1000$$

- **Late Neonatal Mortality Rate:**

This is defined as the number of deaths between 7 and under 28 days divided by the number of live births. Early neonatal death (ENND), defined as the death of a newborn between zero and seven days after birth, represents 73% of all postnatal deaths worldwide. Despite a 50% reduction in childhood mortality, reduction of ENND has significantly lagged behind other Millennium Developmental Goal achievements and is a growing contributor to overall mortality in children aged <5 years. The etiology of ENND is closely related to the level of a country's industrialisation.
Hence, prematurity and congenital anomalies are the leading causes in high-income countries. Furthermore, sudden unexpected early neonatal deaths (SUEND) and collapse have only recently been identified as relevant and often preventable causes of death.

$$\left[= \frac{\text{Deaths at 7-27 days}}{\text{Live Births}} \times 1000 \right]$$

- **Post-Neonatal Mortality Rate:**

This is the number of resident newborns dying between 28 and 364 days of age in a specified geographic area (country, state, county, etc.) divided by the number of resident live births for the same geographic area (for a specified time period, usually a calendar year) and multiplied by 1,000. The post-neonatal mortality rate is usually calculated using the annual number of resident infants who died between 28 and 364 days of age in

the numerator and the total annual number of resident live births during the same year in the denominator.

$$\left[= \frac{\text{Infant death bet 28 days and under 1year}}{\text{Live Births}} \times 1000 \right]$$

However, by matching post-neonatal death certificates to the corresponding birth certificates, much more additional and valuable data are obtained (birth weight, smoking status of mother, when prenatal care began, etc.) for post-neonatal mortality risk analysis.

In less densely populated areas, annual numbers of post-neonatal deaths may be small (<10 or 20) which would result in a post-neonatal mortality rate considered to be too unstable or unreliable for analysis. Adding additional years (three or five-year average annual rates) and/or

expanding the area to be studied should result in a larger number of deaths and more reliable rates for analysis.

.

- **Perinatal Mortality Rate:**

This is Stillbirths plus deaths aged less than 7 days per 1000 total births (live + still). The number of perinatal deaths per 1000 total births. A perinatal death is a fetal death (stillbirth) or an early neonatal death. (No. of perinatal deaths / total # of births (still births + live births)) x 1000. A stillbirth is the death of a fetus weighing 500g or more, or of 22-weeks' gestation or more if weight is unavailable (ICD 10). An early neonatal death (END) is the death of a live newborn in the first 7 days (i.e., 0-6 days) of life.

$$= \left(\frac{\text{No. of still births + deaths aged under 7 days}}{\text{Total births (live + still)}} \right) \times 1000$$

Great variation exists both between and within countries on how the stillbirth component of perinatal mortality is recorded, particularly for early stillbirths that occur at 22- to 27-weeks' gestation. For international comparisons, WHO suggests including only deaths of fetuses weighing at least 1000g, or of 28-weeks' gestation or more if weight is unavailable. Presentations of the PMR should include a clear statement of the definition of perinatal mortality used. In practice, in most developing countries accurate data on birth weight or gestational age are difficult to obtain. In addition, Number of perinatal deaths in a given population in a given reference period (i.e., 12 months) and number of births (live births + stillbirths) in the same population and reference period.

The data sources of PMR are Population-based surveys; vital registration; service statistics. Routine HIS may collect data for this indicator to obtain estimates of the PMR for facilities. Facility data are not recommended for estimating the PMR for the general population because in many settings, many perinatal deaths and live births occur outside the health system, which will cause substantial selection bias.

The PMR is a key outcome indicator for new-born care and directly reflects prenatal, intrapartum, and new-born care. It has also been proposed as a proxy measure of maternal health status and mortality, but a recent study has cast doubt on its use as a proxy for maternal mortality (Akalin et al., 1997).

