

# Generals

- Course Title: EPIDEMIOLOGY
- Module code: PubH-M2101
- Course code: PubH2102
- Credit Hours: 3 credit hrs (5 ECTS)

# Teaching-Learning Method & Materials

- Lectures
- Seminar presentations
- Group discussion
- Printed materials
- Non-projected materials (blackboard)
- Projected materials (LCD, Overhead projectors etc)

# Assessment Methods

- Continuous Assessment – 40%
- Class Participation
- Assignments
- Quizzes
- Tests/mid-examination
- and etc
- Final Written Examination – 60%

## Health science college, PHD, DMU, MPH in Epidemiology Course outline and Schedule for 3rd year environmental health students

| Time      |             | (April 30-May 19/2020)   |
|-----------|-------------|--|
| Monday    | 2:30-4:15   | Introduction to Epidemiology   |
|           | 4:30-6:15   | Communicable diseases Epidemiology   |
|           | 8:30-10:15  | Measures of Diseases occurrence  |
| Tuesday   | 2:30-4:15   | Test- 1  |
|           | 4:30-6:15   | Descriptive epidemiological studies<br>Case report/series<br>Cross-sectional studies<br>Ecological studies |
|           | 8:30-10:15  | Analytical studies: case-control   |
|           | 10:30-11:30 | Analytical studies: cohort   |
|           | 2:30-4:15   | Test- 2  |
| Wednesday | 4:30-6:15   | Analytical studies: Experimental   |
|           | 8:30-10:15  | Measures of association  |

# Health science college, PHD, DMU, MPH in Epidemiology

## Course outline and Schedule for 3<sup>rd</sup> year environmental health students

| Time     |             | ((April 30-May 19/2020))   |
|----------|-------------|--|
| Thursday | 2:30-6:15   | Evaluation of evidence<br>Bias, Chance, confounding<br>(Judgment of causality) |
|          | 8:30-10:15  | Screening  |
| Friday   | 2:30-4:15   | Epidemiological surveillance   |
|          | 4:30-6:15   | Outbreak investigation and management  |
|          | 8:30-10:15  | Independent study  |
|          | 10:30-11:30 | Independent study  |
| Monday   | AM          | Independent study  |
|          | 8:30-10:15  | Final Exam   |

# Principles of Epidemiology

*lecture notes*

By

Daniel Tarekegn

(BSC in PH, MPH in Epidemiology)

PHD, HSC, DMU

# 1. Introduction to Epidemiology

# Learning Objectives

**At the end of this session students will be able to Understand**

1. basic definitions of epidemiological terms
2. basic principles and concepts of epidemiology
3. Historical development of Epidemiology
4. Scope of Epidemiology
5. Purpose/use of Epidemiology
6. Types of Epidemiology
7. Basic Epidemiological assumptions
8. Basic features of Epidemiology

# Definitions

**Health:** A state of complete physical, mental and social well-being and not merely the absence of disease or infirmity (WHO,1948)

**Disease:** A physiological or psychological dysfunction

**Illness:** A subjective state of not being well

**Sickness:** A state of social dysfunction

# Definitions...

## Public health

The science & art of

*Preventing disease,*

*prolonging life,*

*promoting health & efficiency*

through organized *community effort* (Winslow, 1920)

# Definitions...

## Epidemiology

It is the *study* of *frequency*, *distribution*, and *determinants* of *diseases* and *other health-related* conditions in a human *population*.

**and**

the *application* of this study to the prevention of disease and promotion of health.

# Components of the definition

*1. Study*: Systematic collection, analysis and interpretation of data

Epidemiology involves collection, analysis and interpretation of health related data

*Epidemiology is a science*

# Components...

**2. Frequency:** the number of times an event occurs

Epidemiology studies the number of times a disease occurs

It answers the question *How many?*

*Epidemiology is a quantitative science*

## Components...

3. ***Distribution***: Distribution of an event by person, place and time

Epidemiology studies distribution of diseases

It answers the question *who, where and when?*

***Epidemiology describes health events***

## Components...

4. *Determinants*: Factors the presence/absence of which affect the occurrence and level of an event

Epidemiology studies what determines health events

It answers the question how and why?

*Epidemiology analyzes health events*

# Components...

## 5. *Diseases & other health related events*

Epidemiology is not only the study of diseases

The focus of Epidemiology are not only patients

It studies all health related conditions

*Epidemiology is a broader science*

# Components...

## 6. *Human population*

Epidemiology diagnoses and treats communities/  
populations

Clinical medicine diagnoses and treats patients

*Epidemiology is a basic science of public health*

# Components...

## *7. Application*

Epidemiological studies have direct and practical applications for prevention of diseases and promotion of health

Epidemiology is a science and practice

*Epidemiology is an applied science*

# History of Epidemiology

## Seven land marks in the history of Epidemiology

1. *Hippocrates* (460BC): Environment & human behaviors affects health
2. *John Graunt* (1662): Quantified births, deaths and diseases
3. *Lind* (1747): Scurvy could be treated with fresh fruit

## History...

4. ***William Farr*** (1839): Established application of vital statistics for the evaluation of health problems
5. ***John Snow*** (1854): tested a hypothesis on the origin of epidemic of cholera
6. ***Alexander Louis*** (1872): Systematized application of numerical thinking (quantitative reasoning)

## History...

7. *Bradford Hill* (1937): Suggested criteria for establishing causation

*Epidemiological thought emerged in 460 BC*

*Epidemiology flourished as a discipline in 1940s*

## Scope of Epidemiology

*Originally*, Epidemiology was concerned with investigation & management of *epidemics* of communicable diseases

*Lately*, Epidemiology was extended to endemic communicable diseases and non-communicable infectious *diseases*

*Recently*, Epidemiology can be applied to *all* diseases and other health related events

# Purpose of Epidemiology

The ultimate purpose of Epidemiology is prevention of diseases and promotion of health

## *How?*

1. Elucidation of natural history of diseases
2. Description of health status of population
3. Establishing determinants of diseases
4. Evaluation of intervention effectiveness

# Uses of Epidemiology

- ❑ To study history & project the future of health and diseases conditions in the population
- ❑ To diagnose the community for presence, nature & distribution of health & diseases in population ,,
- ❑ To study the working of health services ,,
- ❑ To estimate individual's chances & risks of disease

# Uses of Epidemiology

- ❑ To help complete the diseases clinical picture ,,
- ❑ To identify syndromes from the distribution of clinical phenomena among population
- ❑ To search for causes of health and diseases

# Types of Epidemiology

| <b>Component</b>           | <b>Objectives/ concerns</b>   | <b>Types</b>                  |
|----------------------------|---|-------------------------------|
| Distribution               | <ul style="list-style-type: none"> <li>➤ Frequency/distribution of health events by person, time and place ,,</li> <li>➤ Generate cause-effect hypothesis</li> </ul>            | Descriptive epidemiology      |
| Determinants               | <ul style="list-style-type: none"> <li>➤ Search for causes or risk factors</li> <li>➤ Response to a study of hypothesis</li> <li>➤ Use various epidemiologic methods</li> </ul> | Analytic epidemiology         |
| Health, disease, or injury | <ul style="list-style-type: none"> <li>➤ All health outcomes ,,</li> </ul>  | Disease-specific epidemiology |
| Application                | <ul style="list-style-type: none"> <li>➤ Monitoring and evaluation</li> <li>➤ Planning and policy making</li> <li>➤ Prevention and promotion of health</li> </ul>               | Applied epidemiology          |

# Types of Epidemiology

*But, the Two* major categories of Epidemiology includes;

## *1. Descriptive Epidemiology*

Defines *frequency* and *distribution* of diseases and other health related events by time, place and person

Answers the four major questions: *how many, who, where, and when?*

*Generate hypothesis about possible risk factors*

# Characteristics of Descriptive Epidemiology

*Three* major characteristics of descriptive Epidemiology includes; time, place and person

## *1. Time*

Information organized by time:

1. Shows the trend/ pattern of the disease over time
2. Establishes the usual occurrence of diseases (Endemics)
3. Helps to identify excess/un usual occurrence (epidemics)
4. Uses to predict seasonal and secular trends (projection)

# Characteristics....

## *2. Place*

Information organized by place:

Shows the geographic pattern/variation of the diseases..(urban/rural, birthplace, school, district...)

provides clue in identifying factors of disease occurrence either in the host/environment

Provides insight into the geographic problem extent

# Characteristics....

## *3. Person*

Information organized by personal characteristics:

Help to identify some modifiable factors in order to prevent or control the disease

Data includes:

*Inherent characteristics----(age, sex, race...)*

*Biologic characteristics----(immune status...)*

*Acquired characteristics---(Marital status...)*

*Activities------(occupation, smoking...)*

*Living conditions------(Socioeconomic status...)*

# Types...

## 2. *Analytic Epidemiology*

Analyses *determinants* of health problems

Test a hypothesis of cause-effect relationship

Its key feature is a comparison group

Answers two other major questions: *how?* and *why?*

Generally, Epidemiology answers six major questions:  
*how many, who, where, when, how and why?*

# Basic Epidemiological assumptions

1. Human diseases doesn't occur at *random* or by chance
2. Human diseases have *causal* and *preventive* factors that can be identified by systematic investigation

# Basic features of Epidemiology

1. Studies are conducted on human population
2. It examines patterns of events in people
3. Can establish cause-effect relationship without the knowledge of biological mechanism
4. It covers a wide range of conditions
5. It is an advancing applied science

# **2. Communicable disease Epidemiology**

# Learning Objectives

**At the end of this session students will be able to Understand**

1. basic definitions of Communicable diseases terms
2. basic disease causation theories and models
3. Natural history of diseases
4. Levels of diseases prevention
5. Infectious diseases process
6. Time course of infectious diseases
7. Infection transmission probability

# Definitions

## Communicable diseases:

An illness due to a specific infectious or biological *agent* or its *toxic* products capable of being directly or indirectly *transmitted* from *man to man*, from *animal to man*, from *animal to animal*, or From *environment to man* (*through air, water, food, etc*)

# Definitions

## **Infection:**

the entry and development or multiplication of an infectious agent in the body of man or animals.  
An infection does not always cause illness.

