MODULE: COMMUNITY HEALTH
UNIT: EPIDEMIOLOGY

By Carey Francis Okinda
June-July 2015

OUTLINE

	Topic	Hours
1.	Introduction to Epidemiology	1
2.	Disease – Determinants and Prevention	2
3.	Measurement of Disease Occurrence	2
4.	Descriptive Epidemiology	1
5.	Epidemiological Study Designs	2
6.	Sampling Methods in Epidemiology	2
7.	Population Screening	2
8.	Disease Surveillance	2
	TOTAL	**15**

Lesson 1: INTRODUCTION TO EPIDEMIOLOGY

Objectives

At the end of the lesson the learner will be able to: -
1) Define key words
2) Explain epidemiological concepts and terms
3) Compare the clinical medicine and epidemiology

1.0 INTRODUCTION

- The word derived from 'epidemics' and from Greek words "*epi*"= on or upon; "*demos*" = the common people; "*logos*"= study
- Epidemiology is the study of:
  i) Something that afflicts the population
  ii) Distribution of disease and its determinants in human populations in order to control health problems
  iii) Disease occurrence (frequency, distribution, determinants) in human populations and application of that knowledge to control health problems
  iv) Distribution and determinants of health related states or events in specified human populations and the application of this study to control of health problems.
- Epidemiologists are interested in occurrence of disease with respect to time, place and persons
- The key aspects of the definitions include
  i) Study
     - Involves surveillance, observation, hypothesis testing, analytical research and experiments

- Levels of scientific study of disease involve the
  a) Sub molecular or molecular level (e.g., cell biology, biochemistry, and immunology);
  b) Tissue or organ level (e.g., anatomic pathology)
  c) Level of individual patients (e.g., clinical medicine); and
  d) Level of populations (e.g., epidemiology).

ii) Determinants**(PBSCB)**
- Factors or events that are capable of bringing about change in health which include physical, biological, social, cultural and behavioural factors that influence health

iii) Distribution
- This refers to the analysis of disease by **person, place** and **time**.
- The frequency of disease may vary from one population group to another (Give examples)

iv) Population
- A group of people with a common characteristic, e.g. place of residence, gender, age, or occurrence of event
- To measure frequency of disease in population consider:
  - Number of people affected (cases)
  - Size of the population from which the cases arise
  - Amount of time that the population is followed
- Two major types of populations
  a) Fixed or closed population:
     - A population in which the same individuals are followed from the start of the study (follow up period) until its end
     - Membership is defined by a life event and is permanent
     - Useful for computing cumulative incidence
  b) Dynamic or open population:
     - A population in and out of which individuals move during the follow up period (e.g. the population of a town)
     - Membership based on a changeable condition and is transitory
     - Useful for computing incidence rate or density

v) Health phenomena (morbidity and mortality)
- Epidemiology investigates many different kinds of health outcomes, from infectious diseases to chronic diseases, and various states of health such as disability, injury, limitation of activity, and mortality
- Other health outcomes have included positive functioning of the individual and active life expectancy as well as health-related events including mental illness, suicide, drug addiction and injury

vi) Control of the health problem
- Reduce the risk factors
- Reduce or stop transmission of infectious agents

2.0 SCOPE OF EPIDEMIOLOGY

- Epidemiology is concerned with efforts to:
  1) Describe (enumerate/ count)
     - To describe the health status of a population means to enumerate the cases of disease, to obtain relative frequencies of the disease within subgroups, and to discover important trends in the occurrence of disease.

  2) Explain (etiology
     - To explain the etiology of disease means to discover causal factors as well as to discover modes of transmission.

  3) Predict (risks)
     - To predict the occurrence of disease is to estimate the actual number of cases that will develop as well as to identify the distribution within populations. Such information is crucial to planning interventions and allocation of resources.

  4) Control
     - To control the distribution of ds, the epidemiologic approach is used
       - to prevent the occurrence of new cases of diseases,
       - to eradicate existing cases, and
       - to prolong the lives of those with diseases

3.0 HISTORY OF EPIDEMIOLOGY

- Started off as the study of various disease epidemics
- Concerned with modern epidemics, to include infectious (communicable, molecular), chronic (cancer, cardiovascular), nutritional and non-communicable diseases (injury, mental), clinical (diagnostics, therapy, prognosis) and reproductive health issues

4.0 USES OF EPIDEMIOLOGICAL DATA

1. Examine causation
2. Study natural history of the disease
3. Description of the health status of population (the host)
4. Determine the relative importance of causes of illness, disability and death
5. Evaluation of interventions
6. Identify risk factors
7. Changes of the pattern of infectious diseases
8. Discovery of new infections

5.0 IMPORTANCE OF EPIDEMIOLOGY

1) Describes the geographical distribution and size of disease
2) Identify causes of disease
3) Identifies who is at risk of the disease
4) Determines the risk factors for the disease
5) Design programs to prevent, control and treat the disease
6) Monitor the disease to establish its trends over time

6.0 BASIC CONCEPTS IN EPIDEMIOLOGY

1) Description and quantification of disease events in populations
2) Disease determinants – intrinsic and extrinsic determinants
3) Analysis of epidemiological data
4) Presentation of epidemiological data (using descriptive statistics)
5) Investigation of disease problems

7.0 EPIDEMIOLOGICAL TERMS

	Concept	**Description**
1.	Aetiology	Cause(s) of disease
2.	Agent or risk factor	A biological or non-biological substance whose presence or absence is associated with increased probability that a disease will occur
3.	Carrier	An apparently healthy person who harbours an infectious agent for some time period and with the ability to transmit it to those who do not have it
4.	Endemic	Occurrence of a disease at the expected level within a geographical area Is a disease that occurs in a population with predictable regularity and with only minor deviations from its expected frequency of occurrence
		Holo endemic - a **disease** for which a high prevalence of infection begins early in life and affects most or all children, leading to a state of equilibrium, such that the adult population shows evidence of the **disease** much less frequently than in children
		Hyper endemic - a **disease** that is constantly present at a high incidence and/or prevalence rate and affects all age groups equally (is an endemic disease that affects a high proportion of the population at risk.)
		Hypo endemic - is an endemic disease that affects a small proportion of the population at risk
		Mesoendemic is an endemic disease that affects a moderate proportion of the population at risk

5.	Epidemic	Disease occurrence at a level that is clearly higher than the expected within a geographical area
6.	Host	An organism capable of being infected or affected by a disease causing agent
7.	Immunity	Ability to resist disease attack
8.	Incidence	Number of new cases of a disease within a specified period
9.	Incubation period	Interval between exposure to an infectious agent and the disease it causes The period between infection and clinical onset of the disease Clinical incubation period – Biological incubation period -
10.	Isolation	Voluntary or compulsory separation and confinement of those known or suspected to be infected with a contagious disease agent (whether ill or not) to prevent further infections (transmission-based precautions are imposed.)
11.	Latent period	The time from infection to infectiousness
12.	Life expectancy	The average period that a person may expect to live
13.	Pandemic	Infectious **disease** that has spread through human populations across a large region (multiple continents) or even worldwide.
14.	Population at risk	Total number of people who are exposed to a risk or set of risks responsible for a disease condition
15.	Prevalence	Number of old and new cases of a disease at a point in time
16.	Prognosis	Likely outcome of a disease condition
17.	Quarantine	A **period** of time that must elapse before those exposed to or attacked by a contagious disease can be considered as incapable respectively of developing or transmitting the disease.
18.	Sporadic	Is a disease that is normally absent from a population but which can occur in that population, although rarely and without predictable regularity.
19.	Study sample/population	Subset of the total population from whom information can be generated to reflect characteristics of the whole population
20.	Vital statistics	Record of import information about a population

Define the following carriers – active, convalescent, healthy; incubatory and intermittent carriers

8.0 SOURCES OF EPIDEMIOLOGICAL DATA

1) Census
   - It is a process of collecting, compiling and publishing demographic, economic and social data pertaining at a specified time to all persons in a country.

2) Registration of vital events (vital statistics)
   - Registration of vital events keeps continuous check on demographic changes
   - Complete and accurate registration is a reliable source of health data
   - Vital events include legal registration, statistical recording and reporting of the occurrence of and the collection, compilation, presentation, analysis and distribution of statistics pertaining to vital events such as live births, deaths, foetal deaths, marriages, divorces, adoptions, separations,

3) Notification
   - Provides valuable information about fluctuation of disease frequency and early warning about new occurrences or outbreaks of disease
   - Notifiable diseases under the International Health regulation include cholera, plague and yellow fever. A few others such as Lassa fever, SARS, H1N1 influenza , etc. are placed under international surveillance

4) Hospital records
   - Data constitute a basic and primary source of information about disease prevalence
   - Provide information on geographic sources of patients, age and sex distribution of disease and duration of hospital stay, distribution if diagnosis, association between different diseases, period between disease and hospital admission, distribution of patient according different social and biological characteristics and cost of care.
   - Considered A poor guide to the estimation of disease frequency

5) Disease registers
   - Morbidity registers are a valuable source of information as to duration of illness, case fatality and survival
   - Registers allow follow up of patients and provide continuous account of frequency of disease in the absence of population base, useful information may be obtained from registers on natural course of disease.
   - It can provide data on morbidity from the particular diseases, treatment given and disease –specific mortality

6) Record linkage
   - It is used to describe the process of bringing together records relating to one individual, the records originating in different times and places.

7) Epidemiological surveillance
   - Particular diseases are endemic, targeted for control, elimination and eradication.
   - Surveillance systems are set up to report on occurrence of new cases and efforts to control the diseases

8) Active epidemiological surveillance
   - In-depth search for cases of a few selected diseases likely to cause epidemics e.g. sentinel surveillance for EPI diseases

9) Population surveys
   - Surveys for evaluating health status of a population – community diagnosis and for investigation of factors affecting health and disease- environment, occupation
   - helpful for studying the natural history of disease , obtaining new information about disease aetiology and risk factors

10) Computerized bibliographic database
   - DHS – Demographic Health Survey; DOLPHIN- Data online for Population , Health and Nutrition, GIDEON – Global Infectious Disease and Epidemiology Network

Table 1.1: Advantages and disadvantages of different sources of data

	Source	**Advantages**	**Disadvantages**
1.	Statistical Data	• Enables tracking of major trends • Background information on broad social conditions • Can go back to see trends	• Not always accessible • Causes of death not known • Incomplete statistics • Is not fed back to role-players • Not always reliable
2.	Hospital Clinic Records	• Regular, monitored, fits into overall system • Basic information regarding community • Information on individuals, families and communities • Shows health trends • A statutory obligation to keep records	• Too much data (takes too long to collect • Poor feedback & communication • Very seldom analysed. • Single geographical area. • Information from health sector only • Only as good as record-keeper • Only clinic visits are covered
3.	Other Departments Or Organisations	• Gives a comprehensive picture • Gives inside story on community profile - political power, cultural influences • Can provide information on perceived community needs	• Limited data (not easy to access) • Unknown reliability • Subjectivity • Systems not compatible (age, standards) • Inadequate training
4.	Health Information System	• Good source • Linked up to other processes • Can be interpreted locally. • Can empower people	• Limited e.g. TB • Questionable accuracy • Needs intensive training of health worker • Feedback and interpretation problematic • Lack of incentive to analyse data locally

5.	Annual Reports	• Profile of needs and provision of services • More textured than statistics • Gives direction to other sources	• Poor analysis • Biased and not always inclusive • Too broad - summary • Out of date due to lengthy production time • Boring/not user friendly
6.	Internet	• Huge source, instant access to e.g. literature. • Point of comparison e.g. international	• Limited facilities for access. • Developing world less well represented. • Not always up to date
7.	Text Books	• Expert information • Identify broad range of problems • Provides models/formats	• Out of date • Not always applicable • Too much text • Theory does not inform implementation • Costs/availability
8.	Research Projects	• Current information • Time saving (if relevant) • Can be specific to a problem • Pilots/demonstration projects	• Costly and unsustainable. • Not always applicable or relevant. • Takes time (not often up to date) • Author biased • Creates expectations
9.	District Health Profiles	• Readily available • Gives "bird's eye" view • Provides leads to other sources	• Questionable accuracy and reliability • Questionable validity of out-dated profiles • Takes time to keep updated
10.	Journals	• Current information • Can get back numbers (old copies) • Summarised. • Different points of view	• Difficult to access • Can be biased • More academic than practical • High volume
11.	Surveys	• Cost effective • Focused/specific, pick up hidden information • Can be spread over large areas e.g. national • Correctly planned, much more comprehensive	• Once-off information • Influenced by questionnaire • Might not empower people
12.	Census	• Gives national picture • Standardised denominators & baseline data • Useful for planning	• Not available until years later. • Expensive. • Creates expectations. • Unknown reliability

9.0 PRACTICAL APPLICATION OF EPIDEMIOLOGY

1) Evaluation of health services for baseline data, efficacy, effectiveness and efficiency
2) Policy formulation
3) To diagnose the health of the community and the condition of the people
4) To study the history of the health of populations, and of the rise and fall of diseases and changes in their character giving useful projections into the future
5) To study the working of health services with a view to their improvement
6) To estimate from the group experience what are the individual risks on average of disease, accident and defect, and the chances of avoiding them.
7) To identify syndromes by describing the distribution and association of clinical phenomena in the population.
8) To complete the clinical picture of chronic diseases and describe their natural history
9) To search for causes of health and disease by computing the experience of groups defined by their composition, inheritance and experience, their behaviour and environments

10.0 DESCRIPTIVE VARIABLES FOR THE HEALTH OF THE COMMUNITY

1) Demographic and social variables:
   a) Age and sex distribution
   b) Socioeconomic status
   c) Family structure, including marital status and number of single parent families
   d) Racial, ethnic, and religious composition

2) Variables related to community infrastructure:
   a) Availability of social and health services including hospitals and emergency rooms
   b) Quality of housing stock including presence of lead-based paint and asbestos
   c) Social stability (residential mobility)
   d) Community policing
   e) Employment opportunities

3) Health-related outcome variables:
   a) Homicide and suicide rates
   b) Infant mortality rate
   c) Mortality from selected conditions (cause specific)
   d) Scope of chronic and infectious diseases
   e) Alcoholism and substance abuse rates
   f) Teenage pregnancy rates
   g) Occurrence of sexually transmitted diseases
   h) Birth rate

4) Environmental variables:
   a) Air pollution from stationary and mobile sources

b) Access to parks/recreational facilities
c) Availability of clean water
d) Availability of markets that supply healthful groceries
e) Number of liquor stores and fast-food outlets
f) Nutritional quality of foods and beverages vended to schoolchildren

11.0 ETHICAL AND PROFESSIONAL ISSUES

- Include obligations to study subjects, privacy, confidentiality, access to data, race and ethnicity; conflict of interest and interpreting findings

12.0 CLINICAL MEDICINE VS EPIDEMIOLOGY

- Clinical descriptions of malaria: fever, diarrhoea, vomiting, headache.
- Epidemiologic description: *what age groups* would be most likely affected, time *trends*, *geographic trends* and other variations that affect the distribution of malaria?

Classical (Typical) Epidemiology
- Population-oriented,
- Studies the community origins of health problems, particularly those related to: nutrition, the environment, human behaviour and the psychologic, social and spiritual states of a population**.**
- Classical epidemiologists are interested in discovering risk factors that might be altered/ modified in a population to prevent or delay disease or death.

Clinical epidemiology
- Use similar research designs and statistical tools.
- Science of making predictions about individual patients by counting clinical events in similar patients, using strong scientific methods for studies of groups of patients to ensure that predictions are accurate
- Study patients in health care settings in order to improve diagnosis and treatment of various diseases and the prognosis of patients already affected by a disease.
- Primary goal is to improve clinical decisions.
- Some therefore prefer to call it clinical decision analysis

	Clinician	**Epidemiologist**
1.	Concerned with the health of an individual	Concerned with the health of individuals
2.	Looks for appropriate treatment for a disease	Looks for ways of preventing the disease
3.	Uses expensive and complicated equipment	Uses simple and cheap methods and equipment to study disease characteristics
4.	Concerned with a sick person	Concerned with the community

Lesson 2: DISEASE - DETERMINANTS AND PREVENTION

Objectives

At the end of the lesson the learner will be able to: -

1. Discuss determinants of disease and concepts of immunity
2. Describe the natural history of a disease and the epidemiological triad
3. Discuss levels of disease prevention

1.0 DISEASE

What is disease?

2.0 INFECTION

What are the sources and causes of infection?