Because the PMR includes both fetal deaths and deaths in the first week of life, it avoids conflicting judgments as to whether a fetus exhibited

signs of life and variations in administrative practice regarding whether or not a death should be counted. In many countries, however, vital registration data are not sufficiently complete to allow reliable estimation of the PMR. Techniques now exist for collecting data on stillbirths, live births, and early neonatal deaths in population-based surveys (pregnancy histories) and applied in surveys including the DHS. However, there has been relatively less experience with pregnancy histories than with birth histories because of concerns about the quality of retrospectively reported pregnancy histories. Common problems with data quality include:

- Omission of stillbirths and early neonatal deaths;
- Difficulty in obtaining accurate information on gestational age or birth weight leading to the misclassification of stillbirths as late spontaneous abortions; and
- Heaping of the reported age at death of live births on 7 days, leading to the misclassification of early neonatal deaths as late neonatal deaths.

3.4 Maternal Death

'Death of a woman while pregnant or within 42 days of the termination of the pregnancy irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management, but not from accidental or incidental causes'. *WHO ICD, 10th Revision (1992).*

'Maternal Death' thus includes deaths from abortion, spontaneous or induced, or from an ectopic pregnancy while it excludes deaths in pregnancy or the post-partum period caused by violence or accidents. Maternal deaths Causes related to or aggravated by pregnancy or its management which are either direct obstetric or indirect causes. The Direct Obstetric Deaths are those related to complications of pregnancy, labor or in the 42-day post-partum period (the puerperium) from interventions or from incorrect treatment or omissions in treatment while Indirect Obstetric deaths are those resulting from a pre-existing disease or one that developed during pregnancy and that is aggravated by pregnancy. Before 1975, deaths from indirect causes were not classified as maternal deaths.

Moreover, maternal deaths have about 80 percent direct major causes. These include: Haemorrhage (25%), Sepsis (15%), Unsafe abortion (13%), Eclampsia (8%) and Obstructed labor (7%) while 20 percent are attributed to indirect causes which are Anaemia, Malaria, Cardiovascular diseases, Hepatitis and Diabetes.

- **Measures of Maternal Mortality**

Maternal Mortality can be measured by the following indices: Maternal Mortality Ratio, Maternal Mortality Rate and Life-Time Risk of Maternal Death.

- **Maternal Mortality Ratio**

This is the number of maternal deaths per 100,000 live births. This is the Information on all maternal deaths occurring in a period (usually 1 year) and information on the total number of live births occurring in the same year. A maternal death (as cited in International Classification of Disease or ICD-10, [WHO, 1992]) is the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and the site of the pregnancy, and can stem from any cause related to or aggravated by the pregnancy or its management, but not from accidental or incidental causes. This indicator also measures obstetric risk and risk faced by women when they are pregnant.

$$= \left[\frac{\text{No of maternal deaths in a year x } 100{,}000}{\text{No. of live births}} \right]$$

- **Maternal Mortality Rate**

The maternal mortality rate refers to the risk of mothers dying from causes associated with delivering babies, complication of pregnancies or childbirth. This important statistic is often neglected because it is difficult to calculate accurately. Number of maternal deaths in a year per 100,000 women of reproductive age in that year.

$$= \left[\frac{\text{No. of Maternal Deaths in 1 year}}{\text{No. of women aged 15-49}} \text{ X } 100{,}000 \right]$$

This measures both the obstetric risk and the frequency with which women are exposed to that risk through pregnancy

- **The Lifetime risk of Maternal Death**

The lifetime risk of maternal death is the probability that a 15-year-old girl will die from complications of pregnancy or childbirth over her lifetime; it takes into account both the maternal mortality ratio and the total fertility rate (average number of births per woman during her reproductive years under current. The total life time risk of death from maternal causes as a woman moves through the reproductive ages.

- Estimated from the formula
- $1-(1-MMR)^{1.2\ TFR}$
- Where
  - MMR (Mat. Mortality ratio is expressed as a decimal)
  - TFR (Total Fertility Rate) is adjusted by 1.2 to allow for pregnancies not ending in life births.