## **Contamination:**

the presence of an infectious agent on a body surface, on or in clothes, beddings, toys, surgical instruments or dressings, or other articles or substances including water and food.

# Definitions

## **Infestation:**

The lodgment, development and reproduction of arthropods on the surface of the body or in the clothing, e.g. lice, itch mite.

This term could also used to describe the invasion of the gut by parasitic worms, e.g. ascariasis.

## **Contagious disease:**

An infectious disease that transmitted through contact. E.g. scabies, trachoma, STD, & leprosy.

# Definitions

## Vector of infection:

An insect or any living carrier that transports an infectious agent from an infected individual

or

its wastes to a susceptible individual or food or immediate surroundings.

Both biological and mechanical transmissions encountered.

# Definitions

## Nosocomial infections:

Nosocomial (hospital acquired) infection is an infection originating in a patient while in a hospital or another health care facility.

It has to be a new disorder unrelated to the patient's primary condition.

Example include infection of surgical wounds, hepatitis B and urinary tract infections.

# Definitions

## Opportunistic infection:

This infection takes the opportunity provided by a defect in host defense (immunity) to infect the host and thus cause disease.

For example, opportunistic infections are very common in AIDS.

Organisms like Herpes simplex, Herpes Varicella cytomegalovirus, M. tuberculosis.....

# Definitions

## Zoonosis, Epizootic and Enzootic :

**Zoonosis** is an infection that is transmissible under natural conditions from vertebrate animals to man, e.g. includes rabies, plague, bovine tuberculosis.

An **epizootic** is an outbreak (epidemic) of disease in an animal population, e.g. rift valley fever.

An **Enzootic** is an endemic occurring in animals, e.g. bovine TB.

# Definitions

## Eradiation and Elimination:

*Elimination* is termination of all transmission of infection by the extermination of the infectious agent through surveillance and containment.

*Eradiation* is an absolute process, an “all or none” phenomenon, restricted to termination of infection from the whole world.

The term *elimination* sometimes used to describe *eradiation* of a disease from a large geographic region.

Diseases, which are amenable to *elimination*, in the meantime are polio, measles, and diphtheria.

# Importance of Studying Communicable Diseases Epidemiology

The following points are some of the reasons indicating importance of studying communicable diseases epidemiology

1. Changes of the pattern of infectious diseases
2. Discovery of new infections
3. The high possibility some chronic diseases have an infective origin.
4. Surveillance of infectious disease
5. Identification of source of outbreaks
6. Studies of routes of transmission & natural history of infections
7. Identification of new interventions

# Disease causation

## The cause of a disease

An event, a condition or a characteristic that comes before the disease and without which the disease wouldn't occur

# Theories of disease causality

What causes a disease?

Nineteenth century theories

1. Contagion theory
2. Supernatural theory
3. Personal behavior theory
4. Miasma theory

# Contagion theory

- ❑ This theory was common at the beginning of the 19th century.
- ❑ Most official disease prevention activities were based on the hypothesis that illness is contagious.

## Contagion theory.....

It required:

Keeping sick people away from well people.

The institution of quarantine of ships

Setting up military cordons around infected towns

Isolation of households if they were infected, and

Fumigating /washing bedding & clothing of the sick.

## Problems of Contagion theory.....

Problems confounded the acceptance of this theory were:

There were too many instances where people become ill regardless of their isolation from human contact

*and*

Too many others where brave souls nursed the dying & carried their bodies to the graveyard yet remained well.

## Supernatural theory

Proponents of this theory argue that supernatural forces cause disease.

Disease prevention measures based on this theory were important to the religious people.

The view among them was that disease is a punishment for transgression of God's laws.

## Supernatural theory.....

Because epidemic took a great toll on the poor than the rich,

the healthier rich can employ the supernatural theory as a justification for berating for the poor for sinful behavior;

i.e. presumed idleness, intemperance & uncleanness.

This theory expressed a political philosophy.

## Supernatural theory.....

People could not advocate the belief that sin causes disease without, at the same time, implicitly supporting the idea that government need to redress poverty.

This theory held that disease results from wrong personal behavior to God.

## Personal behavior theory

It gave responsibility to individuals to control their own lives.

People caused their own disease by living fully unhealthy.

It was democratic & ante authoritarian in intent

Believes on the idea that disease was not tied up with the mysterious ways of God.

## Personal behavior theory.....

Hence, improper diet, lack of exercise, poor hygiene and emotional tension become the focus of preventive actions.

This theory does not blame the poor for the illness and in many aspects; it was homage to middle-class life.

## Miasma theory

This theory argues that disease is caused by the odor of decaying of organic materials.

It dates back to the Hippocratic idea that disease is related to climate.

It contrasted sharply from the other three theories since it conceptually separated the source of the disease from the victim of the disease.

# Theories...

## Twentieth century theories

1. Germ theory
2. Lifestyle theory
3. Environmental theory
4. Multi-causal theory

# The Germ Theory

It held the notion that microorganisms cause diseases and it is possible to control diseases using antibiotics and vaccines.

There was criticism on this theory by Thomas Mckeown that stated as the incidence of all major infectious diseases begun to fall several decades before the introduction of vaccines and antibiotics.

Thus rising of living standards was responsible for the reduction of disease not the discovery of antibiotics and vaccines.

# The Life Style Theory

This holds that unhealthy lifestyles are causes for diseases.

This hypothesis blames stress, lack of exercise, the use of alcohol and tobacco improper nutrition for most chronic diseases.

This theory rejects the notion central to the classic germ theory, that a single disease has a single etiology.

# The Life Style Theory.....

Emphasize the interrelatedness of many variables in disease causality, principally those under the control of the individual.

Resembles the germ theory, for it conceives of disease as an individual event

The difference is that prevention demand personal behavior change than physicians' ministrations.

Criticized b/c of lifestyle change needs overall social change

# The Environmental Theory

Environmental theory explains that significant number of chronic disease are caused by toxins in the environment and

it implies that disease prevention, instead of requiring medical treatments or personal hygiene, demands change in the industrial production.

# The Environmental Theory.....

The first aspect of the environmental hypothesis is occupational hazards,

The second concentrates on toxic substances in the air, water and soil (advocates of this theory places particular emphasis on radioactivity),

*and*

The third aspect focus on synthetic additives to foods “organic foods”.

# The Multi Causal Theory

It is also called the web of disease causation.

The theory express that there are multiple factors for a cause of a single disease entity.

But it is incapable of directing a truly effective disease prevention policy as the theories it replaces.

# The Multi Causal Theory.....

Its shortcomings are it gives few clues about how to prevent disease,

the actual prevention policies it implies are inefficient in many ways,

*and*

there is a gap between what it promises and what epidemiologist's deliver.

# Necessary Vs Sufficient Cause

**Necessary:** the disease will not occur without the presence of the factor

Example: Mycobacterium TB for TB

**Sufficient:** the presence of the factor always result in disease

Example: Rabies virus for rabies

# Etiology of a disease

The sum of all factors contribution to the occurrence of a disease

Agent factors + Host factors + Environmental factors = Etiology of a disease

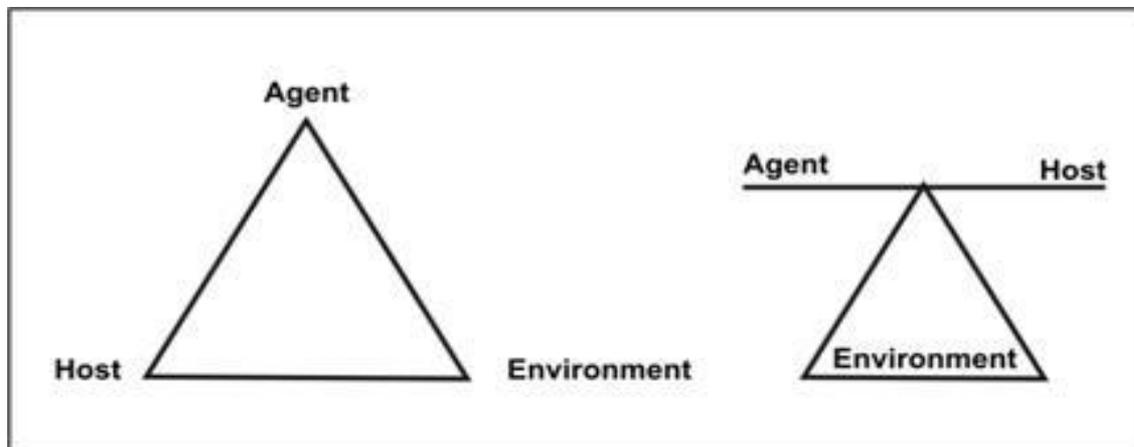
# Disease causation models

How do diseases develop?