3.0 PROCESS OF INFECTION

- Six requirements for the successful invasion of the host by an infectious agent include
  1) Condition in the environment must be favourable to the agent or the agent must be able to adapt in the environment
  2) Suitable reservoirs must be present
  3) A susceptible host must be present
  4) A satisfactory portal of entry into the host
  5) Accessible portal of exit from the host
  6) Appropriate means of dissemination and transmission to a new host

4.0 DETERMINANTS OF DISEASE

- Factors or events that are capable of bringing about change in health (PBSCB) i.e physical, biological, social, cultural and behavioural factors that influence health
- Examples
  - Specific biologic agents (e.g., bacteria, viruses, protozoa).
  - Chemical agents that may act as carcinogens
  - Stress or adverse lifestyle patterns (lack of exercise, or tobacco consumption or diet high in saturated fat
- Determinants of disease can be intrinsic or extrinsic

Intrinsic Determinants

What are the intrinsic and extrinsic determinants of disease?

Extrinsic Determinants

1) Physical - mechanical injuries/trauma, thermal, radiation, noise, psychological stress, light, etc.

2) Chemical - organic/inorganic forms, gases
3) Nutritional - metabolic, primary/secondary, nutritional deficiencies, hormonal imbalances, etc.
4) Biological - all living organisms that cause disease or infection - bacteria, virus, rickettsia, protozoa, fungi, etc. Genetic defects are included here

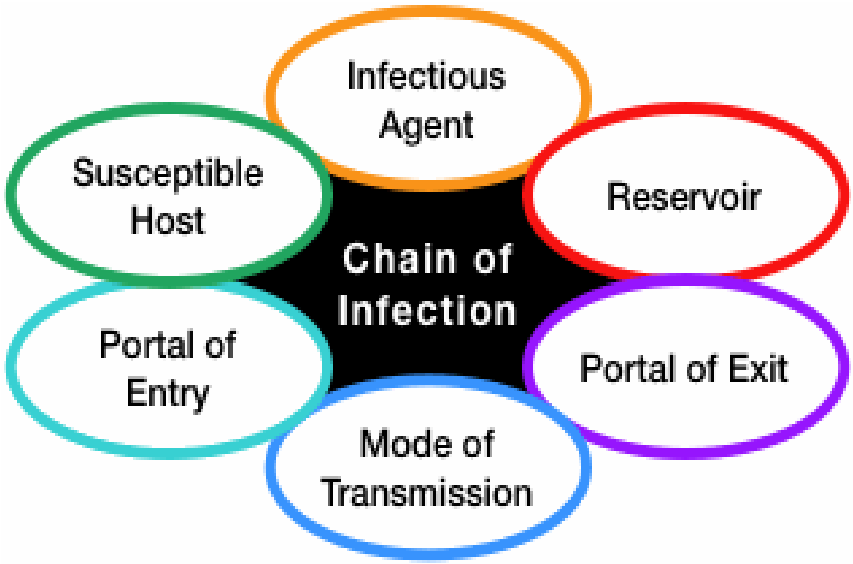
5.0 DYNAMICS OF DISEASE TRANSMISSION

- Disease results from the interaction of the host (a person), the agent and the environment
- Many underlying principles govern transmission of diseases
- Human susceptibility is determined by genetic, nutritional and immunological characteristics

- Discuss the modes of disease transmission

Chain of Transmission

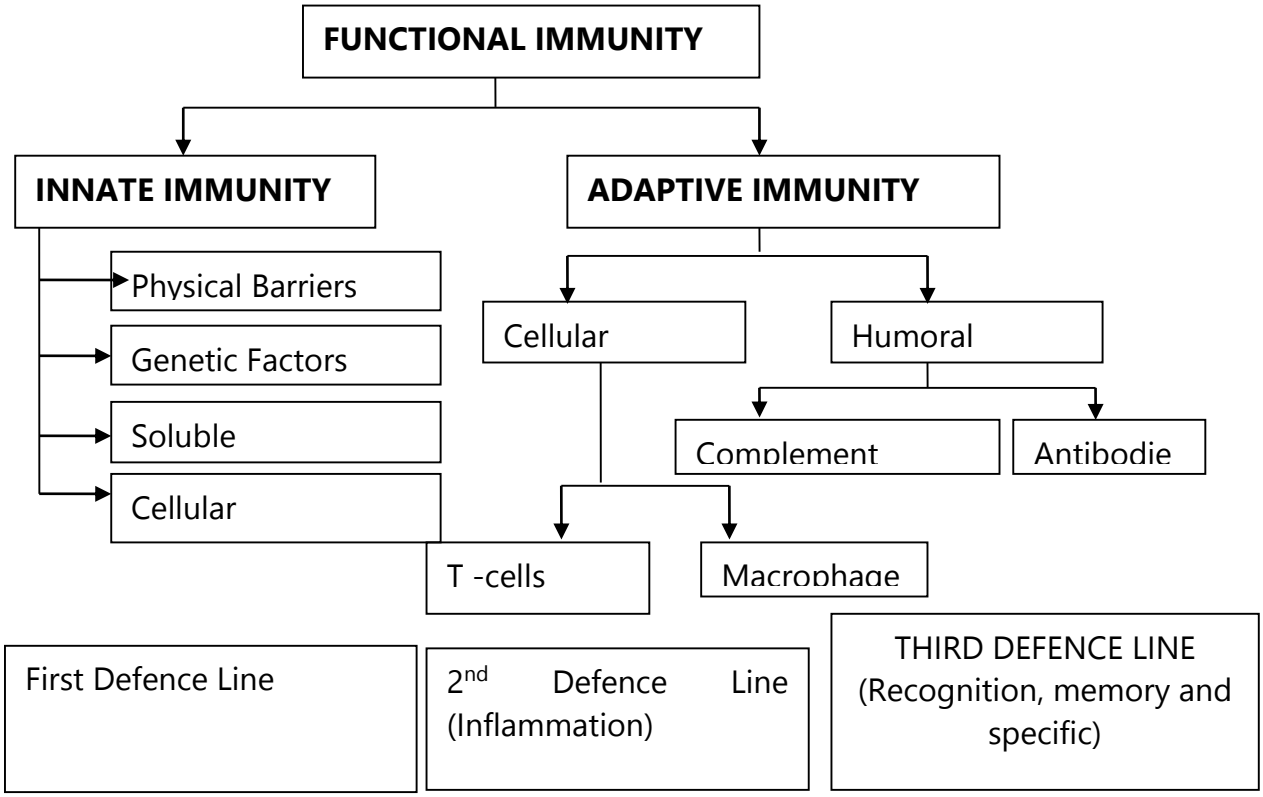
- It is a process that begins when (1) an infectious agent or pathogen (2) leaves its reservoir, source, or host through (3) a portal of exit, (4) is conveyed by some mode of transmission, (5) enters the host through an appropriate portal of entry, and (6) infects a susceptible host
- The now-infected susceptible host becomes a new reservoir and the whole process starts over
- The concept of a chain of infection is essential to our understanding of why we do what we do to prevent infection. If any link of the chain of infection can be broken, the spread of infection can be prevented.



Factors associated with Increased Risk of Human Disease

What are the factors associated with increased risk of human

6.0 CONCEPT OF IMMUNITY



HERD-IMMUNITY

- Is the resistance of a group to an attack by a disease to which the large proportion of the members of the group are immune
- Is the protective effect that arises from a high proportion of the population have had a disease (active natural immunization) or have been immunized (active artificial immunization)
- Occurs when spread of disease in a community reaches a certain proportion of people who are immune in that the likelihood of encountering a susceptible individual to whom the disease can be transmitted is minimal
- Important conditions for establishment of herd immunity include: -
  a) Disease agent must be restricted to a single host species within which transmission occurs
  b) Transmission must be relatively direct from one member of the host species to another
  c) Infections must induce solid immunity
  d) Probability of an infected person encountering every other individual in the population remains the same (constant random mixing of the population)

7.0 HOST-PARASITE RELATIONSHIPS

- Many microorganisms may benefit from mankind by developing a range of host-organism inter-relationship including

1) Symbiosis is the cooperative association between two dissimilar organisms beneficial to both
2) Commensalism is an association between two dissimilar organisms living together benefiting one without harming the other
3) Parasitism is relation between two dissimilar organisms living together benefiting the parasite but harming the host
4) Saprophytism is a relation where organisms live on dead tissues.

Factors Determining Host-Microorganisms Relationships

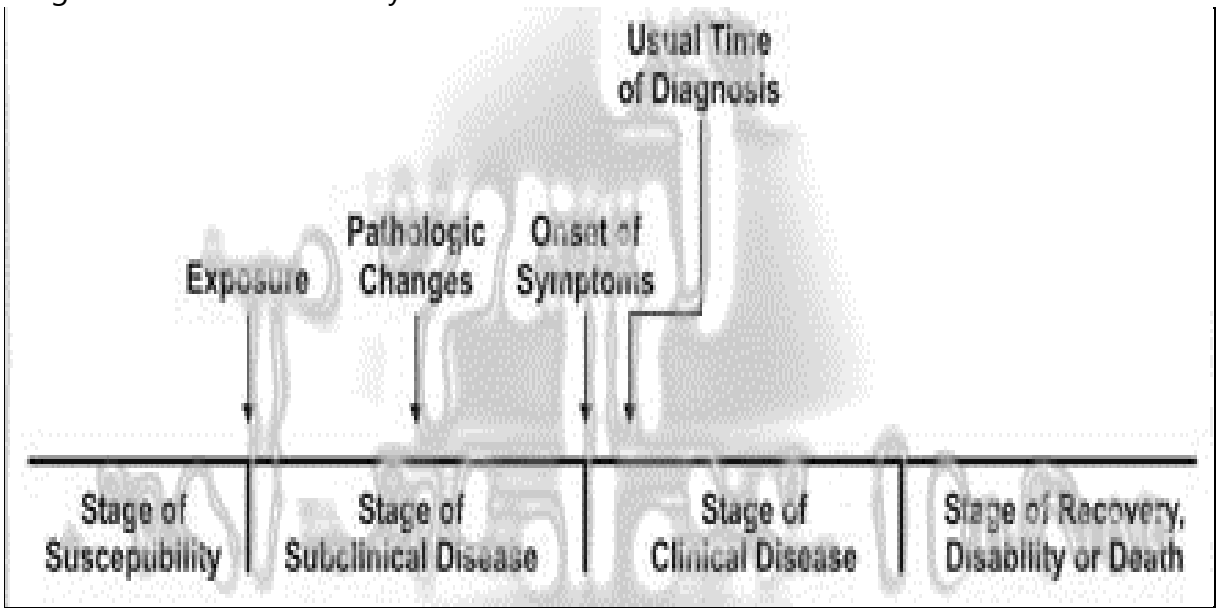
1) Infectious agent factors
   a. Mode of entry - Ingestion (external route), Inoculation (parenteral), Inhalation (respiratory) and Direct contact (contagious infection)
   b. Spread of infection – through phagocytic cells, blood vessels, lymphatics
   c. Production of toxin – exotoxins and endotoxins
   d. Virulence of organisms (adhesions, penetration and colonization, toxins, antimicrobial resistance and host factors)
   e. Products of organisms e.g. enzymes

2) Host factors
   a. Physical barrier e.g. skin
   b. Chemical barrier – mucous secretions, gastric acid
   c. Effective drainage in the respiratory tract, gastrointestinal tract, urinary and genital system
   d. Immune defence mechanisms (immunity) - non-specific (innate) and specific or adaptive immune mechanisms (humoral – antibodies; cellular T and B cells)

8.0 NATURAL HISTORY OF DISEASE

- Every disease in a host follows a potentially predictable life cycle from onset to final outcome (natural history)
- Understanding the natural history of disease is important to clinicians in establishing appropriate treatment and accurate prognosis, and it is vital to public health professionals in developing effective disease prevention and control strategies
- Life cycle or natural history of a particular disease varies from individual to individual, and different diseases but four common stages manifest
  1) Stage of susceptibility
  2) Stage of pre-symptomatic disease
  3) Stage of clinical disease
  4) Stage of diminished capacity/disability/death/chronicity

Diagram 2.1: Natural History of a Disease



The stage of susceptibility

- Precedes the onset of disease (disease has not yet developed)
- Conditions necessary for disease development/occurrence exist (risk factors are present) making the host is susceptible
- Examples
  - Individuals with high serum cholesterol, hypertension, a sedentary lifestyle, and diabetes, for example, have an increased risk of developing coronary heart disease
  - Lack of sleep, excessive stress, and poor eating habits may predispose one to the common cold
  - Susceptibility for kwashiorkor includes inadequate intake of proteins in diet or infectious diseases
- Epidemiologists are continually seeking to identify and confirm risk factors for the major health problems that affect society

Stage of pre-symptomatic disease

- Disease process has begun, but no overt signs or symptoms are evident to the host.
- Characterized by symptoms that tentatively suggest that an individual is about to suffer from a given disease
- For communicable diseases, this stage includes the incubation period, which is the time between the invasion of an infectious agent and the development of the first signs or symptoms of the disease
- For non-communicable diseases the pathological process has begun but the signs and symptoms are not evident
- A carrier of a communicable disease is an individual who has no symptoms of the disease but nevertheless harbours the causative agent

Stage of Clinical Disease

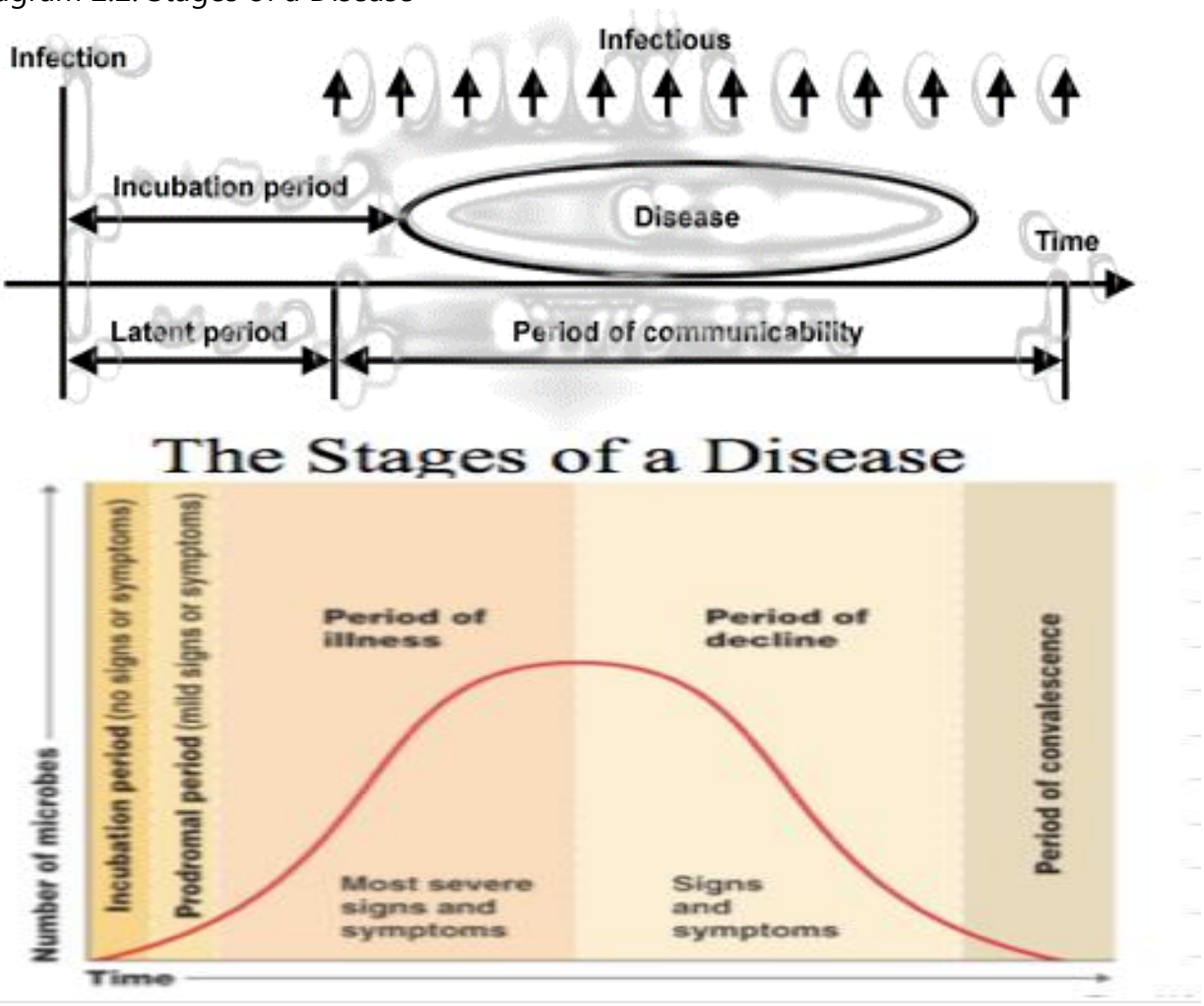
- Disease process is established and the signs and symptoms are evident following the pathological alteration in structure and function of the body tissues, organs and systems e.g. elevated blood pressure headache, fatigue in an individual with hypertension.
- Most people seek medical treatment