4.0 CONCLUSION

Methods for collecting fertility and mortality data and the various issues that arise in using those methods evolve over time. Demographic research leads to the development of new methods. Application of new and existing methods in countries throughout the world adds to knowledge of how different methods work in different contexts. Technological developments create new opportunities for application. Changing economic, political and social conditions in each country change the environment within which data collecting occurs and the needs, interests and sophistication of users.

5.0 SUMMARY

In summary, fertility and mortality in a population depend on socio-cultural history, sanitary conditions, and biological factors. Although the last element is generally similar in all human communities, socio-cultural and sanitary factors vary considerably.

5.0 TUTOR MARKED ASSIGNMENT (TMA)

1. Briefly describe the following fertility measures:
   a) Crude Birth Rate (CBR)
   b) General Fertility Rate (GFR)
   c) Age-Specific Fertility Rate (ASFR)
   d) Total Fertility Rate (TFR)
   e) Gross Reproduction Rate (GRR)
   f) Net Reproduction Rate (NRR)

2. Briefly describe the following mortality measures:
   a) Maternal Mortality Rate
   b) Maternal Mortality Ratio
   c) Neonatal Mortality Rate
   d) Infant Mortality Rate
   e) The Lifetime risk of Maternal Death

6.0 REFERENCE/FURTHER READING

Chika Onwasigwe, (2004). Principles and Methods of Epidemiology. Institute for Development Studies, University of Nigeria, Enugu.

Akinsola O.J., (2019). Lecture Notes on Basic Demography Methods. College of Medicine, University of Lagos, Lagos, Nigeria.

Rothman K.J., Greenland S. & Lash T. (2011). Modern Epidemiology, (3rd ed.).

An Introduction to Applied Epidemiology and Biostatistics, (1992). Second Edition. Center for Disease Control.

Nwogu E.C. & I.S. Iwueze (2009). Textbook on Introduction to Demography

UNIT 3 STANDARDISATION OF RATES

CONTENTS

1.0 INTRODUCTION

Epidemiologists are always mindful of population diversity. Virtually every large population is heterogeneous in regard to socio-demographic (e.g., age, gender, education, religion), geographic, genetic, occupational, dietary, medical history, and innumerable other personal attributes and environmental factors related to health. A population can be viewed as a composite of diverse subgroups (ultimately, subgroups of size one, i.e., individuals, but epidemiologic measures break down at that point). Any overall measure or statistic reflects the value of that measure for each of the subgroups comprising the population.

An overall measure that does not take explicit account of the composition of the population is called crude. Its value will be an average of the values for the individual subgroups, weighted by their relative sizes. The larger the subgroup, the more influence it will have on the crude measure (i.e., "democracy"). Thus, the death rate for a population is a weighted average of the death rates for its component subgroups. Suppose we consider a population of size N as consisting of five age groups, or strata. Each age stratum will have a specific number of people, say ni (i=1 to 5). During the following year, each stratum will experience some number of deaths, say di. The total population size, N, is therefore Σni, the total number of deaths, D, is Σdi, and the crude mortality rate is D/N, which can also be written as a weighted average of the stratum-specific mortality rates, di/ni, as follows:

$D/N = \Sigma di/N = \Sigma\, ni\, (di/ni)/N = \Sigma(ni/N)(di/ni) = \Sigma wi(di/ni)$
where wi are the weights (note that $\Sigma wi = \Sigma(ni/N) = (\Sigma ni)/N = \Sigma ni/\Sigma ni = 1$).

The terms "adjustment" and "standardisation" both refer to procedures for facilitating the comparison of summary measures across groups. Such comparisons are often complicated by differences between the groups in factors that influence the measures of interest but which are not the focus of attention. Adjustment attempts to remove the effects of such "extraneous" factors that might prevent a "fair" comparison.