Three best known models

## *1. Epidemiological triangle*

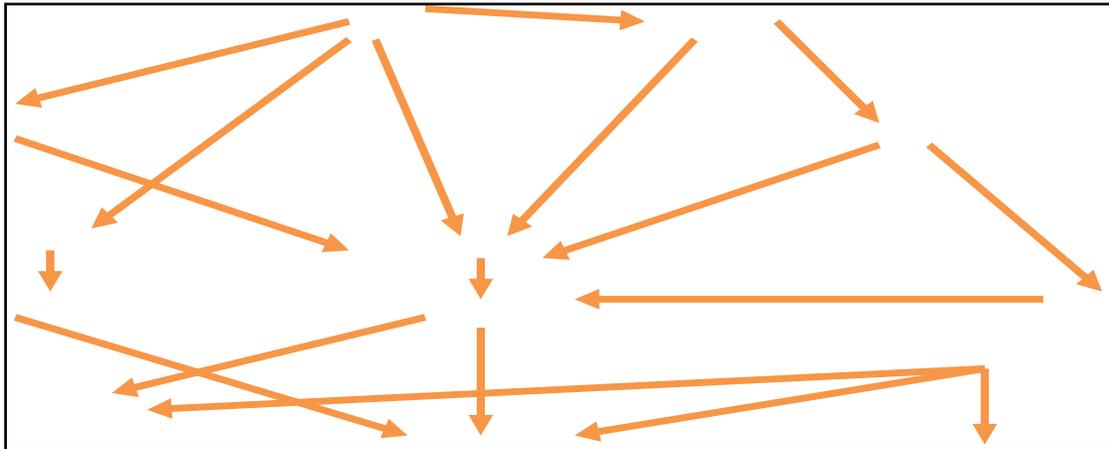
The interaction of an agent and host in an appropriate environment results in disease



# Disease models...

## 2. *Web of causation*

Complex interaction of factors results in disease



Best model for chronic communicable/NCDs.

# Disease models...

## 3. *Wheel model*

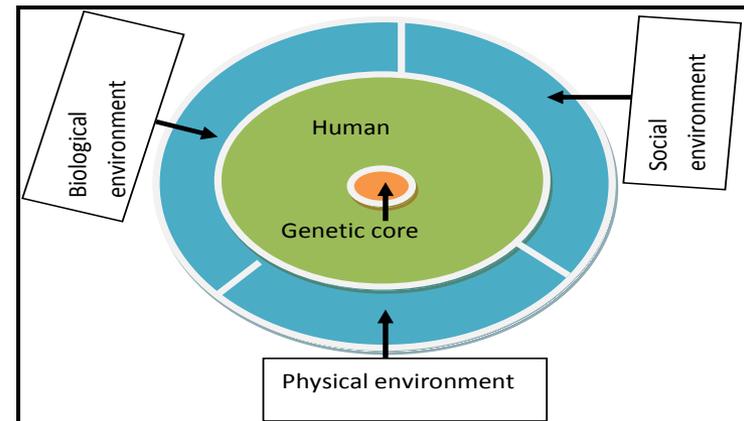
The hub (host) having a genetic makeup as its core, surrounded by an environment schematically divided in to biological, physical and social environment.

Best for a gene linked diseases

E.g., sickle cell anemia

Type-II –DM

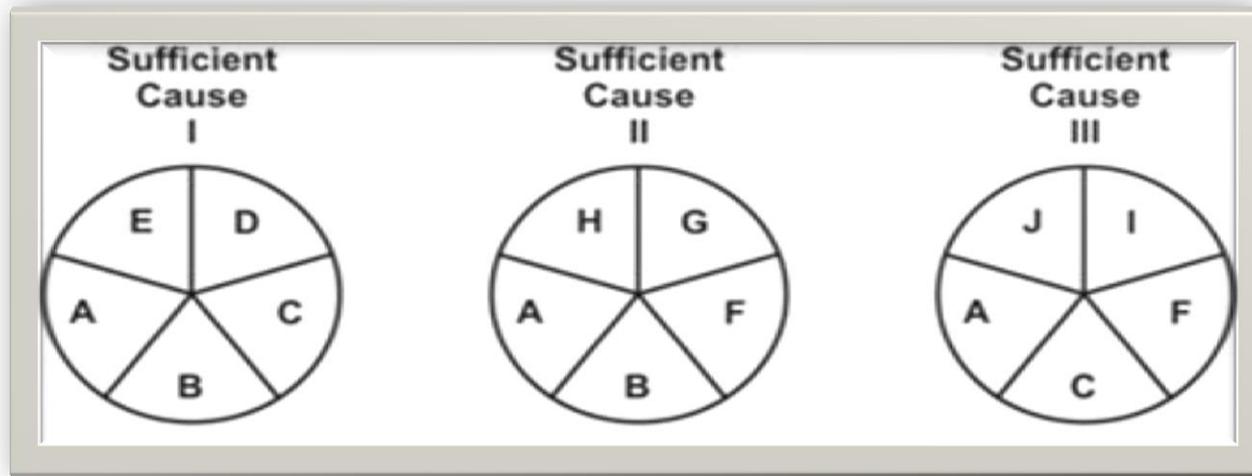
HIV infection



## Disease models...

### 4. Component causes and causal pies model

Because the agent-host-environment model did not work well for many non-infectious diseases, several other models that attempt to account for the multi-factorial nature of causation proposed.



# Natural history of disease

The progression of disease process in an individual overtime in the absence of intervention

Four stages in the natural history of a disease

## *1. Stage of susceptibility*

Presence of risk factors

No disease

## Natural history...

### *2. Stage of sub-clinical disease*

Presence of pathogenic changes (biological onset)

No disease manifestations

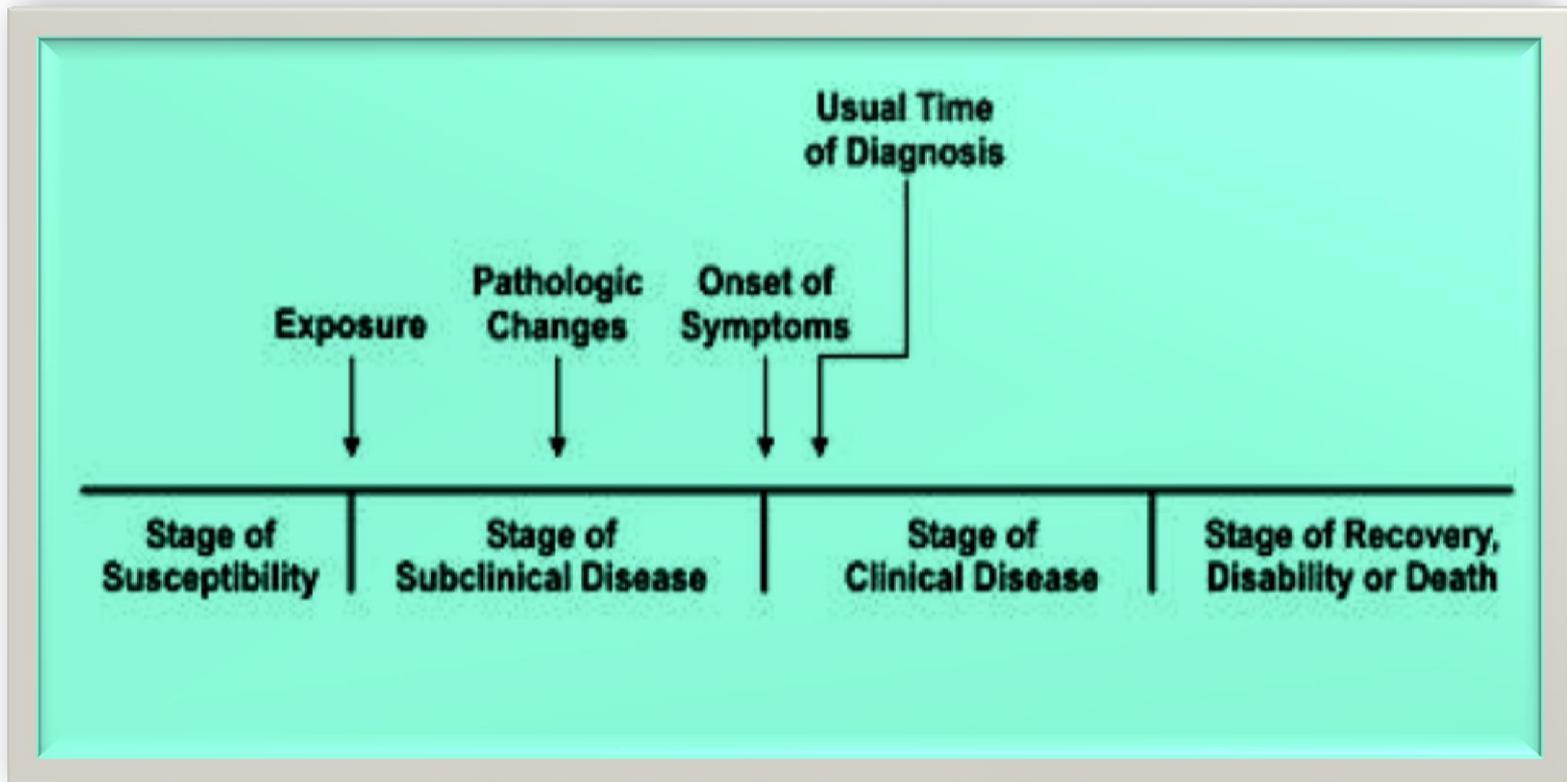
### *3. Stage of clinical disease*

Presence of sign and symptoms (clinical onset)