Stage of Diminished Capacity/recovery/death/chronicity

- Revolves around effects on permanent damage caused by the disease process e.g. disabilities in hypertension, diabetes, syphilis
- If disease is not treated it may cause disability, become chronic or lead to death

ASSIGNMENT
Discuss the natural history of the following diseases – tuberculosis; Diabetes mellitus; Hypertension; HIV; Syphilis; Pneumonia; Amoebiasis; Tetanus; Poliomyelitis; Rheumatic fever

Diagram 2.2: Stages of a Disease



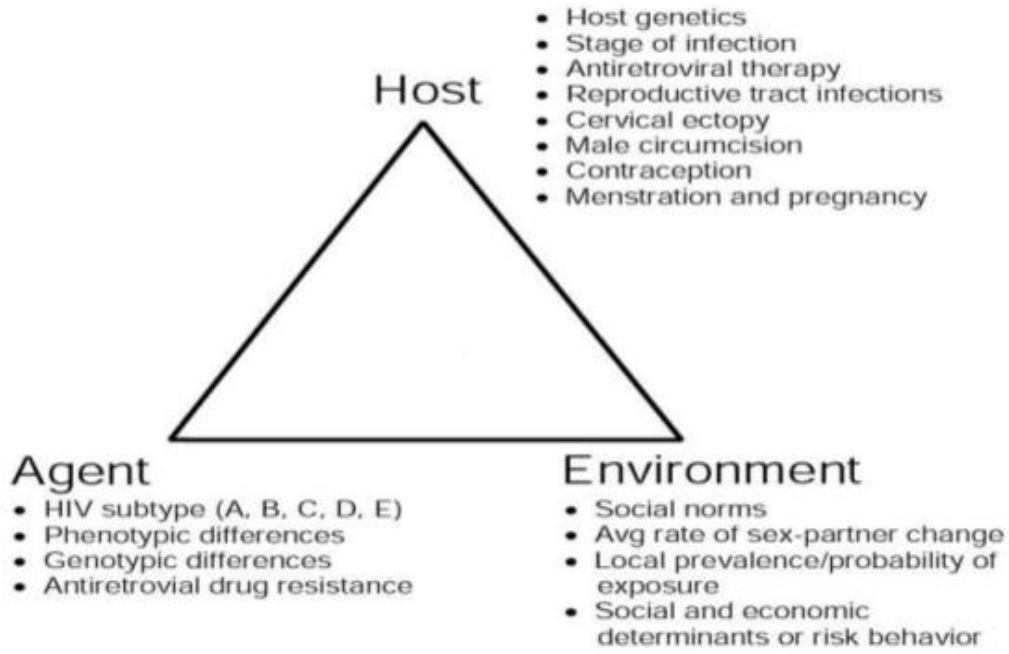
9.0 EPIDEMIOLOGICAL TRIAD (HOST-AGENT-ENVIRONMENT MODEL)

- Amount of disease in a population results from interaction of the 3 components
  - In a state of equilibrium the host and the agent are in a state of balance in the environment and there is no disease
  - When a change occurs that favours the existence and multiplication of the agent disease frequency increases
  - If a change favours the host such as improved nutrition disease frequency decreases
- The epidemiological triad/triangle comprises of : -
  1. A susceptible host (the person at risk for the disease)
  2. A disease agent (the proximate cause)
  3. An environmental context for the interaction between the host and agent

The Host

- The individual human being whose state is determined by the interaction of genetics and environmental factors
- Disease only occurs in a host who is susceptible (lack of susceptibility is due to immunity or inherent factors[1])
- Examples: age, sex, ethnic groups, nutritional status, socio-economic, personal behaviours (smoking, diet, drinking, exercise, sexual behaviours), personality, immunization status and physical states – pregnancy, puberty, fatigue, immunocompromised

Diagram 2.3: Epidemiologic Triad



The Agent

- This is the factor whose presence or absence causes a disease
- Examples - biological agents, chemical agents, nutritive element and physical agents

[1] Ability to resists diseases due to factors other than antibodies e.g. good nutrition, exercises

Environmental factors

- Physical – weather, climate, geology
- Biological – sources of food, water and air, presence of vectors, flora and fauna
- Socioeconomic and cultural – density, crowding, adequate housing, war, sanitation and availability of health care

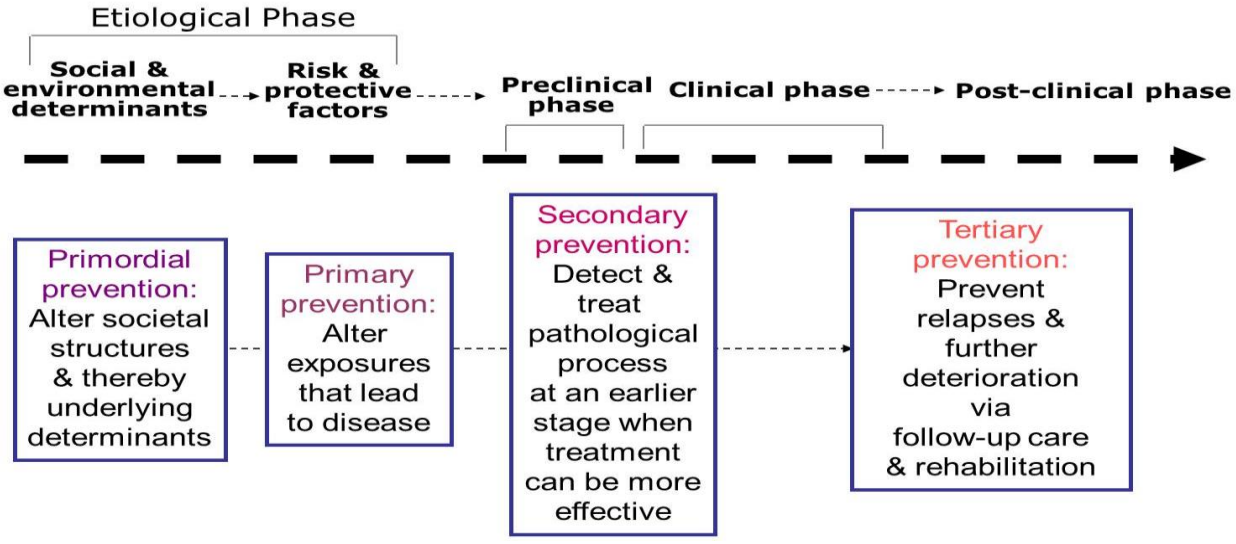
10.0 LEVELS OF DISEASE PREVENTION

- Disease prevention applies to measures taken to prevent diseases before they occur as well as measures taken to prevent disease progression
- There are four levels of disease prevention namely primordial, primary, secondary and tertiary

Primordial Level

- Primordial prevention is defined as prevention of risk factors themselves, beginning with change in social and environmental conditions in which these factors are observed to develop, and continuing for high risk children, adolescents and young adults
- Examples of primordial prevention actions - National policies and programmes on nutrition involving the agricultural sector, the food industry, and the food import-export sector
- Responsibility lies on the government. professional and non-governmental organizations, industry, hospitals, health clinics, health practitioners and health-care workers

Diagram 2.4: Levels of Disease Prevention



Primary Level (Primary Intervention

- Entails measures taken to prevent diseases before they occur
- Strategies emphasize general health promotion, risk factor reduction, and other health protective measures

- Strategies include: -
  - Health education and health promotion programs designed to foster healthier lifestyles and environmental health programs designed to improve environmental quality
  - Specific examples of primary prevention measures include immunization against communicable diseases; public health education about good nutrition, exercise, stress management, and individual responsibility for health; chlorination and filtration of public water supplies; and legislation requiring child restraints in motor vehicles.
  - Examples
    - General health promotion strategies taken at home, working places and in institutions e.g. promotion of good nutrition, provision of basic needs (food, clothing, shelter), recreation facilities
    - Specific measures - immunization, avoidance of substances (e.g. drugs), prevention of accidents

What strategies has the government employed to facilitate primary intervention? Discuss using specific examples

Secondary Level (Curative level)

- Entails early detection of disease followed by prompt intervention
- It is important in preventing complications, disabilities and communicability of disease
- Secondary prevention focuses on early detection and swift treatment of disease
- Its purpose is to cure disease, slow its progression, or reduce its impact on individuals or communities
- A common approach to secondary prevention is screening for disease
- Examples
  - Screening - non-invasive computerized test for the early detection of heart disease, mMammography for breast cancer detection; eye tests for glaucoma; blood tests for lead exposure; occult blood tests for colorectal cancer; the Pap test for cervical concrete breath test for Helicobacter pylori, the bacterium implicated in duodenal and gastric ulcers; and the Prostate-Specific Antigen(PSA) test for prostate cancer
  - Treatment e.g. treatment of hypertension to prevent complications and removal of skin cancer lesions as they occur
- Protects healthy members of the community against certain diseases

What strategies has the government employed to facilitate secondary/curative intervention? Discuss using specific examples

Tertiary Level (Rehabilitative Intervention)

- Consists of various attempts made to improve the quality of life of an individual after a disease has occurred and caused damage to the person
- Involves both therapeutic and rehabilitative measures once disease is firmly established

- Examples include treatment of diabetics to prevent complications; management of chronic heart disease patients with medication, diet, exercise, and periodic examination; improving functioning of stroke patients through rehabilitation by occupational and physical therapy, nursing care, speech therapy, counselling, and so forth, and treating those suffering from complications of diseases such as meningitis, multiple sclerosis, or Parkinson's disease.
- May involve modification of working and home environments and funding affected persons to start IGAs

What strategies has the government employed to facilitate tertiary/rehabilitative intervention? Discuss using specific examples

11.0 CLASSIFICATION OF DISEASES

1) Communicable Diseases

- Communicable diseases are transmissible from one person, or animal, to another
- A communicable disease is defined as an illness that arises from transmission of an **infectious agent** or its toxic product from an infected person, animal or reservoir to a **susceptible host**, either directly or indirectly through an intermediate plant or animal host, vector, or environment.

2) Non-Communicable Diseases

- Non-communicable diseases (NCDs) are chronic medical conditions or diseases which are non-infectious
- NCDs are currently responsible for over 60% of global deaths

12.0 DISEASE OUTBREAK

- A disease outbreak is the occurrence of a number of cases of a disease that is larger than expected for a given time and place.
- Outbreaks are classified according to the mode of spread (i.e. point source of propagation) and major clinical presenting features.

10 STEPS for Investigation of Disease Outbreak

1) Prepare to investigate
   a) Identify outbreak investigation team
   b) Review scientific literature
   c) Notify appropriate state and local entities
   d) Determine if immediate control measures are needed

2) Verify the diagnosis and confirm outbreak
   a) Get laboratory confirmation
   b) Collect stool specimens from ill persons

c) Perform bacteriologic, virologic or parasitic testing at the Georgia Public Health Laboratory (GPHL) Link patients and environmental specimens by DNA fingerprinting/Pulse Field Gel Electrophoresis (PFGE)

3) Case definition
   a) Establish a set of standard criteria for deciding who are the ill persons related to the outbreak ("case-patients")
   b) Narrow or broad (confirm, probable, suspect)
   c) DYNAMIC: may change during investigation

4) Case finding
   a) Conduct systematic search based on case definition
   b) Create line list of possible cases (people exposed)

5) Perform descriptive epidemiology
   a) Tabulate and orient data: person, place, time
   b) Frequencies
   c) Mapping
   d) Epidemic Curve

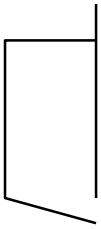
6) Hypothesis generation—the how and the why
   a) Compare with known sources or similar outbreaks
   b) Design questionnaire

7) Evaluate hypothesis thorough statistics
   a) Perform epidemiologic study: cohort, case-control
   b) Compare risk factors among ill (cases) vs not ill (controls)

8) Additional environmental studies
   a) Collect food, water, and/ or environmental samples
   b) Determine what happened with the implicated source or food

9) Implement control/prevention measure
   a) Coordinate with all stakeholders including regulatory partners
   b) Develop strategies to prevent further or future illness

10) Communicate findings
   a) Disseminate outbreak investigation report—internal and external audience
   b) Educate community, ill persons, restaurant staff, and Public Health Staff



TASK

Explain how the following influence disease transmission - herd immunity, incubation period, attack rate, carrier state and epidemics

Lesson 3: MEASUREMENT OF DISEASE OCCURRENCE (VITAL STATISTICS)

Objectives

At the end of the lesson the learner will be able to: -

1. Discuss the measurements of disease occurrence

1.0. INTRODUCTION

- Measurement is fundamental to Epidemiologic practice
- Epidemiological studies are designed to Identify disease determinants, describe and compare disease trends and evaluate public health interventions aimed at controlling health problems

2.0. MEASURES OF DISEASE FREQUENCY

- Measures of disease frequency in mathematical terms – count, proportion (percentage), rate and ratio
- Measures of disease frequency in epidemiology - prevalence and incidence
- Importance of Denominator

$$\frac{a \longrightarrow \text{Numerator}}{b \longrightarrow \text{Denominator}}$$

- Example
  - 500 cases of malaria in Turkana and 120 cases of malaria in Moshi
  - Which one is more infected? Turkana: 200/800,000 = 0.25/1,000 (25%) and Moshi: 120/300,000 = 0.4/1,000 (40%)

Counts

- Number of cases
- On its own offers little information

Ratio

- The division of two numbers (unrelated)
- Numerator not included in the denominator
- Allows comparison of quantities of different nature

Proportion

- The division of 2 numbers (related)
- Numerator always included in the denominator
- Quantities have to be of same nature
- Proportion always ranges between 0 and 1
- Percentage = proportion x 100

Rate

- The division of 2 numbers
- Time included in the denominator
- Speed of occurrence of an event over time
- Rate may be expressed in any power of 10: - 100, 1000, 10000, 100 000…

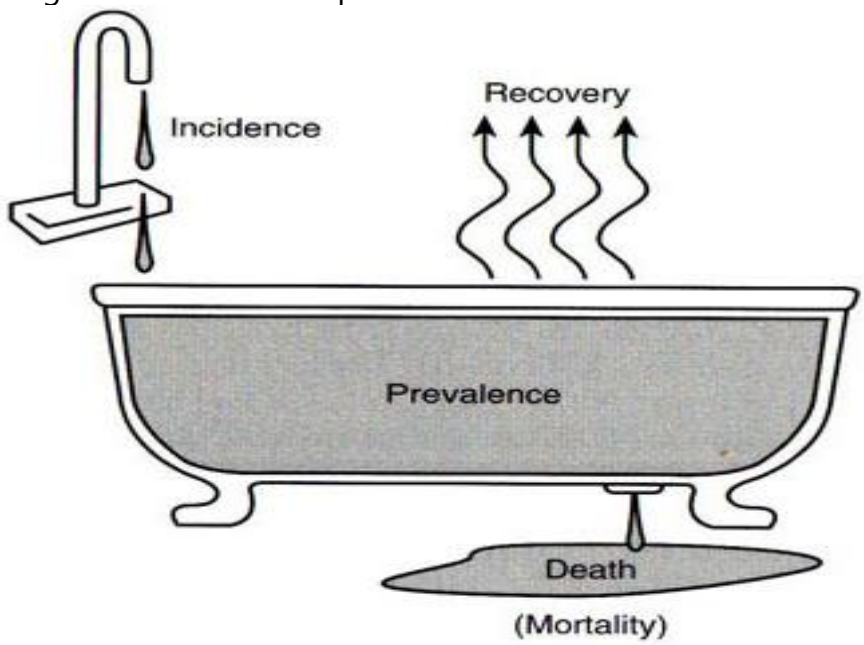
3.0. SOURCES OF MORBIDITY AND MORTALITY DATA

RECALL:
What are the causes of morbidity and mortality?

4.0. MEASUREMENT OF MORBIDITY (DISEASE OCCURRENCE)

- Is the measurement of the amount of disease in a given population?
- It is measured in terms of incidence and prevalence rates.