"Adjustment", the more general term, encompasses both standardisation and other procedures for removing the effects of factors that distort or confound a comparison. Standardisation refers to methods of adjustment based on weighted averages in which the weights are chosen to provide an "appropriate" basis for the comparison (i.e., a "standard"), generally the number of persons in various strata of one of the populations in the comparison, an aggregate of these populations, or some external relevant population. Other kinds of adjustment, some of which also employ weighted averages, will be discussed in the chapter on Confounding.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- define the term standardisation of Rates
- distinguish between Direct and Indirect Standardisation of Rates

3.0 MAIN CONTENT

3.1 Types of Rates

- **Crude Rates:**

These are rates based on the total population as the denominator) e.g. Crude mortality rate, Crude birth rate, and so on. They are computed for an entire (Total) population. Although the numerator may be specific to a defined group, the denominator is based on all persons in the population regardless of exposure or susceptibility status. Crude rates use total mid-interval population estimate to represent the average size of the referent population in the specified time period. They are indifferent to differences that usually exist by age, sex, race, or some category of pre-existing disease conditions (confounding factors/variable). Crude rates from different populations cannot be meaningfully compared. Adjustment for structural differences is necessary.

- **Specific Rates:**

These are rate in which the denominator is a specific sub-group of the population. For example, Infant Mortality Rate (IMR) and Under-Five

Mortality Rate (u5MR). These rates are based on group specific denominators, that is, the denominator is a specific sub-group of the population. Examples are Infant Mortality Rate (IMR) and Under-five Mortality Rate (u5MR). They usually consider the differences among subgroups and are computed by age, sex, race, and other confounding variables. These variables tend to confound risk patterns in a good many diseases and health problems by being closely associated with exposure and/or susceptibility. Examples are Age-specific death rate; cause-specific morbidity rate; age-sex-specific death rate. Fertility rate is an age-specific rate to the extent that the denominator comprises the total number of women of childbearing age in the population.

- **Standardised Rates:**

These are rates computed to adjust (control) for possible confounding factors. This is also called adjusted rates are not real rates. They are fictional rates used to make valid summary comparisons between two or more groups possessing dissimilar age or other structural distributions and exposure characteristics include age, sex, race, income, smoking status, diet, and indeed exposure to various risk factors of diseases. Adjusted rates are standardised summary figures for a defined population by which statistical procedures are carried out to remove the effect of differences in composition of various populations; thus permitting unbiased comparison. In demography, adjustments of rates are carried out to control or neutralise the influence of socio-demographic, socio-economic, and other exposure or susceptibility variables that are known to be strongly associated with the risk of diseases or other health status outcome. Because such rates are fictional, their absolute magnitude depends on standard population chosen. Measures of the inter-population comparisons are usually the standardised mortality or morbidity ratios.

- **Example**

Crude death rate may be affected by each of the following aspects of population composition such as age, urban/rural residence, different occupational composition, different income composition, sex, and marital status. Quite often we compare morbidity and mortality experiences of different populations, or of the same population at different points in time. So, crude rates can be used to measure improvements in health services, Effect of some exposures, nuclear plant in the community, toxic waste dump, Now, which rates should we use? Crude Rates, Specific rates?

Standardised or Adjusted Rates (for total population) OR Specific rates for specific populations (e.g. IMR, u5MR).

Why?

- Comparisons of crude rates between populations may be very misleading if the distribution of key variables in the populations being compared are different.
  (Consider population pyramids of countries).
- Comparison of age-specific rates is appropriate and useful, but quite often there may be a large number of specific groups and such comparison becomes cumbersome.

Population Pyramid: Norway

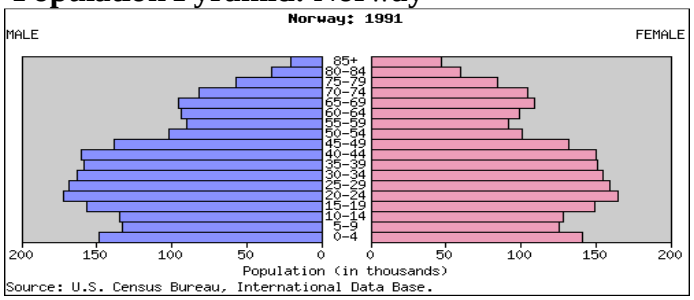


Figure 1.4

Graphical Classification of Population Pyramid (Norway). ***An Introduction to Applied Epidemiology and Biostatistics, Second Edition. Center for Disease Control, USA (1992).***