### *4. Stage of recovery, disability, or death*

# Natural history...

*Stages:*



# Levels of disease prevention

Three major levels of disease prevention

## *1. Primary prevention*

Targeted at healthy people

Objectives are Promotion of health

Prevention of exposure and

Prevention of disease

# Levels of disease...

## 2. *Secondary prevention*

Targeted at sick individuals

Objective is to stop or slow the progression of disease and to prevent or limit permanent damage through early detection & treatment

# Levels of disease...

## *3. Tertiary prevention*

Targeted at people with chronic diseases & disabilities that can't be cured

Objective is to prevent further disability or death and to limit impacts of disability through rehabilitation

# Levels of disease...

Table 2.1 Levels of diseases prevention

| Level      | Stage of disease  | Objective  | Target   |
|------------|---|--|--|
| primordial | Existence of underlying condition leading to causation. E.g. Smoking, environmental pollution | The aim is to avoid the emergence & establishment of the social, economic & cultural patterns of living that are known to contribute to an elevated risk of disease. | Total population<br>selected groups &<br>healthy individuals |
| primary    | Specific causal factors exist. E.g. immunization, measles, polio                              | The causative agent exists but the aim is to prevent the development of disease.   | Total population<br>Selected groups &<br>healthy individuals |
| secondary  | Early stage of disease. E.g. Early detection & Rx of cases of TB & STI                        | The aim is to cure patients & prevent the development of advanced disease  | patients   |
| Tertiary   | Late stage of disease (Rx & rehabilitation) e.g. leprosy                                      | The aim is to prevent sever disability & death.  | Patients   |

# Infectious disease process

There are six components of the infectious disease process constituting chain of disease transmission

1. The agent
2. Its reservoir
3. Its portal of exit
4. Its mode of transmission
5. Its portal of entry
6. Susceptible host

# The agent

It includes viral particles to complex multicellular organisms.

Possible outcomes of exposure to an infectious agent

**Infection:** invasion & multiplication in the host

*Infectivity:* the proportion of exposed who becomes infected

Infection rate =  $\frac{\text{Infected}}{\text{exposed}}$

**Disease:** A clinically apparent infection

# The agent...

Host-agent interaction is characterized by infectivity, Pathogenicity, virulence or immunogenicity.

## A. Infectivity

1. Includes a time range from exposure to infection.
2. It is the ability to cause infection in an exposed susceptible host, Measured by the infection rate (infectiousness).
3. It is the proportion of exposed persons who become infected.
4. The infectivity(infectious ) Rate (IR) is given by:

$$IR = \left( \frac{\text{Number of infected individuals}}{\text{Number of susceptible \& exposed}} \right) \times 100$$

# The agent....

## B. Pathogenicity

Includes a time range from infection to disease.

It is the ability to cause disease in a susceptible host.  
measured by the clinical to subclinical ratio.

The proportion of infected persons who develop  
clinical disease.

*Pathogenicity = Number of clinical cases: Number of subclinical cases*

# The agent...

## C. Virulence

- Includes a time range from disease to disease outcome
- It is the ability to cause severe disease in a susceptible host
- The proportion of clinical disease become severely ill or die
- Measured by case fatality and hospitalization rate

$$\text{Case Fatality Rate} = \left( \frac{\text{Number of fatal cases}}{\text{Total number of cases}} \right) \times 100$$

$$\text{Hospitalization Rate} = \left( \frac{\text{Number of hospitalized cases}}{\text{Total number of cases}} \right) \times 100$$

# The agent...

## D. Immunogenicity

It is the ability of an infection to produce specific immunity

$$\text{Immunogenicity} = \left( \frac{\text{Number of infected individuals who developed specific immunity to the infection}}{\text{Total number of infected individuals}} \right) \times 100$$

A limitation to the measures used as indicator of...

- infectiousness,
- Pathogenicity and
- virulence of infectious agent

...is they does not reflect the intrinsic attributes of the agent.

# Factors which influence the development of disease

Strain of the agent

Dose of the agent

Route of infection

Host age, nutritional status, immune status

Influence of treatment

Influence of season

# Reservoir Vs Carrier

## *Reservoir*

An organism or habitat in which an infectious agent normally lives, transforms, develops and/or multiplies (Human, Animal, Environment)

## *Carrier*

A person who doesn't have apparent clinical disease, but is a potential source of infection to other people

# Types of carriers

***1. Incubatory carriers:*** transmits the disease during incubation period

Example: Measles, mumps

***2. Convalescent carriers:*** transmits the disease during convalescent period

Example: Typhoid fever

## Types of carriers...

3. *Asymptomatic carriers*: transmitting the disease without showing manifestations

Example: polio, Amoebiasis

4. *Chronic Carriers*: transmitting the disease for long time/indefinite transmission

Example: Viral hepatitis, typhoid fever

## Importance of carriers

1. ***Number***- carriers may outnumber cases
2. ***Difficulty in recognition***- carriers don't know that they are infected
3. ***Mobility***- carriers are mobile, cases are restricted
4. ***Chronicity***- carriers re-introduce infection and contribute to endemicity

# Effect of carriers on disease transmission

- *Ice-berg effect* in temperate zone
- *Hippopotamus effect* in tropical zone

These are the fact that carriers constitute a hidden reservoir of infection and that they may outnumber actual cases

# Portal of exit

It is the way the infectious agent leaves the reservoir

Possible portal of exit include all body secretions and discharges

---

## Examples of portal of exits

---

Saliva, Mucus, Tears

Pus

Breast milk

Exudates from wounds

Vaginal, cervical secretions

Excretions ( feces & urine)

Semen, Urethral secretions

Blood and tissue

---

# Modes of disease transmission

## 1. Direct transmission

***Direct contact:*** physical contact with body part of infected person: Touching, kissing, biting, sex

Example: HIV

***Direct projection:*** projection of saliva droplets while coughing, sneezing, spitting, talking, singing etc

Example: Common cold

***Transplacental:*** Transmission from mother to fetus through the placenta

Example: Syphilis

## Modes of disease...

### 2. Indirect transmission

*Vehicle-borne*: transmission through inanimate objects/non-living substances

E.g. HIV by needles

*Air-borne*: transmission by dust or droplet nuclei

E.g. TB, Pneumonia.

## Modes of disease...

*Vector-borne*: infectious agent is conveyed by an arthropod to host literatures brief

*Biological*: there is multiplication and/or development in the vector

*Salivarian*: Injects infected saliva e.g mosquito

*Stercorarian*: infects by infected feaces e.g louse

*Mechanical*: simple transfer without biological stages in the vector e.g flies

# Importance of mode of transmission

A disease often has several modes of transmission

It is important to distinguish between the predominant mode of transmission and those of secondary importance

Identifying primary and secondary modes of transmission is important to identify most effective prevention and control measures

# Portal of entry

The site where an infectious agent enters a susceptible host

The manner of entry of an infectious agent in to a host is one of the factors which determine whether or not the agent will succeed in establishing an infection.

Example, the required portal of entry for the agent of tetanus to cause a disease is the site of an injury.

# Susceptible Host

The susceptible human host is the final link in the infectious process

Host susceptibility can be seen at the individual level and at the community level.

# Herd immunity

It is host resistance at a population level

It is defined as the resistance of a community (group) to invasion and spread of an infectious agent, based on immunity of high proportions of individuals in the community.

It has implications on vaccination programs

# Conditions under which herd immunity best functions

1. Single reservoir
2. Direct transmission
3. Total immunity
4. No carrier state
5. Uniform distribution of immunes
6. No overcrowding

Seldom all fulfilled

## Time course of an infectious disease

*Pre-patent period*: between biological onset and first shedding

*Incubation period*: between biological onset and clinical onset

*Communicable period*: time during which agent is being shed

## Time course of...

*Latent period*: between recovery and relapse in clinical disease

*Convalescent period*: between recovery and time when shedding stops

*Generation period*: between exposure/infection and maximum communicability of exposed host

# Application of time periods

1. *Pre-patent period*: When should we investigate?
2. *Incubation period*: When was time of exposure?
3. *Communicable period*: When should we take care of infectiousness?
4. *Latent period*: When would relapse occur?
5. *Convalescent period*: When after recovery an individual becomes non-infectious?
6. *Generation time*: When is the maximum risk for contacts?