Diagram 3.1: Relationship between Incidence and Prevalence



4.1. INCIDENCE

- Incidence = number of <u>new</u> cases of a disease occurring in a specified time period divided by the number of individuals at risk of developing the disease during the same time
- Incidence of a disease is defined as the number of **new cases** of a disease that occur during a **specified period** of time in a **population at risk** for developing the disease

$$\text{Incidence} = \frac{\text{Number of New cases in a specified period}}{\text{Population at risk during the period}} \times 1000$$

4.2. PREVALENCE

- Prevalence = total number of affected individuals in a population at a specified time period divided by the number of individuals in the population at the time
- This is the proportion of a defined group having a disease at a given point in time
- It is the number of affected persons present in the population at a specific time divided by the number of persons in the population at that time
- Chronic diseases have high prevalence rates
- Low prevalence rate of a disease indicates that the disease is fatal or gets cured easily

$$\text{Prevalence} = \frac{\text{No of cases of a disease in the population in a specified period}}{\text{Number of persons in the population at that specified time}} \times 1000$$

4.3. RISK OF A DISEASE

- Risk is the probability that an individual with some characteristics (e.g. age, race, gender, etc.) will experience a health status change over a specified follow-up time (risk period)
- Estimated by observing events among a population during a specified time
- Measures of risk in epidemiology include absolute risk, relative risk and attributable risk

Absolute Risk

- This is the incidence or prevalence of a disease in a population
- Indicates the magnitude of the risk in a group of people with a certain disease

$$\text{Ratio of risk (or of the incidence rate)} = \frac{\text{Disease risk in exposed}}{\text{Disease risk in non-exposed}}$$

Difference in risks (incidence rate) = Disease risk (exposed) – Disease risk (non-exposed)

Relative Risk (RR) or Odds Risk (OR)

- RR is the ratio of incidence of disease in exposed individuals to the incidence of disease in non-exposed individuals (from a cohort/prospective study)

$$\text{Relative Risk (RR)} = \frac{\text{Risk in exposed}}{\text{Risk in non-exposed}}$$

$$\text{Incidence in exposed} = a/(a + b)$$

$$\text{Incidence in non-exposed} = c/(c + d)$$

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

- Interpretation
  - RR = 1 - Risk in exposed equal risk in non-exposed (No association)
  - RR > 1 - Risk in exposed greater than risk in non-exposed (positive association, possibly causal)
  - RR < 1 - Risk in exposed less than risk in non-exposed (Negative association, possibly protective)

	True Diagnosis		
Test Results	Disease	No Disease	Total
Positive	a (TP) 132	b (FP) 985	a + b (TP + FP) 1117
Negative	c (FN) 47	d (TN) 62, 295	c + d (FN + TN) 62, 342
Total	a + c (TP + FN)	b + d (FP+TN)	a + b + c + d (TP + FP + FN + TN)

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

$$RR = \frac{132/1117}{47/62,342}$$

$$= \frac{0.118}{0.00075}$$

= 157 (Risk in exposed > risk in non-exposed

- OR is the ratio of the odds that cases were exposed to the odds that controls were exposed(from a case control/retrospective study), is an estimate of the RR

Attributable Risk (AR)/Attributable Fraction (AF)

- AR is the amount of disease incidence that can be attributed to a specific exposure
- AF is the proportion of disease incidence that can be attributed to a specific exposure (among those who were exposed)

5.0. MEASUREMENT OF MORTALITY

- This is the measure of the amount death in a community
- Mortality rates include crude death rate (CDR), case specific death rate (CSDR), crude fatality rate (CFR), infant mortality rate (IMR) and maternal mortality rate (MMR)

Crude Death Rate (CDR)

- CDR is the total number of deaths to residents in a specified geographic area (country, state, county, etc.) divided by the total population for the same geographic area (for a Specified time period, usually a calendar year) and multiplied by 1000

$$CDR = \frac{\text{No. of deaths within a year}}{\text{Total mid-year population}} \times 1{,}000$$

Case Specific Death rate (CSDR)

- Measures the number of deaths attributed to a specific cause in a year divided by the total population that year

$$\text{Cause specific death rate} = \frac{\text{Number of death due to a particular cause (defined place and time period)}}{\text{Mid-period population (same place and time period)}} \times 1{,}000$$

Crude Fatality rate (CFR)

- Measures the proportion of episodes of illness that result in death
- Number of people who die from a disease out of those who had the disease within a given period of time divided by the number of cases of that disease in a given period

Infant Mortality rate (IMR)

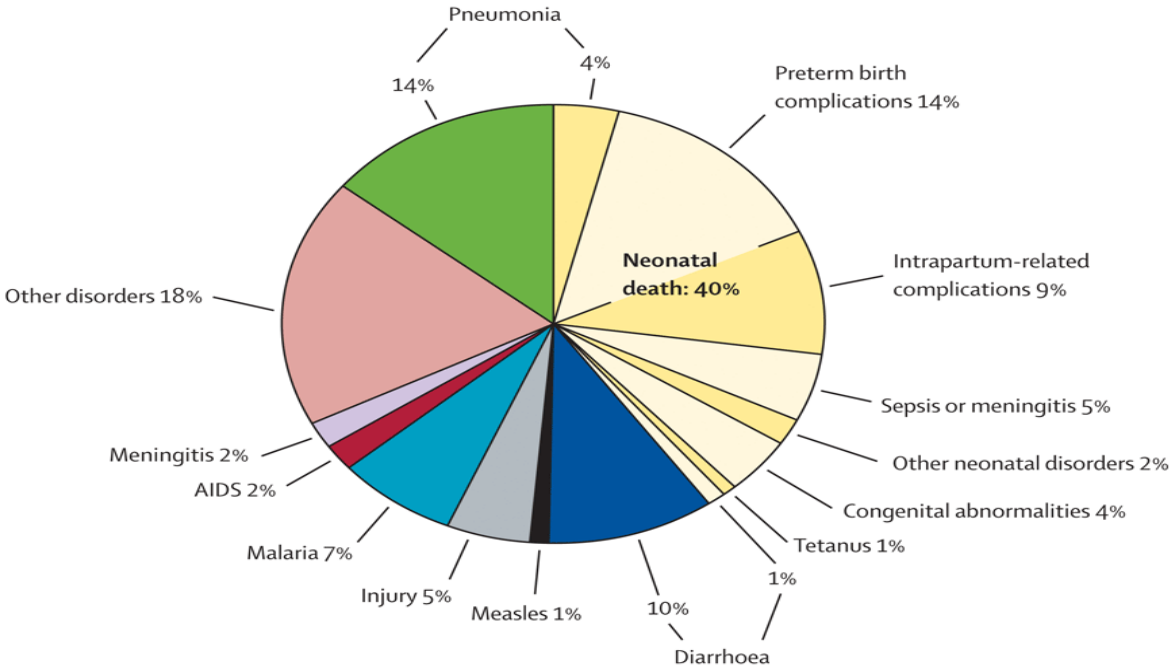
- Number of children who die before they are less than one year old per 1000 live births

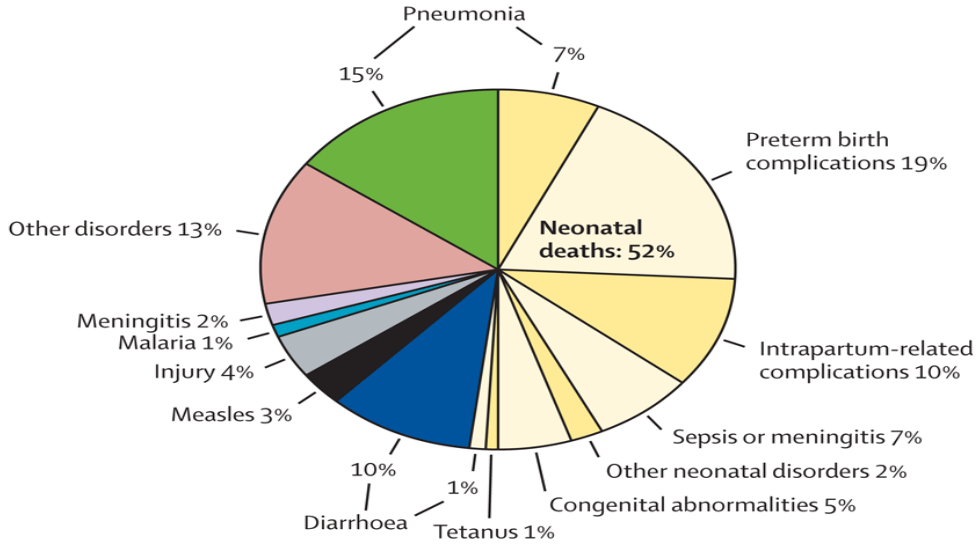
$$\text{Infant mortality rate} = \frac{\text{Number of death to infants <1 year of age (defined place and time period)}}{\text{Number of live births (same place and time period)}} \times 1{,}000$$

Infant mortality rate in Laos in 2005 ~ 70 / 1,000

$$\text{Age-specific death rate} = \frac{\text{Number of death in a particular age group (defined place and time period)}}{\text{Mid-period population (same age group, place and time period)}} \times 1{,}000$$

SUB SAHARAN AFRICA





Maternal Mortality rate (MMR)

- Number of women who die as a result of child bearing in a given year per 1000 live births

MMR = Number of maternal deaths due to child bearing x 1000
Number of live births per year

- **Mortality rate:** Death of a particular disease/event in the total population (e.g., maternal mortality)

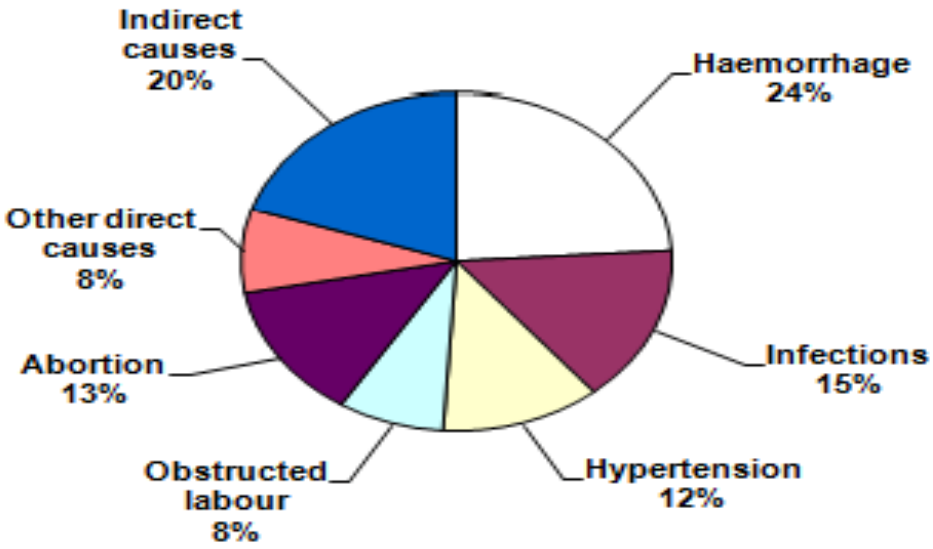
$$\frac{\text{Number of pregnancy – related death (defined place and time period)}}{\text{Number of live birth (same place and time period)}} \times 100{,}000$$

Maternal mortality rate in Laos in 2000 ~ 530/100,000

- The major direct causes of maternal morbidity and mortality include haemorrhage, infection, high blood pressure (eclasmpsia), unsafe abortion, prolonged and obstructed labour and other indirect causes including HIV/AIDS, malaria and TB, heart disease, anaemia

Factors Contributing To Maternal Mortality

1) Difficulty of predicting and/or preventing obstetric complications
2) Lack of access to good quality maternal health services
3) Poor health before and during pregnancy
4) Women's low social and economic status
5) Biological factors e.g. age, parity



6.0. MEASUREMENT OF FERTILITY

Crude Birth rate (CBR)

- Number of live births in a year divided by the total population in a year

$$\text{CBR} = \frac{\text{Number of live births in a year}}{\text{Total population that year}} \times 1000$$

General Fertility rate (GFR)

- Measure of fertility

$$\text{GFR} = \frac{\text{Number of live births in a year}}{\text{Number of women of child bearing age that year}} \times 1000$$

Lesson 4: DESCRIPTIVE EPIDEMIOLOGY

Objectives

At the end of the lesson the learner will be able to: -

1. Explain disease occurrence using the person-place-time model
2. Explain how 'person" variables determine disease occurrence

1.0. INTRODUCTION

- Descriptive epidemiology is the study of the amount and distribution of disease or health status within a population by person, place and time
- Involves determination of incidence, prevalence and mortality rates for diseases in large population groups according to characteristics such as age, sex, race (person) and geographical areas (place) as well as time distribution (cases per day, month, year)
- The **person-place-time** model describes disease occurrence in terms of **who** is affected by the disease (person), **where** the cases are (place) and **when** the disease occurs (time)
- These variables interact in a way that favours or does not favour the development of disease in an individual (the three variables must interact)

2.0. THE PERSON

- Personal characteristics influence one's exposure and susceptibility to disease. These characteristics include: -

1) Age
   - Single most important determinant of disease
   - Diseases can be classified based on age e.g. childhood diseases; diseases of the elderly
   - Diseases differ in terms of severity and frequency according to one's age
   - Age determines occupation
   - Can be used as the basis for determining type of risk factors

2) Sex
   - Influences disease occurrence
   - Rate of death is higher for males than females in all age groups (why?)
   - Some diseases affect specific sex e.g. cancer of the cervix (females) and cancer of the prostate (males)
   - Determines risk factors

3) Occupation
   - Contributes to different rates of morbidity and mortality
   - Determine type and amount of risk factors

4) Marital status and family
    - High marriage rate is a good indicator of prosperity
    - Mortality rates are lower in marred people
    - Some diseases are influence by marital status – e.g. breast cancer is low in married women than unmarried because breast feeding is protective against breast cancer

Discuss how this affects disease occurrence

5) Education
6) Social class
7) Ethnicity(race)
    - Racial differences interact with other factors to determine the occurrence and frequency of diseases
    - Skin colour may influence occurrence of rickets since black skin pigmentation absorbs more sunlight which promotes natural synthesis of vitamin D

3.0. PLACE

- Is the basis for describing the distribution of population and diseases
- Frequency of disease occurrence can be related to the place of occurrence as specified by natural barriers such as mountains, rivers, deserts or political/administrative boundaries because they influence environmental conditions e.g. iodine deficiency is common in highland and swampy places where soil is deficient of iodine; fluorosis is common where water has an excessively high content of fluorine

4.0. TIME

- Disease occurrence is dependent on what time of the month or year e.g. nutrient deficiency diseases become common after a period of drought
- Occurrence of some disease can be predicted as they conform to cyclic disease trends e.g. malnutrition, colds, malaria
- Some cannot be predicted that is they conform to secular trends e.g. Ebola
- Time trends

Outline the descriptive epidemiology of 10 communicable and 10 non-communicable diseases

Lesson 5: EPIDEMIOLOGICAL STUDY DESIGNS

Objectives

At the end of the lesson the learner will be able to: -

1. Describe different types of epidemiological studies/designs
2. State advantages and disadvantages of each study design

1.0. INTRODUCTION

- Epidemiological information is collected by conducting a study in the community
- It is difficult to collect information from all members of the community hence the need to sample the population
- Qualitative designs
  - Are complementary to quantitative designs, are important in study of social determinants of health problems
  - Qualitative methods include focus group discussion (FGDs), interviews, surveys, self-reports, observations and document analysis
  - Sampling is mainly purposive and quality assurance is attained through trustworthiness (credibility, conformability, dependability, transferability) and authenticity(fairness, ontological, educative, tactical, catalytic)
- Quantitative designs
  - Goal is to understand the frequency and causes of health-related phenomena
  - Seek causes by describing associations between exposures (causes) and outcomes
  - Quantitative methods include observational, experimental and mixed
  - Sampling is random (simple, stratified, cluster, etc.) or purposive
  - Quality assurance is through reliability and validity