Population Pyramid: Nigeria

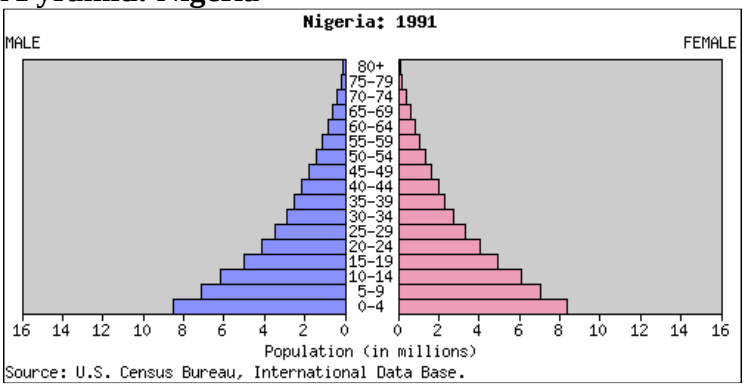


Figure 1.5

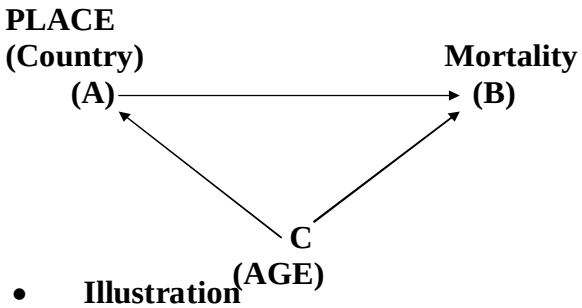
Graphical Classification of Population Pyramid (Nigeria). ***An Introduction to Applied Epidemiology and Biostatistics, Second Edition. Center for Disease Control, USA (1992).***

- Frequently adjustment or standardisation is made for age and sex
- Age adjustment seems to be most commonly undertaken

- **Age as a confounder in mortality**

Age can be a confounder because age is known to be associated with PLACE (i.e. age distribution of places does vary; older people in one

population than the other. Age is also a known risk factor for mortality. Older people are more likely to die than younger people.



- **Illustration**

Two areas A & B have the following mortality data				
	(A)		**(B)**	
Age group	**Pop**	**Deaths**	**Pop**	**Deaths**
0-44	1000	25	4000	120
≥45	4000	160	1000	45
Total	**5000**	**185**	**5000**	**165**
cdr/1000	185 x 1000 / 5000		165 x 1000 / 5000	
	= 37		**= 33**	

- Question: Which area has the higher mortality experience?
- Apparently, Area A, going by the crude death rate: **Or is it?**

If we look again but now at the ASDR in each age group						
	(A)			**(B)**		
Age group	**Pop**	**Deaths**	**ASDR**	**Pop**	**Deaths**	**ASDR**
0-44	1000	25	25	4000	120	30
≥45	4000	160	40	1000	45	45
Total	**5000**	**185**	**(37)**	**5000**	**165**	**(33)**
ASDR = 37				**ASDR = 33**		

- Mortality experience in each age group is lower in A than in B!!
- In order to remove the effect of confounders in our comparison, we undertake standardisation.

- **Principle of Standardisation**

Involves the choice of a standard or reference population the characteristics of this reference population is substituted for each of the two populations being compared:

Thus assigning to them an identical distribution for the variable being standardised.

- **Selection of a Standard/Reference Population**

Standardised rates (indices) are usually influenced by the particular standard used. Distortions can therefore occur if the standard population is quite unlike the observed population. For this reason, standardised rates are fictitious in that they do not refer to any real population. As a general rule select as a standard an age distribution that is similar to the age distribution of the various populations being compared:

- If the two populations are of the same size take the average of the two populations.
- If the sizes are different, then take the average of the relative distribution.