# **3. Measures of frequency and disease occurrence**

# Learning Objectives

**When you have completed this lesson, you will be able to do the following:**

1. Differentiate the four quantitative descriptors (frequency) (Number, Ratios, proportion, and Rates)
2. Calculate and use measures of diseases occurrence
3. Calculate prevalence & incidence of diseases
4. Differentiate point prevalence, period prevalence, cumulative incidence, and incidence density
5. Calculate and use mortality measures
6. Differentiate crude, specific, & adjusted mortality measures

# What are epidemiologic study outcomes

Sometimes referred to as the six 'D's of Epidemiology listed as follows:

- ✓ Death
- ✓ Disease/Illness - Physical signs, laboratory abnormalities
- ✓ Discomfort - Symptoms ( pain, nausea, dyspnea, itching)
- ✓ Disability - Impaired ability to do usual activities
- ✓ Dissatisfaction - Emotional reaction (e.g., sadness, anger)
- ✓ Destitution - Poverty, unemployment, socio-economic status

The first two, death, and disease are the most commonly used outcomes in epidemiologic study.

# What are measures of disease occurrence?

These are measurements of the  
frequency/magnitude/amount  
of disease  
in populations

# How do we measure diseases?

Four *quantitative* descriptors

1. Numbers/counts
2. Ratios
3. Proportions
4. Rates

Are also called frequency measures

# Descriptors

- Provide useful information about:
  - ✓ the probability of occurrence of health events,  
and
  - ✓ population at high risk of acquiring the disease.
  - ✓ They are also important in designing an appropriate public health intervention.

# Descriptors

## *a. Numbers/ counts*

Use of actual number of events

As noted, one of the basic tasks in public health is identifying and counting cases.

Usually derived from case reports submitted by health-care workers and laboratories to the health department, or epidemiological studies.

# Descriptors

## *a. Numbers:*

Calculating the magnitude of disease occurrence with a count is simple and useful for certain purposes, such as allocating health resources.

However, simple counts do not provide all the information a health department needs.

For some purposes, the counts must be put into context, based on the population in which they arose.

# Descriptors

## *a. Numbers:*

Hence, for such purposes, it is more helpful to have a denominator under the count that indicates the size of the study population.

The remaining four measures address this issue.

# Descriptors

## *b. Ratios:*

Quantifies the magnitude of one occurrence  $X$ , in relation to another event  $Y$  as  $X/Y$

The relative size of two quantities (relative importance).

No specific relationship is necessary between the numerator and denominator

Numerator not necessarily included in the denominator

# Descriptors

## *b. Ratios:*

a special fraction in which the numerator includes only individuals who meeting one criterion

*(e.g. the case definition)*

*and*

the denominator includes only individuals in the study population who meet another criterion

*(e.g. do not meet the case definition but are at risk).*

# Descriptors

## *b. Ratios:*

Not dependent upon time.

less frequently used *as measure of disease frequency*

Example-1: 1 case of colon cancer for every 1 case of breast cancer.

Example-2: 2 female cases of major depression to 1 male case of major depression.

Example-3: Ratio of TB cases in community A to B is 1:10

# Descriptors

## *b. Ratios:*

After either the numerator divided by the denominator, the numerator or denominator is usually set one,

The result is often expressed as the result "to one" or written as the result “ :1 ”

Note that in certain ratios, the numerator and denominator are different categories of the same variable, such as:  
*(males and females, or 20–29 years and 30–39 years of age)*

# Descriptors

## *b. Ratios:*

In other ratios, the numerator and denominator are completely different variables, such as:

the number of hospitals in a city and the size of the population living in that city.

### *Method for calculating a ratio*

$$\text{Ratio} = \frac{(\text{Number or rate of events, items, persons, etc. in one group})}{(\text{Number or rate of events, items, persons, etc. in another group})} \times 10^n$$

# Descriptors

## *EXAMPLE: Calculating a Ratio*

### *1. Different Categories of Same Variable*

Between 1971 and 1975, as part of the National Health and Nutrition Examination Survey (NHANES), 7,381 persons ages 40–77 years were enrolled in a follow-up study.

At the time of enrollment, each study participant was classified as having or not having diabetes.

During 1982–1984, enrollees were documented either to have died or were still alive. The results summarized as follows.

# Descriptors

## *EXAMPLE: Calculating a Ratio*

### *1. Different Categories of Same Variable*

*Table-2.1 Death from Diabetics and Non Diabetics by sex in a follow-up study.*

| <b>Events</b>             | <b>Original Enrollment<br/>(1971–1975)</b> | <b>Dead at Follow-Up<br/>(1982–1984)</b> |
|---------------------------|--|--|
| <b>Diabetic men</b>       | 189  | 100                                      |
| <b>Non-diabetic men</b>   | 3,151                                      | 811                                      |
| <b>Diabetic women</b>     | 218  | 72                                       |
| <b>Non-diabetic women</b> | 3,823                                      | 511                                      |

*1. Calculate the ratio of non-diabetic to diabetic men?*

*2. Calculate the ratio of non-diabetic to diabetic women?*

# Descriptors

## 2. Calculating Ratios for Different Variables

**Example-A:** ANRS of Ethiopia with 19,626,000 persons has 19 Hospital. Calculate the ratio of Hospitals per person?

1.  $19 / 19,626,000 \times 10^n = 0.0000009681$  Hospitals per person.
2. For easily understood the result, set  $10^n = 10^7 = 10,000,000$ .
3. The ratio becomes:  $0.0000009681 \times 10,000,000 = 9.681$  Hospitals per 10,000,000 persons.
4. You could also divide each value by 9.681, and express this ratio as 1 Hospital for every 1,032,947 persons.

(Source: *HHRI 2005 EFY, Health facility to population Ratio by Region*)

# Descriptors

## 2. *Calculating Ratios for Different Variables*

### **Example-B:**

Ethiopia's IMR estimated for 2013(EFY) was 52/1,000 LBs

Kenya's IMR estimated for 2013(EFY) was 54/1,000 LBs.

Calculate the ratio of the IMR in Ethiopia to that in Kenya.

$$52 / 54 \times 1 = 1:1.04.$$

Thus, Kenya's IMR was estimated to be 1.04 times as high as Ethiopia's IMR for 2013 EFY.

*(Source: HHRIs, 2005 EFY, Population Health, & Environment Data & Estimates for the Countries & Regions of the World EFY 2013).*

# Descriptors

## *c. Proportions:*

A ratio in which the numerator is included in the denominator  
It is comparing of a part to the whole population in which the occurrence takes place.

$$A/(A+B);$$

A fraction in which the numerator (A) includes only individuals who meet the case definition and

The denominator totals (A+B) the numbers of individuals who meet the case definition plus those in the study population who do not meet the case definition and are at risk.

# Descriptors

## *c. Proportions:*

Not dependent upon time, expressed as a fraction, decimal, or a percentage.

Indicates the fraction of the population that affected by the disease or condition.

Common descriptive measures, linked to estimating risk.

Its result ranges between 0 and 1 or 0-100%.

Percentage=proportion x100.

# Descriptors

## *c. Proportions:*

Example-1: 30% of persons over 50 years of age have screened for colon cancer.

Take care of those who should be included for calculation

Calculating the proportion of women with cervical cancer requires a special consideration. Cervical cancer only occurs in women with a cervix.

A woman who has had a complete hysterectomy is no longer at risk for developing cervical cancer.

# Descriptors

## *c. Proportions:*

Method for calculating a proportion

$$\text{Proportion} = \frac{\text{Number of persons or events with a particular characteristics}}{\text{Total number of persons or events, of which the numerator is a subset}} \times 10^n$$

For a proportion,  $10^n$  is usually 100 (or  $n = 2$ ) and is often expressed as a percentage.

The statements "one fifth of the residents became ill" and "twenty percent of the residents became ill" are equivalent.

# Descriptors

## *d. Rates:*

a proportion with time element.

It measure the occurrence of an event overtime

In epidemiology, a rate is a measure of the frequency with which an event occurs in a defined population over a specified period.

# Descriptors

## *d. Rates:*

Put disease frequency in the perspective of the size of the population,

Particularly useful for comparing disease frequency

In different locations, at different times, or

Among different groups of persons with potentially different sized populations;

It is a measure of risk.

e.g U5 measles cases in 2000 per U5 population in 2000

## Which community is more affected?

Community A has 100 cases of disease X  
and

Community B has 1000 cases of disease X,

which community is more affected?

Answer --- Can't be known. Why?