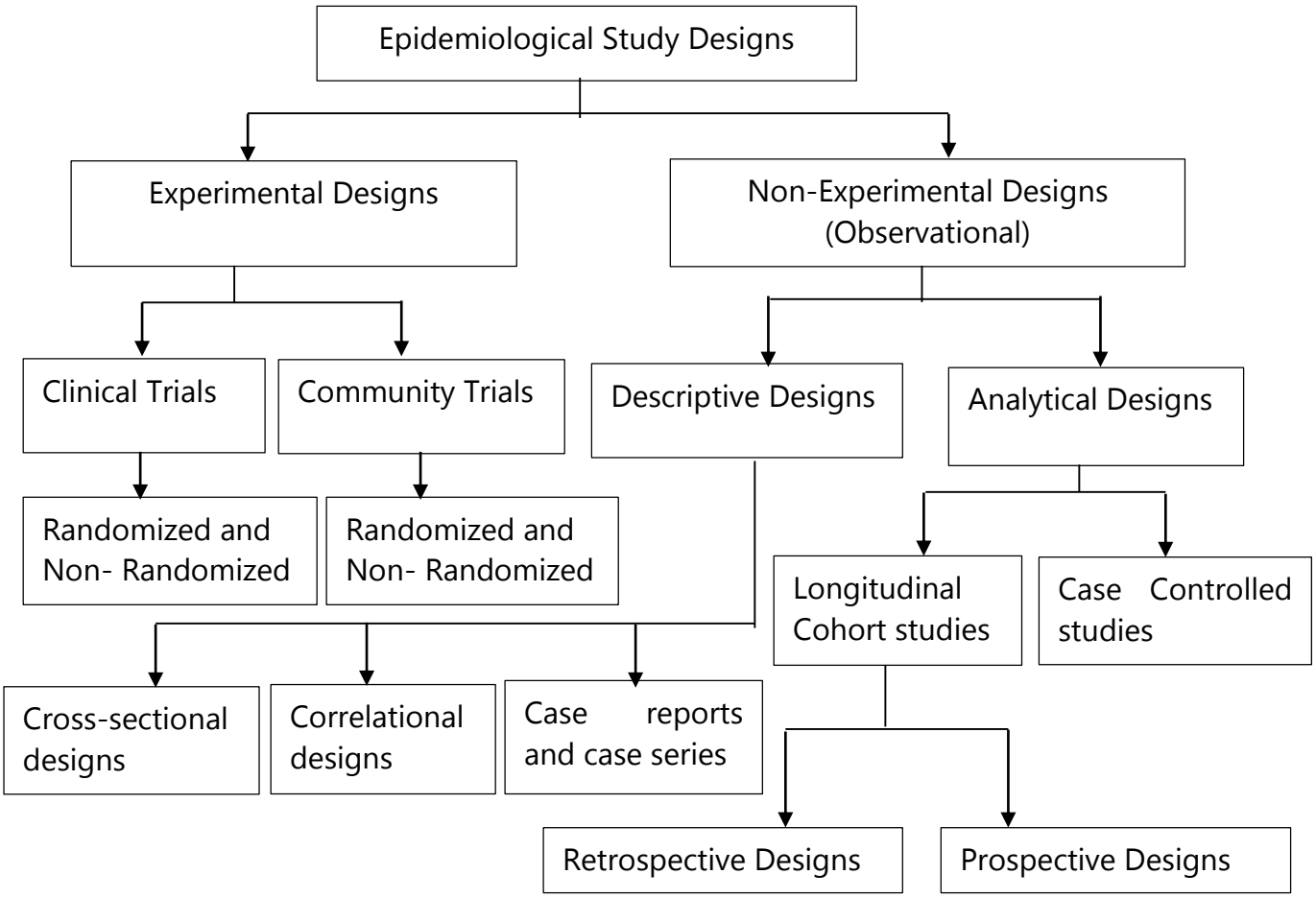
2.0. CHOOSING A DESIGN

1. Purpose of the study
   a. Hypothesis-testing versus hypothesis generating
   b. Finding signal versus quantifying the signal
2. Available resources
3. Need for data collection
4. Choice of outcome
5. Ability to draw valid causal inference

3.0. EPIDEMIOLOGICAL STUDY DESIGNS

- Broadly grouped as **experimental** and **non-experimental** (**observational**).
- Observational studies are further sub classified as **descriptive** and **analytical**
- Analytical designs can be case-controlled or cohort while descriptive can be cross-sectional, correlational and case series

- Experimental studies can be clinical or community trials (randomized or non-randomized)



4.0. OBSERVATIONAL (NON-EXPERIMENTAL) DESIGNS

4.1. DESCRIPTIVE STUDIES

- Provide information on patterns of disease occurrence
- Descriptive statistics generated can be correlated with clinical observations or laboratory studies to generate hypotheses
- Often provide clues about disease causation that can be pursued by more sophisticated research designs
- Three critical dimensions for describing a health condition are – person, place & time

Advantages

1) Uses routinely collected, readily available data
2) Less expensive and time-consuming
3) Good for assessing prevalence and patterns of disease occurrence
4) Useful in the formulation of research hypotheses – suggestive of risk factors

Disadvantages

1) Usually cannot test epidemiologic hypotheses
2) Lacks comparison group

3) Cannot usually discern a temporal relationship between an exposure and disease
4) Not useful for rare events
5) May be subject to selection bias due to refusal, death, etc.

Types

- Main types of descriptive studies include correlational studies, case reports and case series and cross-sectional studies

4.1.1. CORRELATIONAL STUDIES

- Measure of association: correlation coefficient (r)
- Linear association between exposure and outcome, ranging from -1 to 1
- Examples
  - Correlation of rate of a given disease and average amount of caloric intake, proportion of smokers, or median income
  - Death rates from coronary artery disease correlate with per capita cigarette sales

Uses of Correlational Studies

1) To suggest disease causation
2) To describe broad social and cultural attributes affecting health
3) Surveillance
4) To evaluate disease control measures Correlational Studies Summary

Advantages

1) Quick and relatively inexpensive
2) May be able to use readily available data
3) Useful in hypothesis generation

Disadvantages

1) Does not provide information about the relationship between risk factor levels and disease in individuals
2) Ecologic fallacy – association observed between variables on an aggregate level does not necessarily represent the association at an individual level

4.1.2. CASE REPORTS AND CASE SERIES

- Case reports and case series are observational, descriptive research designs.
- A case report is an in depth discussion of a patient's diagnosis, the intervention(s) used, and the outcome(s) achieved
- Case reports are typically conducted on one or two patients.
- A case series is a descriptive analysis of a series of patients with common characteristic(s) such as having a similar diagnosis or receiving a similar intervention

- Case reports and case series are most useful for describing the potential effectiveness of new interventions, for describing the effectiveness of interventions on unusual diagnoses, and for describing unusual responses (either good or bad) to interventions
- Case series are often conducted prospectively
- Primary distinction between case reports/series and the single-subject experiment is that the researcher does not manipulate the intervention in a case report/series but merely describes/documents what happened during the normal course of the intervention.

CASE SERIES

- Collections of individual case reports
- May occur in a relatively short time period
  1. Can indicate the beginning or presence of an epidemic
  2. Hypothesis formulation - through investigation of the experiences of the affected individuals
  3. Identification of possible causal factors – analytic study to compare experiences of the case series with a group of individuals who did not develop the disease
- Clinical case-series are of value in epidemiology for:
  i) Studying symptoms and signs
  ii) Creating case definitions
  iii) Clinical education, audit and research

- Population based
  - When a clinical case-series is complete for a defined geographical area for which the population is known, it is, effectively, a population based case-series consisting of a population register of cases.
  - Epidemiologically the most important case-series are registers of serious diseases or deaths, and of health service utilisation, e.g. hospital admissions.
  - Usually compiled for administrative and legal reasons.

- Population
  - Full epidemiological use of case-series data needs information on the population to permit calculation of rates
  - Key to understanding the distribution of disease in populations and to the study of variations over time, between places and by population characteristics
  - Case-series can provide the key to sound case control and cohort studies and trials

Requirements for interpretation

- To make sense of case-series data the key requirements are:
  i) The diagnosis (case definition) or, for mortality, the cause of death
  ii) The date when the disease or death occurred (time)
  iii) The place where the person lived, worked etc. (place)

iv) The characteristics of the person (person)
v) The opportunity to collect additional data from medical records (possibly by electronic data linkage) or the person directly
vi) The size and characteristics of the population at risk
vii) Case-series data can be linked to other health data either in the past or the future, e.g. mortality data can be linked to hospital admissions including at birth and childhood, cancer registrations and other records to obtain information on exposures and disease.
viii) Cases may also be contacted for additional information.
ix) This type of action may turn a case-series design into a cohort design.

Advantages

1) Population case-series permit two arguably unique forms of epidemiological analysis and insight
2) Paint a truly national and even international population perspective on disease.
3) The disease patterns can be related to aspects of society or the environment that affect the population but have no sensible measure at the individual level e.g. ozone concentration at ground level and the thickness of the ozone layer in the earth's atmosphere.
4) Useful in the formulation of research hypotheses – suggestive of risk factors
5) Important step in recognizing new diseases or risk factors

Disadvantages

1) Case report is based on the experience of one individual − The presence of any "risk factor" may be coincidental
2) Can't use to test for valid statistical association (No comparison group)
3) Can merely raise the question of an association

4.1.3. CROSS SECTIONAL DESIGNS (Community health studies, surveys)

- Is an observational study of a defined population at a single point in time or over a defined period of time
- Used to determine the prevalence of disease in the community
- In a cross-sectional study prevalent (existing) cases are identified rather than incident (new) cases
- Prevalent cases may not be representative of all cases in this population
- Often generate descriptive information on the prevalence of health outcomes or determinants of health

Advantages

1) Convenient and inexpensive, economic

2) Quick
3) Can consider several exposures and several diseases/allows study of several diseases
4) Can generate hypotheses
5) Usually represents the general population
6) Useful for estimation of the population burden, health planning and priority setting of health problems

Disadvantages

1) Cannot establish whether the exposure preceded disease or disease influenced exposure
2) Can identify only prevalent cases rather than incident cases
3) Possible bias since only survivors are available for study
4) May under-represent diseases with short duration
5) Liable to survivor bias
6) Possible measurement error;
7) Not suitable for rare conditions;

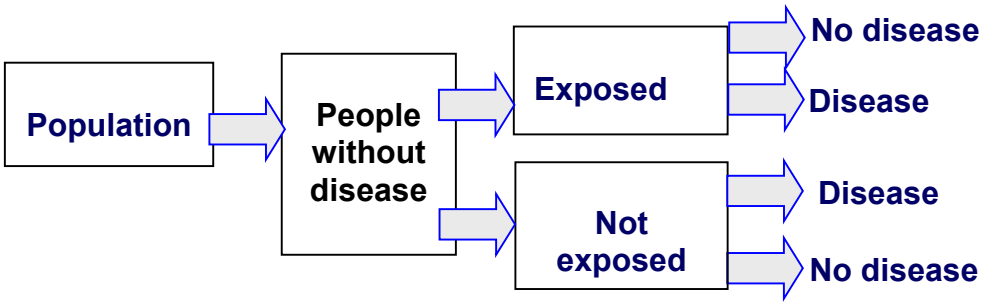
4.2. ANALYTIC DESIGNS

- Focus on searching for the underlying causes with the main purpose being to uncover the source and mode of spread of disease
- Are designed specifically to test hypotheses that have usually been generated from descriptive studies

4.2.1. LONGITUDINAL COHORT STUDIES

- Are observational where the goal of the investigator is to observe outcomes as they occur, rather than to experimentally manipulate outcomes by applying an intervention
- Begins by identifying a group of people who are initially free of the outcome of interest, but who vary in terms of their degree of exposure to various factors that may cause or prevent the outcome
- Subjects are then followed over time to determine whether the outcome of interest occurs
- Although cohort studies are observational, they can sometimes offer strong evidence about the effectiveness of an intervention, particularly when it is not ethical or efficient to randomize subjects to different therapies
- Involves data collection from different sample at each data collection point in a population that is constant e.g. checking dropout rate in schools
- Cohort studies
    i) Can be large or small
    ii) Can be long or short
    iii) Can be simple or elaborate
    iv) Can be local or multinational

v) For rare outcomes need *many people* and/or *lengthy follow-up*
vi) May have to decide what characteristics to measure *long in advance*



Intuitive approach to studying disease incidence and risk factors

1) Start with a population at risk
2) Measure characteristics at baseline
3) Follow-up the population over time with surveillance or re-examination
4) Compare event rates in people with and without characteristics of interest

Assumptions

1) Exposed and non-exposed groups are representative of a well-defined general population
2) Absence of exposure well defined
3) Outcome assessment comparable between exposed and non-exposed

Advantages

1) Can establish population-based incidence
2) Accurate relative risk (risk ratio) estimation
3) Can examine rare exposures (asbestos > lung cancer)
4) Temporal relationship can be inferred (prospective design)
5) Time-to-event analysis is possible
6) Can be used where randomization is not possible
7) Magnitude of a risk factor's effect can be quantified
8) Selection and information biases are decreased
9) Multiple outcomes can be studied (smoking > lung cancer, COPD, larynx cancer)

Disadvantages

1) Lengthy and expensive
2) May require very large samples
3) Not suitable for rare diseases
4) Not suitable for diseases with long-latency
5) Unexpected environmental changes may influence the association

6) Nonresponse, migration and loss-to-follow-up biases
7) Sampling, ascertainment and observer biases are still possible

Retrospective Cohort Designs

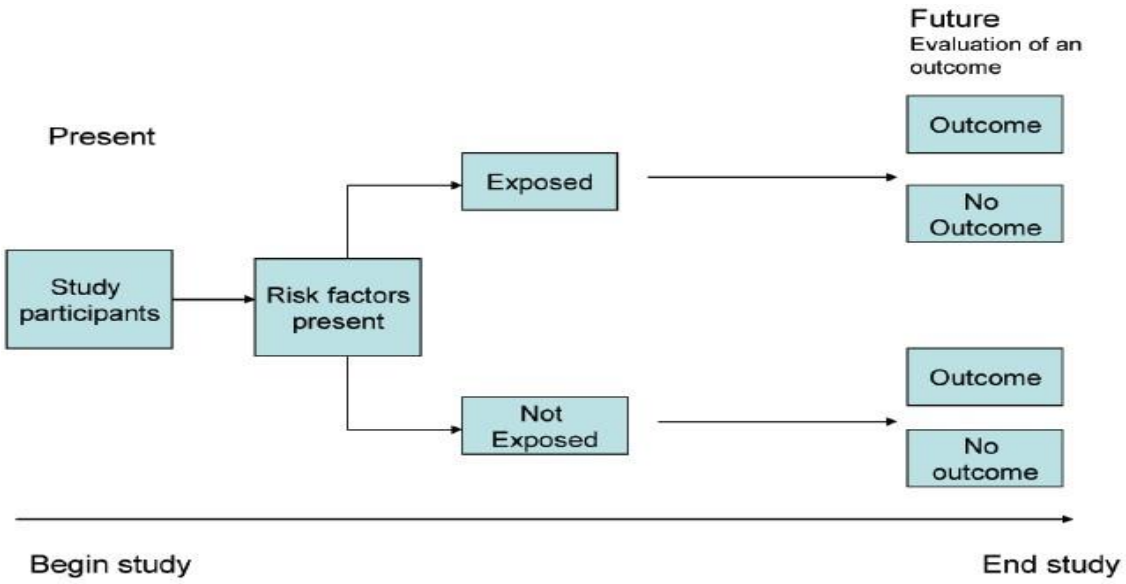
- Also called historical cohort; non-concurrent prospective cohort
- The cohort is identified from past records and followed from the time of those records up to some defined point in the recent past or up to the present time
- Study outcomes are assessed at this defined point and the investigator goes back in time to determine the subjects' exposures and risk factor characteristics
- An investigator accesses a historical roster of all exposed and non-exposed persons and then determines their current case/non-case status
- The investigator initiates the study when the disease is already established in the cohort of individuals, long after the measurement of exposure
- Requires sound data on exposure status for both cases and non-cases at a designated earlier time-point
- Investigator
  - uses existing data collected in the past to identify the population and the exposure status (exposed/not exposed groups)
  - Determines at present the (development) status of disease
  - Investigator spends a relatively short time to assemble study population (and the exposed/not exposed groups) from past data
  - Determine disease status at the present time (no future follow-up)
- Potential problems
  i) Selection bias
  ii) Misclassification
- Advantages
  i) Can be useful when reliable records are available
  ii) Investigators can determine the case status of the entire group at the present time, then use the exposure records to assess the relationship between exposure and disease

Prospective Cohort Designs

- The cohort is assembled in the present and followed into the future
- Investigator starts the study (from the beginning) with the identification of the population and the exposure status (exposed/not exposed groups)
- Follows them (over time) for the development of disease
- Study outcomes are recorded after baseline characteristics of subjects have been assessed
- Takes a relatively long time to complete the study (as long as the length of the study). Study outcomes are recorded after baseline characteristics of subjects have been assessed.