- **Methods of Standardisation**

There are two methods of standardisation:

- Direct
- Indirect

3.2 DIRECT STANDARDISATION

The age structure of the standard population is selected.

- The ASDR of each of the observed population is applied to the standard age structure to yield expected deaths in the standard population,
- The crude death rate in the standard population at the ASDR of each of the observed population can thus be calculated.
  The calculated crude death rate in the standard population is now called the Standardised Death Rate, SDR.

- **Direct Age Adjustment**

Specific rates in the populations of interest are applied to a standard population

- Information on age distribution of populations of interest
- Information on age distribution of events in the populations of interest
- Information on age distribution of the standard population

- **Procedure for the Direct Method**
- Applying the asdr of population A to the Std population, obtain SDR_A
- Also applying the asdr of population B to the Std population, obtain SDR_B
- SDR_A and SDR_B can now be easily compared.
- **Direct Method of standardisation**

	Population A				Population B		
Age Group	No. of Persons	No. of Deaths	Age - specific Death Rate/1000		No. of Persons	No. of Deaths	Age - specific Death Rate/1000
<15	15000	30	2		20000	40	2
15-44	20000	120	6		25000	250	10
45+	15000	300	20		5000	100	20
Total	50,000	450			50000	390	
CMR =450/50000 = 9 per 1000					390/50000 =7.8 per 1000		

	(A)				**(B)**	
Age group	**No. of Persons**	**ASDR/1000 (Pop A)**	**Expected Death (A)**		**ASDR (B)**	**Expected Death (B)**
<15	35000	2	70		2	70
15-44	45000	6	270		10	450
≥45	20000	20	400		20	400
Total	**100,000**		**740**		**165**	**920**
SDR_A = 740/100,000 = 7.4 per 1000					**SDR_B = 920/100,000 = 9.2 per 1000**	

- **Why does Pop B now have higher standardised rate?**

ASDR (A)	EXPECTED	ASDR (B)	EXPECTED
2	70	2	70
6	270	10	450
20	400	20	400

- **Points to note**
- If specific rates in populations of interest are similar, the adjusted rates will remain similar.
- Populations with higher specific rates eventually have higher standardised rates.
- The use of different standard populations does not change the direction of the difference between the populations of interest.

Population A					Population B		
Age Grou	No. of Perso	No. of Deaths	specific Death Rate/1000		No. of Perso	No. of Deaths	specific Death Rate/100
15-19	1600	43	27		1000	5	5
20-24	3000	6	2		1500	3	2
25-29	6000	18	3		2500	90	36
30-39	800	10	13		5000	15	3
40+	800	24	30		10000	10	1
Total	12,200	101			20,000	123	
CMR = Total Deaths * 1000 / Total Population					CMR = Total Deaths * 1000 / Total Population		
101/12,200= 8.3 per 1000					CMR=123/20,000= 6.15		

Age Group	No. of Persons	ASDR (A)	Expected Deaths (A)		No. of Persons	ASDR (B)	Expected Deaths (B)
15-19	5000	27	135		5000	5	25
20-24	14000	2	28		14000	2	28
25-29	25000	3	75		25000	36	900
30-39	4000	13	52		4000	3	12
40+	2000	30	60		2000	1	2
Total	50000		350		50000		967
	SDR (A) = 350/50,000= 7.0 per 1000				SDR (B) = 967/50,000= 9.7 per 1000		

3.3 Indirect Standarisation

i. A standard set of ASDR is selected. The crude death rate of this standard population must also be known.
ii. The actual total deaths in each area must be available as well as the age distribution of the area.
iii. The method boils down to adjusting the crude death rate of the Standard population by a factor.
iv. The factor is the ratio of the recorded number of deaths to the expected number of deaths.

- **The Indirect procedure is as follows:**

i. Multiplication of the appropriate population by the set of standard rates gives the expected number of events.
ii. Summation of these over all age groups gives the total expected number of events
iii. Dividing the total observed no. by the expected no. of events gives the Standardised Mortality Ratio SMR_1 for population 1, ditto SMR_2 for population 2
iv. The ratio multiplied by the CDR of the Std population gives the SDR: - the Standardised Death Rate.