Number could not help to compare. Why? (Use Rate)

## When we call..

- When we call a measure *a ratio*, we mean a non-proportional ratio
- When we call a measure *a proportion*, we mean a proportional ratio that doesn't measure an event overtime
- When we call *a rate*, we mean a proportional ratio that does measure an event in a population overtime

# Types of rates

1. ***Crude rates***: Apply to the total population in a given area
2. ***Specific rates***: Apply to specific subgroups in the population (age, sex etc) or specific diseases
3. ***Standardized rates***: used to permit comparisons of rates in population which differ in structure (e.g age structure)

Two methods of standardization: Direct, indirect

# Types of Epidemiologic measures

**Table 2.2 Epidemiologic Measures Categorized as Ratio, Proportion, or Rate**

| <b>Condition</b>               | <b>Ratio</b>   | <b>Proportion</b>   | <b>Rate</b>  |
|--------------------------------|--|---|--|
| <b>Morbidity<br/>(Disease)</b> | Risk ratio<br>(Relative risk)<br>Rate ratio<br>Odds ratio<br>Period prevalence | Attack rate<br>(Incidence proportion)<br>Secondary attack rate<br>Point prevalence<br>Attributable proportion | <b>Person-time incidence rate</b>  |
| <b>Mortality<br/>(Death)</b>   | Death-to-case ratio<br>Maternal mortality ratio                                | Proportionate mortality   | <b>Crude mortality rate<br/>Case-fatality rate<br/>Cause-specific mortality rate<br/>Age-specific mortality rate<br/>Maternal mortality rate<br/>Infant mortality rate</b> |
| <b>Nativity<br/>(Birth)</b>    |  |   | <b>Crude birth rate<br/>Crude fertility rate</b>   |

# Morbidity rates

Morbidity rates are rates that are used to quantify the magnitude/frequency of diseases

Two common morbidity rates are:

1. Incidence rates

(Cumulative incidence, incidence density)

2. Prevalence

(Period prevalence, point prevalence)

# Morbidity rates

**Table 2.3 Frequently Used Measures of Morbidity**

| Measure  | Numerator   | Denominator   |
|--|---|---|
| Incidence proportion<br>(or attack rate or risk)<br>(Cumulative incidence) | Number of new cases of disease during specified time interval                 | Population at start of time interval  |
| Secondary attack rate  | Number of new cases among contacts  | Total number of contacts  |
| Incidence rate<br>(or person-time rate)<br>(Incidence density)             | Number of new cases of disease during specified time interval                 | Summed person-years of observation or average population during time interval |
| Point prevalence   | Number of current cases (new and preexisting) at a specified point in time    | Population at the same specified point in time                                |
| Period prevalence  | Number of current cases (new and preexisting) over a specified period of time | Average or mid-interval population  |

# 1. Incidence rate

The proportion of a population that develops a disease overtime

The risk/probability of an individual developing a disease overtime

The rapidity with which new cases of a disease develop overtime

The proportion of unaffected individuals who on average will contract the disease overtime

## A. Cumulative incidence

It assumes that the entire population is at risk and is followed for a specified time of period.

It is used if the population is more or less stable. Hence, the denominator is considered to be all of the population at the beginning of the period.

$$\text{Cumulative Incidence} = \frac{\text{Number of new cases of a disease during a specified period}}{\text{Population at risk in the same Period}}$$

## A. Cumulative incidence

Example:1 Cumulative incidence demonstrated by the following figure 2.2.

Among twelve individuals at the beginning of the follow-up month-one about three individuals developed disease during the one-year period of follow-up period, What is the cumulative incidence? It is 0.25 per year.

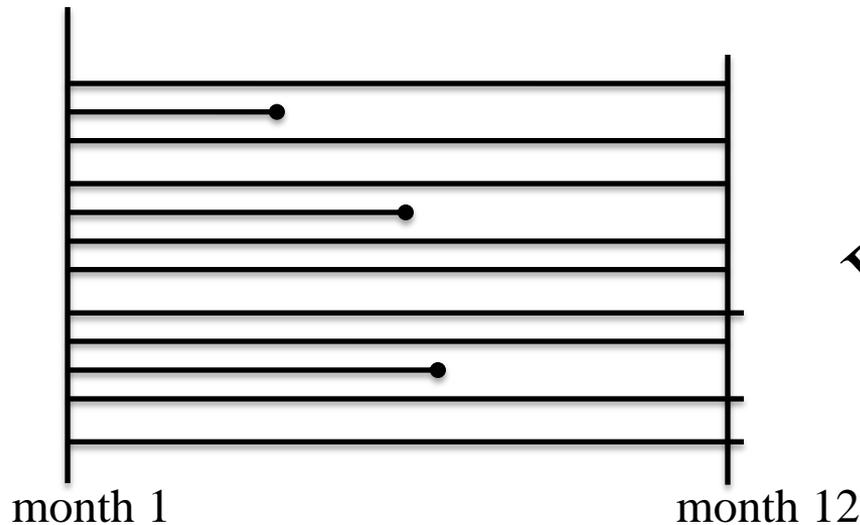


Fig-2.2 Cumulative incidence

*Risk*

$$CI = \frac{3}{12} \text{ per year} = 0.25$$

## B. Incidence density ( rate)

Number of new cases of disease during a period per total person-time of observation.

It must take in to account number of individuals who become ill in a population and time periods experienced by member of the population during which the events occur.

The denominator of this measure is calculated using person time units.

$$\text{Incidence Density} = \frac{\text{Number of new cases of a disease during a specified period}}{\text{Person-time at risk in the same Period}}$$

## B. Incidence density ( rate)

Example:2 Look at table 2.4 and calculate P-Time?

| # of people                    | period at risk    | person-Yrs      |
|--------------------------------|-------------------|-----------------|
| 50                             | 1 year            | $50 * 1 = 50$   |
| 40                             | 6 months (1/2 yr) | $1/2 * 40 = 20$ |
| 20                             | 3 months (1/4 yr) | $1/4 * 20 = 5$  |
| <b>Total person years = 75</b> |                   |                 |

**Person-time-** is the sum of length of time period passed free of illness (at risk) of each individual member of the study.

## B. Incidence density ( rate)

Example-3: The following table presented the five individuals in a dynamic cohort of eleven years follow up period from 1990 to 2000, Coded by A, B, C, D, and E.

- ❑ Individuals leveled as A, C and D did not develop the disease of interest after 6 years (from 1990 to 1995), 10 years (from 1990 to 1999), and 8.5 years (from mid-year of 1990 to 1998) follow-up period, respectively.
- ❑ However, Individuals leveled as B and E developed the disease of interest after 6 years (from 1992 to 1997), 5 years (from 1991 to 1995) follow-up period, respectively.

## B. Incidence density ( rate)

Example-3: look at figure 2.3 below and calculate ID?

| ID.No                     | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 00 | Time at Risk      |
|---------------------------|----|----|----|----|----|----|----|----|----|----|----|-------------------|
| A                         |    |    |    |    |    |    |    |    |    |    |    | 6.0 years         |
| B                         |    |    |    |    |    |    |    |    |    |    |    | 6.0 years         |
| C                         |    |    |    |    |    |    |    |    |    |    |    | 10.0 years        |
| D                         |    |    |    |    |    |    |    |    |    |    |    | 8.5 years         |
| E                         |    |    |    |    |    |    |    |    |    |    |    | 5.0 years         |
| <b>Total Years a risk</b> |    |    |    |    |    |    |    |    |    |    |    | <b>35.5 years</b> |

— Time followed  
X disease onset

$$\text{ID} = 2/35.5 \text{ person-years}$$

$$= 0.056/ \text{ person-years}$$

## C. Attack rate

- ❖ It is a special type of incidence rate which is used when the time interval considered is short.
- ❖ it is a cumulative incidence during an outbreak usually expressed for the entire epidemic period, from the first to the last case.

$$AR = \frac{\text{Number of new cases of a specified disease reported during an epidemic period of time}}{\text{Population at risk during the same time interval}}$$

## D. Secondary attack rate

It is another type of incidence rate

Measures the number of new cases among the contacts of known cases within the total contact population at risk.

Can be calculated as follows

$$\text{SAR} = \frac{\text{Number of new cases of a specified disease among contacts of known cases}}{\text{Size of contact population at risk}}$$

# Uses of incidence rates

Fundamental tool for etiologic studies of acute and chronic disease

Direct measure of risk

## Practical challenges in measuring incidence rate

1. Identification of population at risk

Population at risk constitutes all those free of the disease and susceptible to it

2. Population is not static/it fluctuates/as a result of births, deaths and migration

3. People are at risk only until they get the disease and then no more at risk

## Practical solution to the challenges

1. Use the *total population* as a denominator

This gives an estimate of the incidence rate and not the actual incidence rate

2. Use *person-time* at risk

Incidence density = number of new cases of a disease over a specified period / person-time at risk

## 2. Prevalence rate

It measures the proportion of a population with a disease during a specified period or at a point in time

Two types

1. Point prevalence rate
2. Period prevalence rate

## A. Point prevalence rate

Measures the proportion of a population with a disease at a point in time

The amount of disease in a population is constantly changing,

Thus this may not be useful for assessment of disease with short generation period.