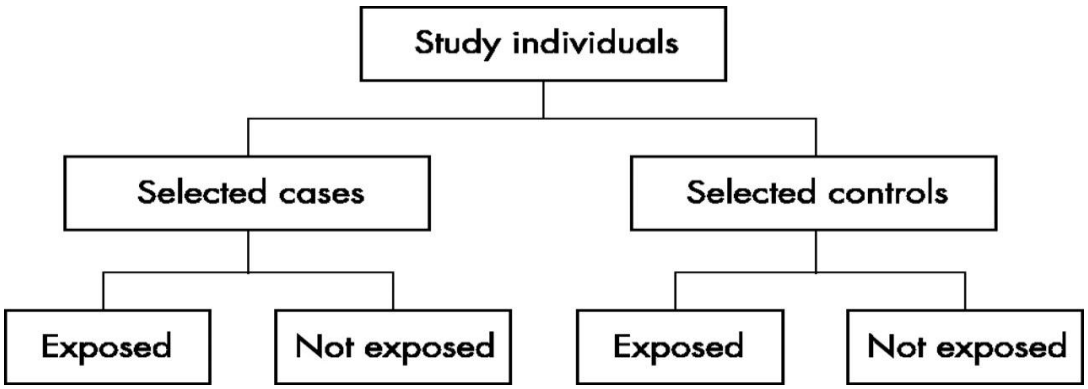
BIAS

1) Selection bias - select participants into exposed and not exposed groups based on some characteristics that may affect the outcome
2) Information bias
   a. Collect different quality and extent of information from exposed and not exposed groups
   b. Loss to follow-up differs between exposed and not exposed (or between disease and no disease)
3) Misclassification bias - misclassify exposure status or disease status



4.2.2 CASE CONTROLLED STUDIES

- Identify participants based on their disease/outcome status and compare presence of risk factor
- Subjects who are known to have the outcome of interest are first identified as cases, and then a comparison(control) group, is assembled
- Both groups are then evaluated and the distribution of past exposures, risk factors, and/or subject characteristics in the case and control groups are compared



- In a retrospective case-control study
  - Assessment of the exposure or risk factor occurs after subjects are classified as cases or controls
- In a prospective case-control study
  - All measurements of the exposure or risk factor variables are recorded before subjects are classified as cases or controls

Assumptions

1) Cases representative of all cases of disease
2) Controls drawn from the same population as cases (and at risk for the outcome)
3) Exposure data collected similarly in cases and controls

Case Selection

- Cases are identified on the basis of their disease/phenotype, representative of all individuals who develop disease
- Distinguishing incident from prevalent or recurrent cases important
- High participant rates important

Controls Selection

- Controls are representative of the general population who do not develop the disease
  - Selected from population at risk to become case
  - Families, population registries, neighbourhood
- Who is the population at risk?
- How do you know they don't have the disease?

Examples

- Aspirin and Reye's syndrome in children
- Oral contraceptives and reduced risk of ovarian/endometrial cancer

Advantages

1) Cheap, easy and quick studies
2) Multiple exposures can be examined
3) Rare diseases and diseases with long latency can be studied
4) Suitable when randomization is unethical (alcohol and pregnancy outcome)
5) Suitable for rare outcomes
6) Suitable for outcomes with long induction period
7) Need fewer people in some cases
8) Readily evaluate multiple exposures
9) Convenient

Disadvantages

1) Case and control selection troublesome
2) Subject to bias (selection, recall, misclassification)
3) Direct incidence estimation is not possible
4) Temporal relationship is not clear
5) Multiple outcomes cannot be studied
6) If the incidence of exposure is high, it is difficult to show the difference between cases and controls
7) Not easy to estimate attributable fraction
8) Reverse causation is a problem in interpretation - especially in molecular epidemiology studies

5.0. EXPERIMENTAL DESIGNS

- A population is selected for a planned trial of a regimen, whose effects are measured by comparing the outcome of the regimen in the experimental group versus the outcome of another regimen in the control group
- Such designs are differentiated from observational designs by the fact that there is manipulation of the study factor (exposure), and randomization (random allocation) of subjects to treatment (exposure) groups.
- Investigator studies the effect of a factor under his/her control by determining which group to expose to the factor under study

WHY PERFORMED

1) Provide stronger evidence of the effect (outcome) compared to observational designs, with maximum confidence and assurance
2) Yield more valid results, as variation is minimized and bias controlled
3) Determine whether experimental treatments are safe and effective under "controlled environments" (as opposed to "natural settings" in observational designs), especially when the margin of expected benefit is doubtful / narrow (10 - 30%)

5.1 COMMUNITY TRIALS

- Experiments involving communities as a whole are called community trials
- The group as a whole is collectively studied as opposed to clinical trials where an individual within a group is studied
- In a community trial,
  - Investigator selects two communities that have very similar characteristics
  - Obtains community assent from leaders
  - Conducts a survey in each community to measure incidence and prevalence of the disease of interest and prevalence of suspected risk factors
  - Introduce interventions

- Conduct a survey
- Note the net difference

- Example - introduction of fluorides into water supply in order to determine whether it decreases frequency of dental caries

Conducting a community trials

- Six steps
  1) Development of a protocol
     - State the rationale, procedures and organization
     - Detailed description of methods of assessment
     - Methods of data analysis
     - Provide contingences

  2) Community selection and recruitment
     - Communities should be similar in as many aspects as possible
     - Select interventions and control
     - Determine the number and size of the communities
     - Selected communities should be stable with little migration
     - Based on 5 criteria
       i) Unusual prevalence of the disease
       ii) Unusual prevalence of suspected risk factors
       iii) Administrative convenience
       iv) Favourable community relations
       v) Availability of background demographic information
     - Recruitment involves
       i) Obtain consent from leaders, officials and health professionals
       ii) Inform community members about the study – newspapers, local radio and TV stations, barazas

  3) Establishment of a baseline and community surveillance
     - Examples - changes in mortality rates, incidence of disease and changes in risk factors for disease
     - Once an investigator selects outcomes of interest then he/she must select the baseline for the outcomes
     - Community surveillance (6 steps)
       i) Development of a protocol – solicit community support
       ii) Community outreach
       iii) Establish a diagnostic criteria
       iv) Ascertainment of cases
       v) Validation
       vi) Data management

4) Intervention selection and assignment
    - Specify the intervention to be used e.g. education
    - If the intervention has any injurious effect it is has to be stopped
    - Data collection also allows surveillance
5) Oversight and data monitoring (involves data collection and analysis)
6) Evaluation (data analysis and making inferences and conclusions)

Randomized Community Trials

- Communities are selected using simple random sampling

Non- Randomized Community Trials

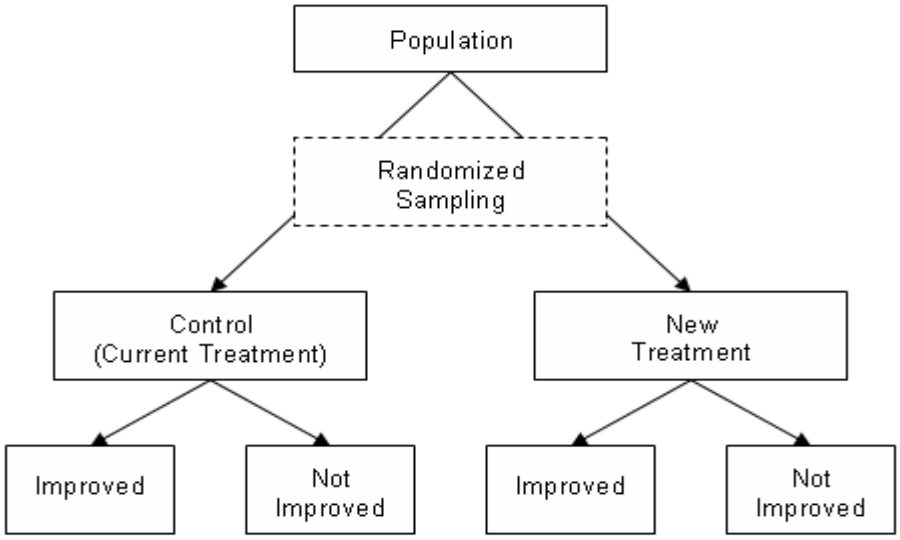
- Communities are selected by non-probability methods
- An element of bias is possible to ensue

5.2 CLINICAL TRIALS

- Is an experiment that evaluates the effects of one or more therapeutic interventions in groups of human subjects
- Clinical trials - an individual within a group is studied
- Effect of an intervention is evaluated by comparing a group of subjects receiving a standard intervention to a group receiving a new intervention. If no standard intervention exists, then a control group receiving no intervention or a placebo intervention can be used as the comparison.
- Assignment to a given intervention group may be random, as when a coin is tossed or a random number generator is used, or it may be entirely non-random.
- Random assignment or randomization improves the ability to attribute differences between groups to differences in the types of interventions received. Although randomization is preferable because it helps to control for factors other than the intervention that may be responsible for differences in outcomes, non-randomized assignment may be necessary when randomization is not feasible or ethical.
- Additionally, when little evidence is available to justify the difficulty and expense of a fully-randomized clinical trial, a non-randomized clinical trial might be the first step in that direction.
- Non-randomized clinical trials are sometimes referred to as “quasi-experimental” clinical trials or “non-equivalent control group” designs because the characteristics of subjects in non-randomized groups will tend to be non-equivalent. The estimation of intervention effects in non-randomized clinical trials may be biased if group differences in subject characteristics are not controlled for in the data analysis

Randomized Clinical Trials

- A comparative study in which study subjects are assigned by a formal chance mechanism between two or more intervention strategies (gold standard for inferring causality)
- Also called "randomized controlled trials, , experimental studies
- Participant assigned to intervention group by a formal chance mechanism
- Assumptions
    - Exposure must be potentially modifiable
    - Primary outcomes are relatively common, occur relatively soon



-

Methods of randomization

- Several choices, from "flipping a coin" to stratified randomization
- Blinding/masking
    - Participant, study investigator (and anybody else involved in follow-up)
    - Ideally, double-blinded
- Analysis: intention-to-treat

Advantages

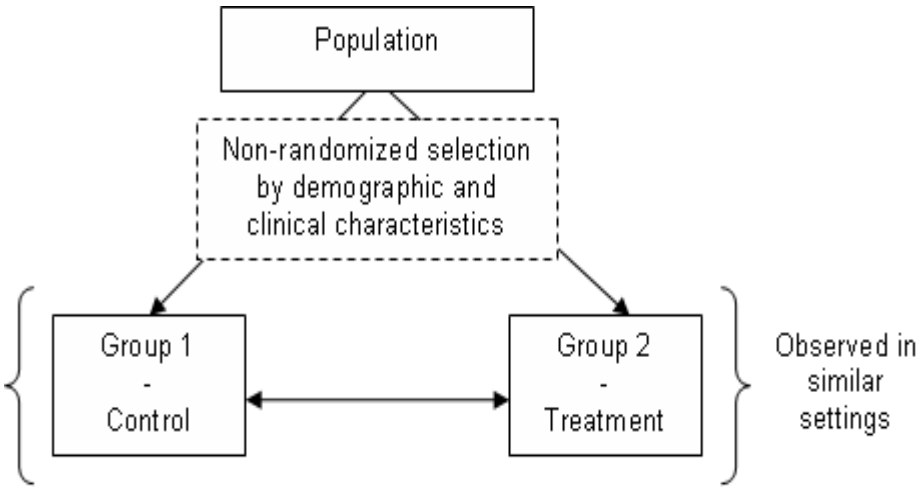
1) Similar distribution of baseline characteristics in comparison groups
2) Protection against confounders, both known and unknown
3) Able to directly estimate risk
4) Allows comparison of multiple outcomes
5) Best evidence study design
6) No inclusion bias (using blinding)
7) Controlling for possible confounders
8) Comparable Groups (using randomization)

Disadvantages

1) Limitations on types of interventions
2) Costly
3) Not suitable for rare outcomes
4) Not suitable for outcomes requiring long or extensive follow-up
5) Potential challenges to the generalizability of findings
    - Eligibility: strict inclusion/exclusion
    - Adherence/withdrawal issues
4) Large trials (may affect statistical power)
5) Long term follow-up (possible losses)
6) Compliance
7) Expensive
8) Public health perspective?
9) Possible ethical questions

Non-Randomized Studies

A clinical trial in which the participants are not assigned by chance to different treatment groups. Participants may choose which group they want to be in, or they may be assigned to the groups by the researchers.



Lesson 6: SAMPLING METHODS AND DATA ANALYSIS IN EPIDEMIOLOGY

Objectives

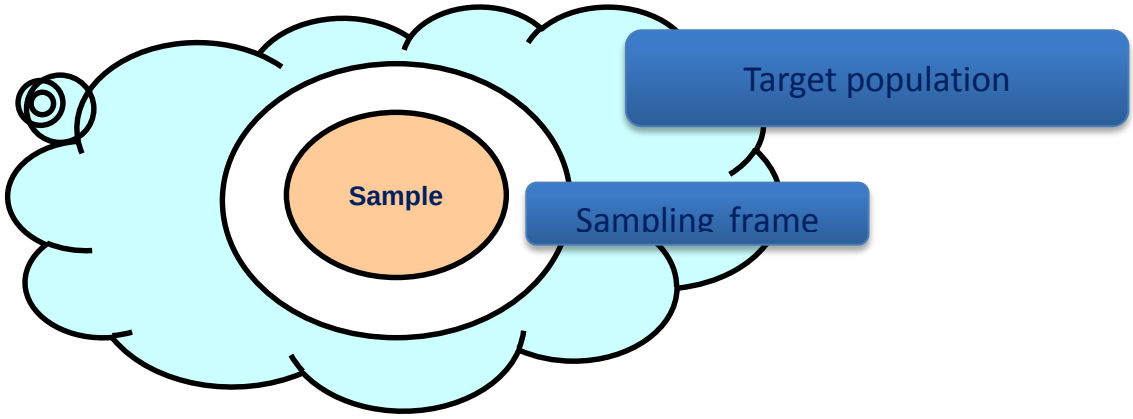
At the end of the lesson the learner will be able to: -

1. Define key words
2. Describe methods of sampling in epidemiological studies
3. Describe methods of data analysis and presentation in epidemiological studies

SAMPLING METHODS

1.0. INTRODUCTION

- Sampling is the process or technique of selecting a sample of appropriate characteristics and adequate size.
- Definitions
  - i) A population is defined by:
    - Its nature (an individual, housing, a firm etc.)
    - Its intrinsic characteristics (gender, housing type, industries)
    - Its localisation (city, neighbourhood etc.)
    - who, where and when
  - ii) Sampling unit – the basic unit around which a sampling procedure is planned e.g., a person, group (household, school, district, etc.) or a component (e.g. eye, physiological response)
  - iii) Sampling frame – list of all of the sampling units in a population
  - iv) Sample – collection of sampling units from the eligible population



2.0. SAMPLING METHODS

1) Random (Probability) Sampling
   a) Simple random sample
   b) Stratified random sample
   c) Cluster sample
   d) Adaptive cluster sample
   e) Multistage sample

2) Non-random Sample
   a) Convenience sample
   b) Systematic sample
   c) Consecutive sample
   d) Quota sample
   e) Volunteer sample
   f) Capture-recapture

2.1. PROBABILITY SAMPLING METHODS

1) Simple Random sampling

- Simple random sampling *without replacement* a simple random sample is one in which each of the possible samples of elements taken from a population of elements has the same probability of selection
- In a simple random sample *without replacement*, any element selected in a sample cannot be selected again for the same sample
- Advantages
  i) Simple process and easy to understand
  ii) Easy calculation of means and variance
  iii) Sampling error easily measured
- Disadvantages
  i) Not most efficient method, that is, not the most precise estimate for the cost
  ii) Requires knowledge of the complete sampling frame
  iii) Cannot always be certain that there is an equal chance of selection
  iv) Non respondents or refusals

2) Systematic sampling

- Sampling units are spaced regularly throughout the sampling frame, e.g., every 3rd unit would be selected
- May be used as either probability sample or not
  - Not a probability sample unless the starting point is randomly selected
  - Non-random sample if the starting point is determined by some other mechanism than chance
- Principle - Select sampling units at regular intervals (e.g. every 20th unit)
- Procedure
  - Arrange the units in some kind of sequence
  - Divide total sampling population by the designated sample size (e.g. 1200/60=20)
  - Choose a random starting point (for 20, the starting point will be a random number between 1 and 20)
  - Select units at regular intervals (in this case, every 20th unit), i.e. 4th, 24th, 44th etc.
- Advantages
  i) Sampling frame does not need to be defined in advance

ii) Easier to implement in the field
iii) If there are unrecognized trends in the sample frame, systematic sample ensure coverage of the spectrum of units

- Disadvantages
  i) Variance cannot be estimated unless assumptions are made
- Example: Estimate HIV prevalence in children born during a specified period at a hospital

3) Stratified random sample

- The sampling frame comprises groups, or strata, with certain characteristics
- A sample of units are selected from each group or stratum
- Used for a population with distinct subgroups
- Procedure
  - Divide (stratify) sampling frame into homogeneous subgroups (strata) e.g. age-group, urban/rural areas, regions, occupations
  - Draw random sample within each stratum
- Advantages
  i) Assures that certain subgroups are represented in a sample
  ii) Allows investigator to estimate parameters in different strata
  iii) More precise estimates of the parameters because strata are more homogeneous, e.g., smaller variance within strata
  iv) Strata of interest can be sampled most intensively, e.g., groups with greatest variance
  v) Administrative advantages
- Disadvantages
  i) Loss of precision if small number of units is sampled from strata
  ii) Sampling error is difficult to measure
  iii) Different strata can be difficult to identify
  iv) Loss of precision if small numbers in individual strata (resolved by sampling proportional to stratum population)