Besides,

- Age distribution of Standard Population
- Age distribution of events in the Standard Population*
- Age distribution of the population of interest

* ***Not necessary in direct adjustment***

Example					
				Expected Deaths	
Age group	**Std Rate Per 1000**	**Pop 1 ('000)**	**Pop 2 ('000)**	**Pop 1**	**Pop 2**
(1)	(2)	(3)	(4)	5= 2x3	6= 2x4
0-4	7.5	560	440	4200	3300
5-24	2.25	1800	1620	4050	3645
25-44	3.15	1820	1800	5733	5670
45-64	17.50	1440	1560	25200	27300
≥65	117.50	620	700	72850	82250
cdr = 18.55					
Actual Deaths		116360	113170	112033	122165
Total Deaths					
SMR= Actual / Expected		116360 / 112033 =1.038	113170 / 122165 0.926		
SDR_1 = 1.038 x 18.55 = 19.3			**SDR_2 = 0.926 x 18.55 = 17.2**		

Standard Population					Pop. (A)		
Age Group	No. of Persons	No. of Deaths	ASDR		No. of Persons	Deaths	ED (A)
<15	13000	35	2.7		400		1.08
15-44	8000	100	13		500		6.25
45+	2000	150	75		600		45
Total	23,000	285			1500	195	52.33
CMR=285/23,000 =12.4 per 1000						CDR= 195/1500 =130 per 1000	
						SMR=O/E =195 /52.33 =3.7	

3.3.1 Interpretation of Standardised Mortality Ratio

When SMR is:

- 1: The risk of death in the population of interest is the same as that of the standard population.
- >1: The risk of death in the population of interest is greater than that of the standard population.
- < 1: The risk of death in the population of interest is lower than that of the standard population.

4.0 CONCLUSION

One of the uses of crude rates is comparison of the levels of an event in two or more populations. However, such comparisons are hampered because the true difference two crude rates may be obscured by population composition due to the factors on bases of which data was classified. Therefore, to compare the levels of an event in two or more populations undistributed by compositional factor, the contribution to the crude rates due to compositional factor need to be eliminated or controlled. This is achieved by a procedure called standardisation.

5.0 SUMMARY

Crude rates should not be used as a basis for comparison especially in situations where the populations have different age structure (or any other variable related to the risk of the outcome of interest). Standardisation is a procedure for controlling the effect(s) of compositional factors on crude rates. The factors often controlled for in demographic analysis include age, sex, race, place of residence, level of education, wealth quintile etc.

6.0 TUTOR MARKED ASSIGNMENT ((TMA)

1. (a) Compare and contrast direct and indirect methods of standardisation of rates

   (b) The table below shows population numbers and deaths by age and sex in a segment of a population

Age group	Population (000)		Number of deaths	
	Males	Females	Males	Females
1–4	1422	1380	1637	1325
5–14	3062	2968	1390	920
15–24	2430	2318	2816	1437
25–44	4101	4023	9690	5942
45–64	2755	2753	36581	18535

From this segment of the population:

Calculate

i. The crude death rates for males, females and the total population
ii. The age-specific death rates for males, females, and the total population
iii. an index of excess male mortality for each age group

7.0 REFERENCES/FURTHER READING

Chika Onwasigwe, (2004). Principles and Methods of Epidemiology. Institute for Development Studies, University of Nigeria, Enugu.

Akinsola O.J. (2019). Lecture Notes on Basic Demography Methods. College of Medicine, University of Lagos, Lagos, Nigeria.

An Introduction to Applied Epidemiology and Biostatistics, (1992). (2nd ed.). Centre for Disease Control.

Nwogu EC & IS Iwueze (2009). Textbook on Introduction to Demography

Akinsola OJ, (2019). Lecture Notes on Basic Demography Methods. College of Medicine, University of Lagos, Lagos, Nigeria.