It is not a rate, but a true proportion

## A. Point prevalence rate

$$\text{Point PR} = \frac{\text{All persons with a specific condition at one point in time}}{\text{Total population at a given point in time}}$$

or

$$\text{Point PR} = (\text{IR})(\text{ADD})$$

where, IR = incidence Rate and

ADD = average Disease Duration

## B. Period prevalence rate

Measures the proportion of a population with a disease in a specified time period

$$\text{Period PR} = \frac{\text{All persons with a specific condition overtime period}}{\text{Average(mid - year)population in the same period}}$$

## Uses of prevalence rate

To study chronic diseases

To plan health facilities and manpower

To monitor diseases control programs

To track changes in diseases patterns over time

# Incidence Vs prevalence

Incidence rate considers only new cases of a disease

Prevalence rate considers all (new + old) cases of a disease

Incidence rate considers population at risk as a denominator

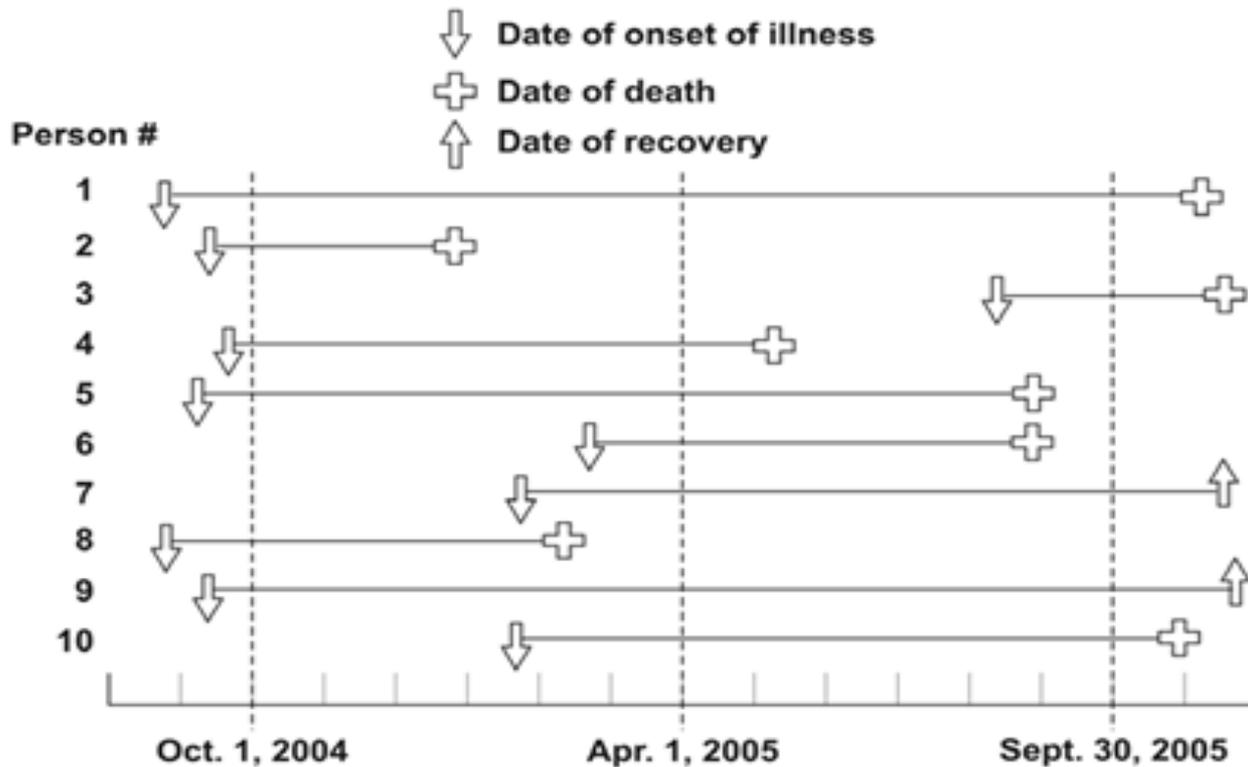
Prevalence rate considers total population as a denominator

Incidence & period prevalence rates require follow up studies

Point prevalence rate requires cross-sectional study

# Incidence Vs prevalence

Example-4: Figure 2.3 New Cases of Illness from October 1, 2004–September 30, 2005



# Incidence Vs prevalence

Example-4:

Figure 2.3 represents 10 new cases of illness over about 15 months in a population of 20 persons.

Each horizontal line represents one person.

The down arrow indicates the date of onset of illness.

The solid line represents the duration of illness.

The up arrow and the cross represent the date of recovery and date of death, respectively.

# Incidence Vs prevalence

Example-4: From the above figure calculate the following

**Example A:** Calculate the incidence rate from October 1, 2004, to September 30, 2005, using the midpoint population (population alive on April 1, 2005) as the denominator. Express the rate per 100 population.

**Example B:** Calculate the point prevalence on April 1, 2005. Point prevalence is the number of persons ill on the date divided by the population on that date. On April 1, seven persons (persons 1, 4, 5, 6, 7, 9, and 10) were ill.

**Example C:** Calculate the period prevalence from October 1, 2004, to September 30, 2005. The numerator of period prevalence includes anyone who was ill any time during the period. In Figure 2.3, the first 10 persons were all ill at some time during the period.

# Relationship between prevalence and incidence rates

Prevalence  $\sim$  incidence

Prevalence  $\sim$  Average duration

Prevalence = Incidence  $\times$  Average duration

An increase in prevalence rate may not necessarily be due to an increase in incidence rate,

it could be due to an increase in average duration of a disease due to decrease in death and/or recovery rates

# Relationship between prevalence and incidence rates

High prevalence may indicate

Increase in survival due to change in virulence of the agent or in host resistance factors

Improvement in medical care

Increase in the incidence rate

Low prevalence may indicate

Rapidly fatal process

Rapid cure of disease

Low incidence rate

# Mortality rates

These rates measures magnitude of deaths in a community

Some are crude like the crude death rate

Others are cause-specific mortality rate

Some others are adjusted like standardized mortality ratio

# Common Mortality rates

- Crude death rate
- Age-specific mortality rate
- Sex-specific mortality rate
- Cause-specific mortality rate
- Proportionate mortality ratio
- Case fatality rate
- Fetal death rate
- Perinatal mortality rate
- Neonatal mortality rate
- Infant mortality rate
- Child mortality rate
- Under-five mortality rate
- Maternal mortality ratio

# Mortality rates...

## A. Crude rate

- ✓ apply to the total population of a given area
- ✓ It is easier to obtain ( calculate )
- ✓ It is difficult to interpret in comparing different countries
- ✓ It is less informative
- ✓ Actual summary rates

## B. Specific rates

- ✓ It is difficult to calculate
- ✓ It is easier to interpret in comparing different countries
- ✓ It is more informative
- ✓ Apply to homogeneous subgroups

# Mortality rates...

## C. Adjusted rates

- ✓ Are summary rates
- ✓ Permit unbiased comparison
- ✓ Easy to interpret
- ✓ Fictitious rates
- ✓ Absolute magnitude depends on standard population
- ✓ Opposing trends in subgroups masked
- ✓ Undergone statistical transformations

# Mortality rates...

## A. Crude death rate

It is the mortality rate from all causes of deaths for the population.

Numerator is all deaths

$$\text{CDR} = \left( \frac{\text{Total number of deaths reported during a given time interval}}{\text{Estimated mid interval population}} \right) \times 1,000$$

# Mortality rates...

## B. Specific Rates

### 1. Cause- specific mortality rate

1. It is the mortality rate from a specified cause for a population.
2. The numerator is the #of deaths attributed to a specific cause
3. The denominator is similar with the crude death rate
4. It is a measure of risk of dying from a certain cause

$$\text{CSMR} = \left( \frac{\text{Number of deaths from a specific cause during a given time}}{\text{Estimated mid interval population}} \right) \times 100,000$$

# Mortality rates...

## 2. Case- fatality rate

It measures the risk of dying of a disease among those with the disease.

It indicates how much the disease is fatal and the causative agent is virulent

$$\text{CFR} = \left( \frac{\text{Number of deaths from a specific disease during a given time}}{\text{Number of cases of the disease}} \right) \times 100$$

# Mortality rates...

## 3. Proportionate-mortality ratio

- It is not a measure of risk of death.
- It is a measure of importance of certain disease as a cause of death.
- A higher proportionate mortality ratio doesn't necessarily imply higher risk of death from the disease.

$$PMR = \left( \frac{\text{Number of deaths from a specific disease during a given time}}{\text{Total number of deaths from all causes in the same time}} \right) \times 100$$

# Mortality rates...

## 4. Fetal Death Rate

$$\text{FDR} = \left( \frac{\text{Number of fetal deaths of 28 Wks or more gestation reported during a given time}}{\text{Total number of fetal deaths of 28 Wks or more gestation reported and live birth at the same time}} \right) \times 1,000$$

# Mortality rates...

## 5. Perinatal Mortality Rate

$$\text{PMR} = \left( \frac{\begin{array}{c} \text{Number of fetal deaths of 28 Wks or} \\ \text{more gestation plus} \\ \text{Number of infant death under 7 days} \\ \text{reported during a given time} \end{array}}{\begin{array}{c} \text{Total number of fetal deaths of 28 Wks or} \\ \text{more gestation reported and} \\ \text{live birth at the same time} \end{array}} \right) \times 1,000$$

## 6. Neonatal Mortality Rate

$$\text{NMR} = \left( \frac{\begin{array}{c} \text{Number of deaths under 28 days of age} \\ \text{reported during a given time} \end{array}}{\begin{array}{c} \text{Number of live births reported} \\ \text{during the same time} \end{array}} \right) \times 1,000$$

# Mortality rates...

## C. Adjusted rates

### Standardization

- ❑ Def<sup>n</sup>: A method used to compare the occurrence of death or mortality for a country or region.
- ❑ In essence, adjustment that weights an over all figure( disease occurrence or mortality) according to the risk factor profile of the country or region.