4) Cluster sampling

- Clusters of sampling units are first selected randomly
- Individual sampling units are then selected from within each cluster
- Principle
  - Whole population divided into groups e.g. neighbourhoods
  - A type of multi-stage sampling where all units at the lower level are included in the sample
  - Random sample taken of these groups ("clusters")
  - Within selected clusters, all units e.g. households included (or random sample of these units)
  - Provides logistical advantage

- Advantages
  i) The entire sampling frame need not be enumerated in advance, just the clusters once identified
  ii) More economical in terms of resources than simple random sampling
- Disadvantages
  i) Loss of precision, i.e., wider variance, but can be accounted for with larger number of clusters
- Example: Estimate the prevalence of dental caries in school children
  1. Among the schools in the catchments area, list all of the classrooms in each school
  2. Take a simple random sample of classrooms, or cluster of children
  3. Examine all children in a cluster for dental caries
  4. Estimate prevalence of caries within clusters than combine in overall estimate, with variance

5) Multistage sampling

- Similar to cluster sampling except that there are two sampling events, instead of one
- Primary units are randomly selected
- Individual units within primary units randomly selected for measurement
- Example: Estimate the prevalence of dental caries in school children
  1. Among the schools in the catchment area, list all of the classrooms in each school
  2. Take a simple random sample of classrooms, or cluster of children
  3. Enumerate the children in each classroom
  4. Take a simple random sample of children within the classroom
  5. Examine all children in a cluster for dental caries
  6. Estimate prevalence of caries within clusters than combine in overall estimate, with variance

2.2 NON-PROBABILITY SAMPLING METHODS

1) Convenience (Accidental, Haphazard) Sampling

- A non-random collection of sampling units from an undefined sampling frame
- Advantages
  i) Convenient and easy to perform

- Disadvantages
  i) Not statistical justification for sample
  ii) Biased

- Examples - case series of patients with a particular condition at a certain hospital, "normal" graduate students walking down the hall are asked to donate blood for a study, children with febrile seizures reporting to an emergency room
- Investigator decides who is enrolled in a study

- A consecutive sample
    - A case series of consecutive patients with a condition of interest
    - Consecutive series means ALL patients with the condition within hospital or clinic, not just the patients the investigators happen to know about
    - Advantages
        i) Removes investigator from deciding who enters a study
        ii) Requires protocol with definitions of condition of interest
        iii) Straightforward way to enrol subjects

    - Disadvantage
        i) Non-random

2) Snowball

- The researcher asks the respondents to give referrals to other possible respondents
- Used in studies in populations where it is difficult to get respondents e.g. drug addicts, homeless people, individuals with HIV/AIDS, prostitutes
- Such populations are hard to reach and/or hidden because they exhibit some kind of social stigma

3) Purposive sampling (Judgemental/Selective)

- We sample with a purpose in mind
- Focus on particular characteristics of a population that are of interest
- Types – maximum variation, homeogneous, typical case, extreme(deviant) case, critical case, expert and total population sampling

4) Quota sample

- Researcher selects respondents according to a some fixed quota (representative group)
- The representative individuals are chosen out of a specific group

5) Volunteer sample

6) Capture-recapture

3.0. DATA ANALYSIS

Lesson 7: DISEASE SCREENING IN POPULATIONS

Objectives

At the end of the lesson the learner will be able to: -
1. Describe types of screening and screening tests
2. Discuss characteristics of screening tests

1.0. INTRODUCTION

- Different kinds of testing in medicine
  1) "Diagnostic" - specifically looking for a suspected condition which is tested for and confirmed or excluded
  2) "Case-finding" - usually in an investigation of exposed people, to sort the exposed and ill from the exposed and well (E.g., test people who were in contact with a case of tuberculosis, or check blood pressure of patient who is overweight)
  3) "Screening" - usually no specific exposure or indication that the individual has disease. (E.g. routine PSA testing in middle-aged males)

2.0. SCREENING

- Defined as a relatively quick means of detecting potential diseases before it has become manifest
- Screening does not establish a diagnosis but helps to identify individuals who appear to have a given disease and who should, therefore, undergo further testing to determine if the disease is actually present.
- Screening for diseases detection is valuable where early identification of a disease leads to more effective treatment and a better prognosis for the individual e.g. screening for breast, colorectal, cervical cancer and screening for hypertension, diabetes and rubella.
- Depending on its intent, screening can take place at different stages in the natural history of disease and can represent different levels of prevention.
- Screening is a mode of primary or secondary prevention
- Screening programs, especially mass screening, can be expensive and time consuming
- Require valuable resources
- Examples of disease and conditions for screening include neonatal hypothyroidism, hypertension, breast cancer, cervical cancer, syphilis, gonorrhoea, cystic fibrosis, colorectal cancer, diabetes, tuberculosis, diminished visual acuity, hearing impairment, elevated cholesterol, lead poisoning, HIV infection, Down's syndrome, iron deficiency anaemia, neural tube defects and obesity
- Examples of diagnostic/screening tests include mammography (preclinical breast cancer), alpha-fetoprotein (baby with neural tube defect), electrocardiography (heart disease), serological testing (e.g. infection) and somatic cell counts (subclinical mastitis in cattle)

3.0. REASONS FOR SCREENING

1) Identify risk factors for disease
2) Reduce morbidity and mortality
3) Improved quality of life

4.0. CONSIDERATIONS FOR SCREENING

1) Nature of the disease - the disease should
   a) Be a significant public health problem that can be detected before symptoms develop
   b) Be one that the public views as important enough to submit to screening
   c) One that is likely to yield enough cases for the screening program to be cost effective
   d) Have a relatively high prevalence in pre-symptomatic stage.

2) Nature of the screening test
   a) Should be accurate (valid and precise)
   b) Relatively quick and simple to apply
   c) Easily interpreted
   d) Safe
   e) Acceptable to those being screened
   f) Relatively inexpensive

3) Nature of the follow-up tests
   a) Must be available to confirm the suspected diagnosis
   b) Should be readily accessible, accurate, generally safe, and acceptable to those with presumed disease from the standpoint of comfort and cost

4) Nature of the treatment
   a) Must be available, accessible, and acceptable to patients, and it must be more effective than if it or another treatment were initiated at a later stage of the disease
   b) Should result in significantly better prognosis when the disease is treated in its early stages before symptoms arise

5.0. TYPES OF SCREENING

- Screening for disease detection is often categorized into the following four types:

1) Mass screening

- Aimed at large population groups that vary widely in their risk of the disease e.g. screening for visual impairment in elementary schools is another example of mass screening

2) Selective screening

- Applied only to groups at high risk for the disease e.g. screening for elevated blood lead levels among inner-city children is an example of selective screening. So is screening for tuberculosis among prison inmates.
- Expected to detect more potential cases of a given disease than mass screening because of the difference in risk profiles between the populations being screening
- Is sometimes referred to as targeted screening.

3) Multiphase screening

- Employs multiple screening tests at the same time
- May be used to detect the possibility of more than one disease or condition.
- Paramilitary exams, for example, may use multiphase screening to test for possible diabetes, hypertension, and hearing impairment.

4) Case finding

- Occurs in a clinical setting when patients visit their physician (or other health provider) for general consultation or unrelated problems, and physician takes the opportunity to request one or more routine screening tests.
- Places the responsibility for follow-up on the physician performing or supervising the screening test. Therefore, case finding is more likely to result in follow-up than other types of screening
- Many individuals identified as having elevated blood pressure during a mass screening, for example, may not seek the recommended follow-up, but a physician finding elevated blood pressure during a routine examination will ordinary schedule additional tests.
- Examples of case finding include screening for cervical cancer using Pap tests, heart abnormalities using an electrocardiogram, weight changes using a calibrated scale, and diabetes using blood tests or urine samples. In addition, optometrists and ophthalmologists routinely screen patients for glaucoma
- Case finding has been referred to as opportunistic screening

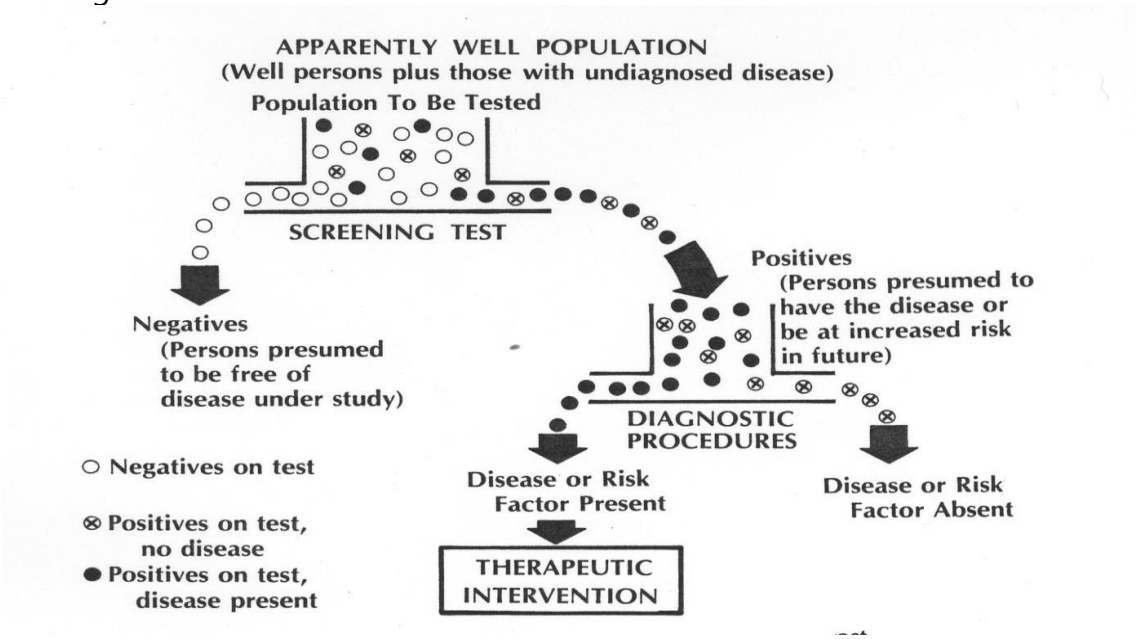
6.0. PRINCIPLES FOR SCREENING PROGRAMS

1. Condition should be an important health problem
2. There should be a recognizable early or latent stage
3. There should be an accepted treatment for persons with condition
4. The screening test is valid, reliable, with acceptable yield
5. The test should be acceptable to the population to be screened
6. The cost of screening and case finding should be economically balanced in relation to medical care as a whole

7.0. CHARACTERISTICS OF A GOOD SCREENING TEST

7.1. Validity of Screening Tests

- Defined as its ability to distinguish between who has a disease and who does not
- Has 2 components ¬ **sensitivity** (ability of a test to identify correctly those who have the disease or true positives) and **specificity** (ability of a test to identify correctly those who do not have a disease or true negatives)
- The validity of a screening test is measured by its ability to correctly categorize persons who have pre-clinical disease as test-positive and those without pre-clinical disease as test-negative
- The precision of a screening test refers to its reliability, that is, its consistency from one application to the next
- A positive test implies that the disease is likely to present, and follow-up diagnostic tests are therefore advisable
- A negative test implies that the disease is unlikely to be present, and follow-up diagnostic are therefore not indicated.



Sensitivity of Screening Test

- The probability of a positive test result given that the individual tested actually has the disease
- A test with high sensitivity will have few false negatives

		Disease	
Mammography		Cancer	No cancer
	Positive	132 (TP)	985 (FP)
	Negative	47 (FN)	62,295 (TN)
	Total	179 (TP + FN)	63,280 (FP + TN)

Sensitivity = TP/(TP + FN)

= 132 / 179 = 73.7%

Specificity of Screening Test

- The probability of a negative test result given that the individual tested actually does not have the disease
- A test that has high specificity will have few false positives

EXAMPLE: Mammographic screening in the Health Insurance Plan (HIP) New York

Mammography		**Disease**	
		Cancer	No cancer
	Positive	132 (TP)	985 (FP)
	Negative	47 (FN)	62,295 (TN)
	Total	179 (TP + FN)	63,280 (FP + TN)

Specificity = TN/(FP + TN)

= 62,295 / 63,280 = 98.4%

Relating Specificity and Sensitivity

	True Diagnosis		
Test Results	Disease	No Disease	Total
Positive	a (TP)	b (FP)	a + b (TP + FP)
Negative	c (FN)	d (TN)	c + d (FN + TN)
Total	a + c (TP + FN)	b + d (FP+TN)	a + b + c + d (TP + FP + FN + TN)

- Specificity + false positive rate = 1
  - (d/(b+d) + b/(b+d) = 1)
  - If the specificity is increased, the false positive rate is decreased
  - If the specificity is decreased, the false positive rate is increased
- Sensitivity + false negative rate = 1
  - (a/(a+c) + c/(a+c) = 1)
  - If the sensitivity is increased, the false negative rate is decreased
  - If the sensitivity is decreased, the false negative rate is increased

Predictive Value

- Predictive value of a positive test is the likelihood that a person with a positive test has the disease (the probability that an individual actually has the disease given that he or she tests positive)

- Predictive value of a negative test is the likelihood that a person with a negative test does not have the disease (the probability that an individual is disease-free given that he or she tests negative)
- Predictive value depends on the prevalence of disease

EXAMPLE: Mammographic screening in the Health Insurance Plan (HIP) New York

		Disease		
Mammography		Cancer	No cancer	Total
	Positive	132 (TP)	985 (FP)	1117 (TP + FP)
	Negative	47 (FN)	62,295 (TN)	62,342 (FN + TN)
	Total	179 (TP + FN)	63,280 (FP + TN)	

Positive Predictive Value (PPV) - proportion of test positives that truly have the condition

PPV = TP/(TP+FP)

= 132/1117 = 11.8% = 12% (88 false positives out of 100 positive tests)

Negative Predictive Value (NPV) - proportion of test negatives that truly do not have the condition

NPV = TN/(TN+FN)

= 62, 295/62, 342 = 99.9%

7.2. Reliability (Precision) Of Screening Test

- Reliability is the ability of a test to give consistent results when performed more than once on the same individual under the same conditions
- Can be improved through
  1) Variation in the method due to variability of test chemicals or fluctuation in the item measured (e.g., diurnal variation in body temperature or in relation to meals)
  2) Standardize fluctuating variables
  3) Use standards in laboratory tests, run multiple samples whenever possible
  4) Observer variation (train observers and use more than one observer and have them check each other)

7.3. Yield of Screening Test

- Is the amount of previously unrecognized disease that is diagnosed and brought to treatment as a result of the screening program
- Results from

1) Sensitivity - you must detect a sufficient population of disease to be useful
2) Prevalence of unrecognized disease - screen high risk populations
3) Frequency of screening - screening on a one time basis does not allow for the natural history of the disease, differences in individual risk, or differences in onset and diseases have lead time
4) Participation and follow-up - tests unacceptable to those targeted for screening will not be utilized

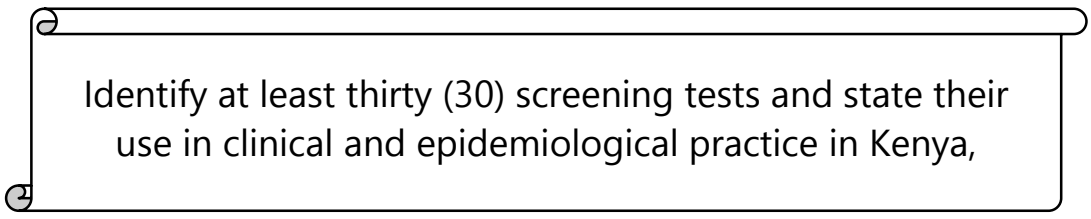
7.4. Efficacy of Screening Test

- Benefits of a treatment, procedure, or service among those who use it compared to those who don't
- Measured by the degree to which it benefits those who are screened compare to those who are not
- Most evaluations of the efficacy of screening are not based on randomized controlled trials, however, often because of costs, feasibility, or ethical limitations
- More often than not, evaluations of screening utilize observational methods or other nonrandomized designs that are more subject to potential sources of bias, such as volunteer bias, lead time bias, and length bias

8.0. BIAS IN SCREENING

1) Referral Bias, compliance (volunteer bias)
   - Volunteers or compliers are better educated and more health conscious – thus they have inherently better prognosis
2) Length Bias
   - Screening selectively identifies those with a long preclinical and clinical phase (i.e., those who would have a better prognosis regardless of the screening program)
3) Lead Time Bias
   - The apparently better survival that is observed for those screened is not because these patients are actually living longer, but instead because diagnosis is being made at an earlier point in the natural history of the disease
4) Over-diagnosis Bias (a misclassification bias)
   - Enthusiasm for a new screening program may result in a higher rate of false positives and give false impression of increased rates of diagnosis and detection

9.0. SCREENING TESTS



Lesson 8: DISEASE SURVEILLANCE

Objectives

At the end of the lesson the learner will be able to: -

1) Explain the purpose of surveillance
2) Describe the various epidemiological surveillance systems
3) Discuss the uses of epidemiological surveillance data

1.0. INTRODUCTION

- Disease surveillance is an information-based activity involving the collection, analysis and interpretation of large volumes of data originating from a variety of sources.
- Surveillance is
  i) The ongoing, systematic collection of data related to health events; their verification, analysis, interpretation, and the dissemination of information to those who need to know in order to reduce morbidity and mortality & to improve health (WHO)
  ii) Ongoing systematic collection, analysis, and interpretation of health data essential to the planning, implementation, & evaluation of public health practice, closely integrated with the timely dissemination of the data to those who need to know (CDC)
- Epidemiological information can be used to develop prevention strategies according to person (groups at risk), place and time (peaks at a particular season)
- In emergencies, epidemiology has three elements:
  i) Descriptive epidemiology
     - Determines the distribution of a disease among displaced populations
     - Describes the health problem, its frequency, who, where, and when
     - Examples
       a) Monitoring the health status of a population to detect cholera cases, such as, by age, sex, location, water source and duration of stay in a dispersed population or camps
       b) Conducting a nutritional survey to determine the prevalence of acute malnutrition among children under five
  ii) Analytical epidemiology
     - Compares those who are ill with those who are not in order to identify the risk of disease or protective factors (determinant of a disease)
     - Examines how the event (illness, death, malnutrition, injury) is caused (e.g. environmental and behavioural factors) and why it is continuing
     - Standard mathematical and statistical procedures are used
     - Example**:** Investigating an outbreak of an unknown disease in a displaced population settlement
  iii) Evaluation epidemiology
     - Examines the relevance, effectiveness and impact of different programme activities in relation to the health of the affected populations
     - Example **-** evaluating a malaria control programme for displaced populations

2.0. PURPOSE

1) Morbidity and mortality reporting
2) Monitoring disease trends
3) Describing natural history of diseases
4) Documenting of distribution and spread of disease.
5) Establish long term trends in disease occurrence.
6) Detect epidemics, outbreaks or new syndromes
7) Identify high risk groups or areas.
8) Facilitating planning of control and disease prevention.
9) Evaluation of intervention measures
10) Resource allocation and planning
11) Setting research priorities
12) Monitor implementation and effectiveness of programmes

3.0. SOURCES OF INFORMATION

1) Health facilities
2) Death and birth registers
3) Laboratories
4) The community self
5) Special search e.g. Chickenpox
6) Investigation of Outbreaks
7) Surveys

4.0. TYPE OF INFORMATION

The Person

- When available, demographic characteristics such as gender, age, race/ethnicity, occupation, education level, socio-economic status, sexual orientation, immunization status can reveal disease trends
- Example: looking at *Streptococcus pneumoniae*, a common cause of community-acquired pneumonia and bacterial meningitis, examining distribution of cases by race provides important information about burden of disease in different populations

Place

- Best to characterize cases by place of exposure rather than by place at which cases reported as the two may differ and place of exposure is more relevant to epidemiology of a disease;
- Example: travellers on a cruise ship exposed to a disease just prior to disembarking but become symptomatic and are diagnosed after return to various home locations; Example: person exposed to disease in small rural town but referred to tertiary care centre 100 miles away where disease is diagnosed and reported
- Data by geographic location can be presented in a table

- Inferential analysis can also be done using multilevel modelling, other statistical methods

Time

- Compare number of cases reported in time period of interest (weeks, months, years) to number of cases reported during similar historical period
- Usually a delay between disease onset and date when disease is reported, so preferable to use date of onset, if available, rather than date of report especially helpful for examining data not likely to have much short term variation example: there is limited variation in number of AIDS cases reported each month
- Provide valuable qualitative information; disease outbreaks often obvious from visual inspection of data, may not require a quantitative analysis

5.0. PROCESS

- Detect priority diseases
- Plan for surveillance
- Collect and compile information (tally daily/frequencies tables, prepare disease maps, investigate all cases e.g. cholera, total all cases daily, monthly)
- Analyse and interpret data
- Supervise and give feedback
- Monitoring and Evaluation

6.0. SURVEILLANCE SYSTEMS

6.1. Criteria of Evaluating Surveillance Systems

- Generally evaluated according to the following attributes such as simplicity, flexibility, acceptability, sensitivity, positive predictive value, representativeness (completeness, public/private), timeliness and adequately resourced (cost of training, travel, supplies, equipment and services)

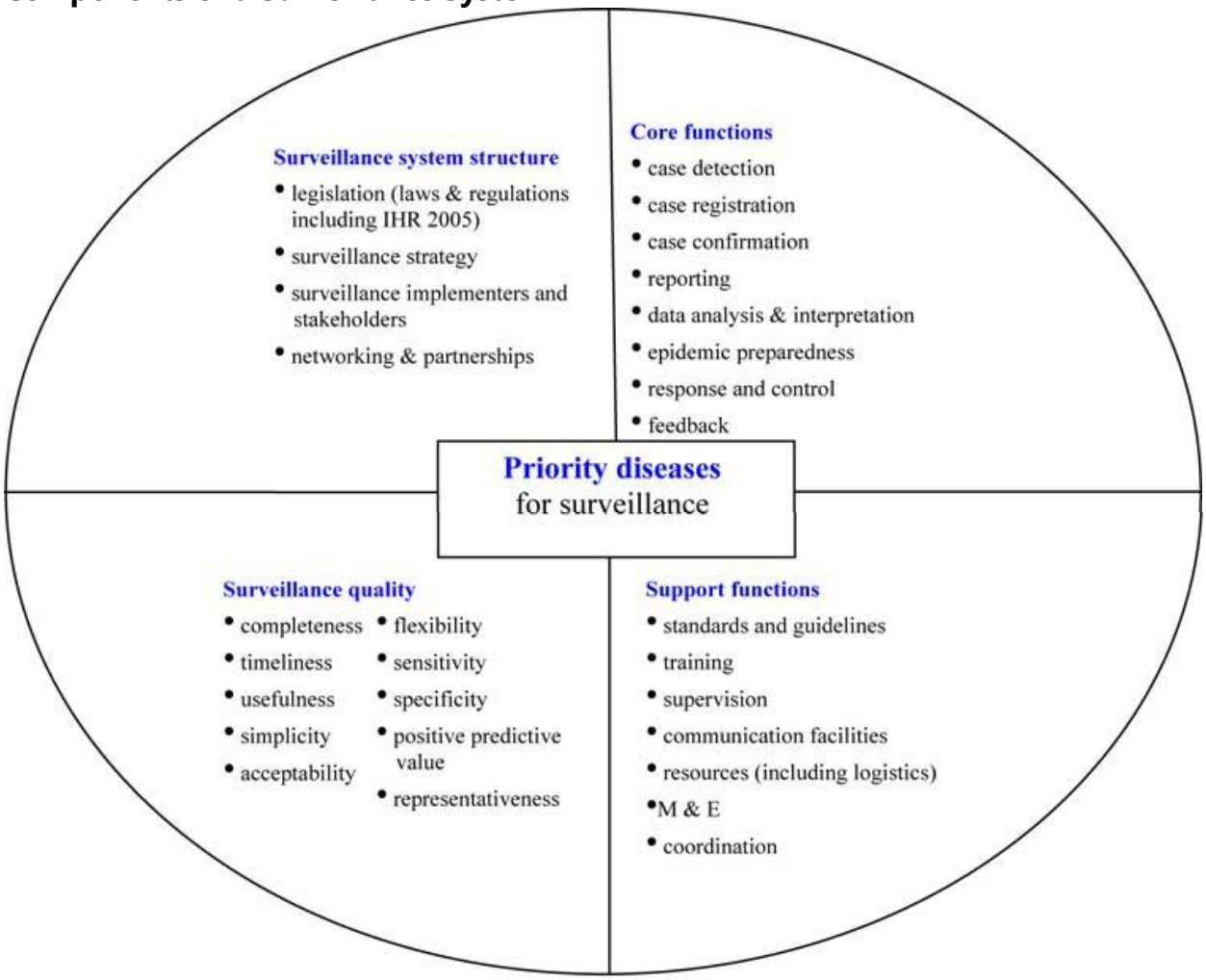
6.2. Surveillance Systems

- Comprises of five components namely: - defining target events/diseases; defining structure, core functions, support functions and measuring surveillance attributes
  1) First Component: Defining Target Events/Diseases
  - Prioritized events selected for surveillance
    - Based on disease burden, case fatality, potential to control, potential to cause outbreaks, potential to cross borders
  - Defining the objective of the control
    - Control - to reduce mortality and morbidity
    - Elimination - prevent outbreaks & limit indigenous transmission of infectious agent
    - Eradication - to subtract the infectious agent from the global earth.

Examples

Immediate Notification	Weekly or monthly notification
Polio, measles, cholera and food poisoning	

Components of a Surveillance System



2) Second Component - Defining the Structure

a. Legislation (laws and regulation)
b. Surveillance strategy
   - Indicator-based - specific indicators reaches specified thresholds, number of cases, weekly incidence and relative increase
   - Event-based - unofficial reports, rumours, health and non-health sources
c. Alerts – verification of outbreak or false alert
d. Implementers and stakeholders
e. Data and information required
f. Case definition – based on common understanding by professionals for targeted diseases/events and clinical, epidemiological, laboratory and other para-clinical findings/confirmed/probable/suspected cases

3) Third Component – Core Functions

Core Functions	Details
Case detection	By clinicians -data sources
Case registration	By data sources, in medical records or registers
Case confirmation	Laboratory tests to confirm cases or discard them
Case notification	By physicians to surveillance officers, on regular or irregular basis; immediately, weekly or monthly; written on paper or electronically
Case investigation	By surveillance officers to collect additional data and to collect samples
Data analysis	Transform data into information & displayed it by time, place & person.
Information Communication (feedback)	Feed-back to data sources; forward to higher level and dissemination to professionals and the public

4) Fourth Component – Support Functions
   - Standards and guidelines, training sessions, supervision, coordination, laboratory support, monitoring and evaluation, information and Communication technology

5) Fifth Component – Attributes of the Surveillance

Attribute	Examples
Quantitative attributes	Completeness; Timeliness, Positive predictive value; Sensitivity
Qualitative attributes	Usefulness; Simplicity, Acceptability; Flexibility; Representativeness

6.3. Key Elements of A surveillance System

1) Detection and notification of health events
2) Investigation and confirmation
3) Data collection
4) Analysis and interpretation of data
5) Feedback and dissemination of results
6) Response – action for prevention and control

6.4. Types of Surveillance

1) Passive surveillance

- Involves collection of data from existing unsolicited reports of the diseases(s)
- Health authorities do not stimulate reporting by reminding health care workers to report disease nor providing feedback to individual health workers.
- Most common type of surveillance in humanitarian emergencies & communicable diseases
- Health workers are trained on how to complete the surveillance forms (collected periodically)
- Data is often incomplete because there are few incentives for health workers to report

- Sources include reports such as notification forms filled by nurses of health department, solicited reports.
- It is good for conditions that have clear symptomatology e.g. measles.
- It is less costly than active surveillance

2) Active surveillance

- Involves active finding cases of the disease; for example, calling medical facilities (e.g., laboratories or emergency departments) or sending field teams to hospitals to extract information from hospital records
- Reporting frequency by individual health workers is monitored; health workers who consistently fail to report or complete the forms incorrectly are provided specific feedback to improve their performance. There may also be incentives provided for complete reporting.
- Requires substantially more time and resources and is therefore less commonly used in emergencies
- Often more complete than passive surveillance
- Used if an outbreak has begun or is suspected to keep close track of the number of cases
- Community health workers may be asked to do active case finding in the community in order to detect those patients who may not come to health facilities for treatment.
- Is used for disease which can be easily missed e.g. malaria with few asymptomatic infections. It remind health workers to be on the lookout for these conditions.
- It need laboratory confirmation because of its nonspecific clinical syndrome

3) Sentinel surveillance

- A sentinel surveillance system selects, either randomly or intentionally, a small group of health workers from whom to gather data
- The health workers then receive greater attention from health authorities than would be possible with universal surveillance
- Requires more time and resources, produce more detailed data on cases of illness because the health care workers have agreed to participate and may receive incentives
- Best type of surveillance if more intensive investigation of each case is necessary to collect the necessary data
- Use data from a few selected sites rather than data from all sites e.g. clinic
- Since it is from selected sites it provide accurate and more complete data on the sentinel population they are more quick available more reliable and less costly.
- Most of them are passive surveillance

4) Syndromic surveillance

- Is the collection and analysis of non-specific data from multiple data sources to detect a possible change or trend in the health of a population

- Involves collection and analysis of syndrome-related data, but has expanded to include almost any non-specific data from multiple sources that may indicate a potential biologic event has occurred
- Data sources may include: data from hospital emergency departments or other emergency encounters, physician office visits, over-the-counter pharmaceutical sales, and school absenteeism records

5) Outbreak Surveillance

- Undertaken to establish the source and cause of the outbreak with the aim of quick control with targeted interventions
- Also provide information on vaccine efficacy, age specific attack rate and case fatality

6.5. Epidemiological indicators

- Indicators are measures that reflect the state of a population in terms of health and socioeconomic status
- Indicators may be quantitative or qualitative in nature
- Quantitative indicators are easily calculated from numeric information such as the total number of people, the number of people according to age and sex etc. Examples of quantitative indicators include incidence, prevalence, morbidity and mortality rate
- Qualitative indicators that measure people's attitudes and knowledge are more difficult to measure. These indicators might be critical in explaining unexpected values of quantitative indicators. Social processes influencing health outcomes might also be elucidated by using qualitative indicators.
- Examples of qualitative indicators include
  1) Awareness of the value of immunization—low awareness may explain the high incidence of measles in a population living within five kilometres from a health facility
  2) Adherence to preventive interventions against HIV/AIDS—poor compliance from youths in preventive interventions (e.g.; A lack of understanding about the "Abstinence, Be Faithful, Use a Condom" (ABC) programme)) might explain the increasing prevalence of HIV/AIDS in a population
  3) Equity in distribution of resources—inequitable distribution of food might explain the increased mortality detected in a subgroup of a population
  4) Barriers to seeking treatment for malaria—barriers to seeking treatment such as unaffordable health services, might explain an increase in malaria-specific mortality.
- Qualitative indicators commonly used for assessing programme outcomes:
  1) Access - the proportion of the target population that can use the service or facility
  2) Coverage - the proportion of the target population that has received service
  3) Quality of services - the actual services received compared with the standards and guidelines
  4) Availability - amount of services compared with total target population. This should be based on minimum standard requirements

7.0. DATA PRESENTATION

1) Line graphs for displaying data by time
2) Maps for presenting data in geographic context
3) Graphical displays such as histograms, frequency polygons, box plots, scatter diagrams, bar charts, pie charts, or stem-and-leaf displays
4) Spot maps
5) Single/multivariable tables

8.0. USES OF SURVEILLANCE DATA

1) Priority setting and planning
2) Resources mobilization and allocation
3) Prediction and early detection of epidemics
4) Early and adequate detection and response
5) Monitoring and evaluation of intervention programmes
6) Identify high risk groups
7) Increase knowledge of vectors, dynamics of transmission
8) Establish long-term trends and pattern in disease occurrence
9) Serve a background which allows for detection on unusual pattern of a disease
10) Trigger disease control effort

9.0. CHALLENGES

1) Failure to report on time
2) Incomplete and late reporting
3) Inadequate data analysis
4) Failure to use available information to check trends
5) Poor feedback to health workers and communities
6) Duplication of efforts
7) Underutilization of surveillance information in decision making

10.0. PROBLEMS

1) Poor knowledge of conditions that are notifiable
2) Notification forms are too complicated and too many details are requested
3) Lack of feedback on reported cases
4) Lack of adequate training in surveillance
5) Understaffing (workers too busy to spend a lot time filling in the notification forms)
6) Incomplete filling of notification forms
7) Administrative delays of reporting hamper also timely and effective disease control
8) Lack of regular evaluation and validation of surveillance systems and processes
9) Active surveillance is not implemented widely because of its higher cost
10) Biased selection of sentinel sites (results are unrepresentative of the total population)