# Mortality rates...

## Types of standardization

1. **Direct**; depends on use of a population whose age-structure is standardized.
2. **Indirect**: depends on standard stratum-specific rates.

# Mortality rates...

Example

Mortality in UK vs. Kenya

|          | UK     |                  |      | Kenya  |                  |      |
|----------|--------|------------------|------|--------|------------------|------|
|          | Deaths | pop <sup>n</sup> | Rate | Deaths | pop <sup>n</sup> | Rate |
| All ages | 654.6  | 58,649           | 11.2 | 250.1  | 29,008           | 8.6  |

# Mortality rates...

## Deaths by age

|          | UK     |                  |      | Kenya  |                  |      |
|----------|--------|------------------|------|--------|------------------|------|
| Age      | Deaths | pop <sup>n</sup> | Rate | Deaths | pop <sup>n</sup> | Rate |
| 0-29     | 29     | 22,287           | 1.3  | 90     | 16,825           | 5.4  |
| 30-59    | 105.9  | 25,219           | 4.2  | 71     | 10,443           | 6.8  |
| 60+      | 521.6  | 11,143           | 46.8 | 88.2   | 1,740            | 50.7 |
| All ages | 656.4  | 58,649           | 11.2 | 250.1  | 29,008           | 8.6  |

# Mortality rates...

## Direct standardization

| Age      | Standard pop <sup>n</sup> | Uk expected               | Kenya Expected            |
|----------|---------------------------|---------------------------|---------------------------|
| 0-29     | 56,000                    | 0.0013 X 56,000<br>=72.8  | 0.0054 X 56,000<br>=302.4 |
| 30-59    | 33,000                    | 0.0042 X 33,000<br>=138.6 | 0.0068 X 33,000<br>=224.4 |
| 60+      | 11,000                    | 0.0468 X 11,000<br>=514.8 | 0.0507 X 11,000<br>=557.7 |
| All ages | 100,000                   | 726.2                     | 1084.5                    |

# Mortality rates...

## Comparative mortality figure

✓ It is an age standardized ratio

$$\text{CMF} = \frac{\text{Age standardized mortality in Kenya}}{\text{Age standardized mortality in UK}} = \frac{1084.5}{726.5} = 1.49$$

### Interpretation

- The expected mortality of Kenya is 49% higher than the expected mortality of UK if the age-structure is standardized similarly for both using the above standard population.

# Mortality rates...

## Indirect standardization

| Age      | UK Mortality | Kenya popn | Kenya Expected Deaths |
|----------|--------------|------------|-----------------------|
| 0-29     | 0.0013       | 16,825     | 21.9                  |
| 30-59    | 0.0042       | 10,443     | 43.9                  |
| 60+      | 0.0468       | 1,740      | 81.4                  |
| All ages |              |            | 147.2                 |

# Mortality rates...

## Standardized Mortality ratio (SMR)

$$\text{SMR} = \frac{\text{observed deaths}}{\text{Expected deaths}} = \frac{250.1}{147.2} \times 100 = 170\%$$

### Interpretation

- the number of observed deaths in Kenya is about 70 % higher than the number expected if Kenya had the same mortality experience as the UK.

# Mortality rates...

If  $SMR > 1$

- ❑ More deaths are observed in the smaller population than would be expected on the bases of rates in the larger (standard) population.

If  $SMR < 1$

- ❑ Fewer deaths are observed than expected.
- ❑ This method is used to compare two populations, in one of which the ASMR are not known or are excessively variable because of small numbers.

# 4. Epidemiologic Study Designs

# Session Objectives

After completing this session learners will be able to

- ❖ Understand epidemiological study designs
- ❖ Differentiate the type of study design used by investigators
- ❖ Identify the merits and demerits of each study designs
- ❖ Apply epidemiological study designs to conduct a study.

# Study design

Study design is the arrangement of conditions for the collection and analysis of data to provide the most accurate answer to a question in the most economical way.

# Types of Epidemiologic study designs

- I. Based on objective/focus/research question
  1. Descriptive studies
    - Describe: who, when, where & how many
  2. Analytic studies
    - Analyse: How and why

# Types...

## II. Based on the role of the investigator

### 1. Observational studies

- The investigator observes nature
- No intervention

### 2. Intervention/Experimental studies

- Investigator intervenes
- He has a control over the situation

# Types...

## III. Based on timing

### 1. One-time (one-spot) studies

- Conducted at a point in time
- An individual is observed at once

### 2. Longitudinal (Follow-up) studies

- Conducted in a period of time
- Individuals are followed over a period of time

# Types...

## IV. Based on direction of follow-up/data collection

### 1. Prospective

- Conducted forward in time

### 2. Retrospective

- Conducted backward in time

# Types...

## V. Based on type of data they generate

### 1. Qualitative studies

- Generate contextual data
- Also called exploratory studies

### 2. Quantitative studies

- Generate numerical data
- Also called explanatory studies

# Types...

## VI. Based on study setting

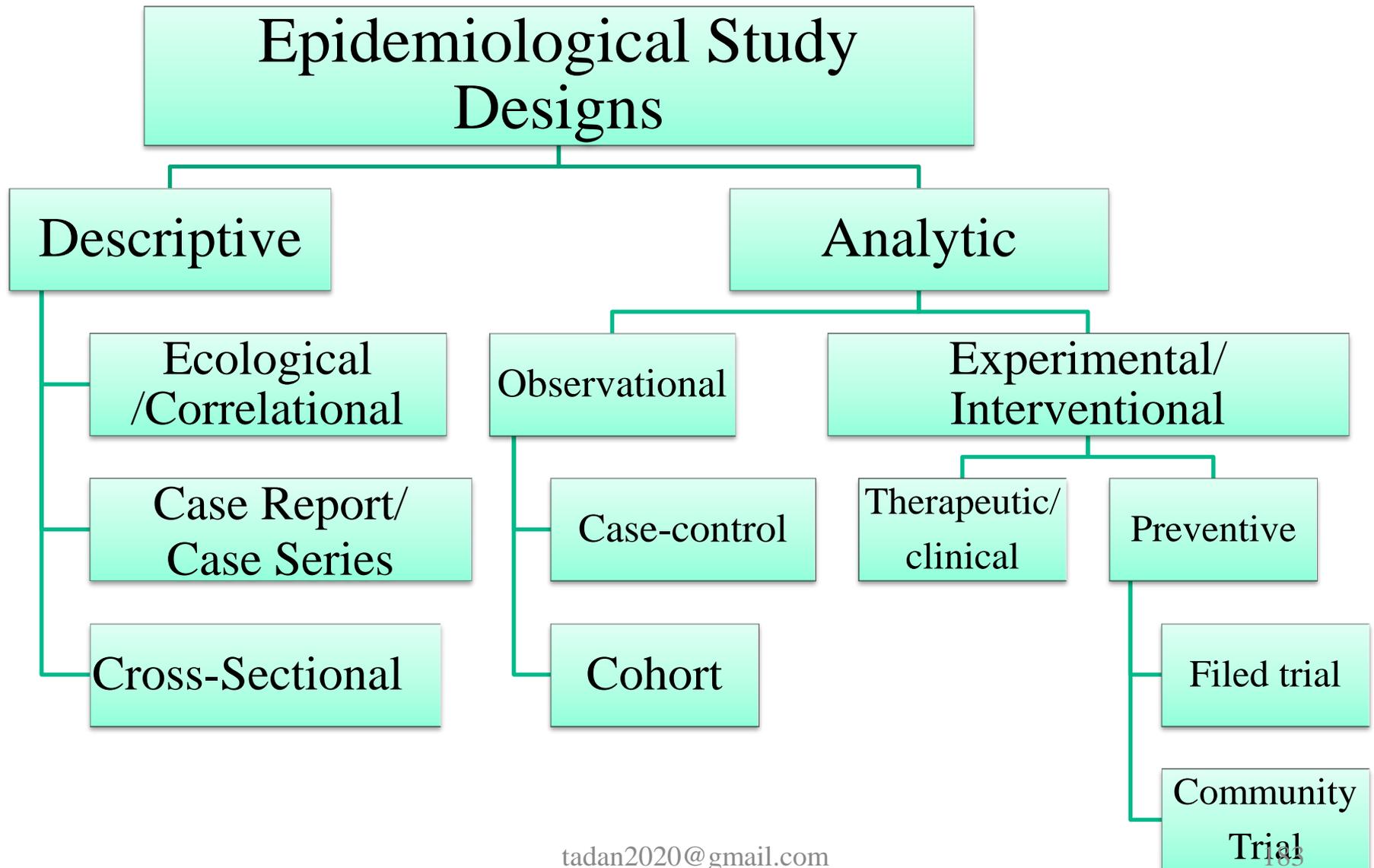
1. Community-based studies
  - Conducted in communities
2. Institution-based studies
  - Conducted in institutions
3. Laboratory-based studies
  - Conducted in major laboratories

# Types...

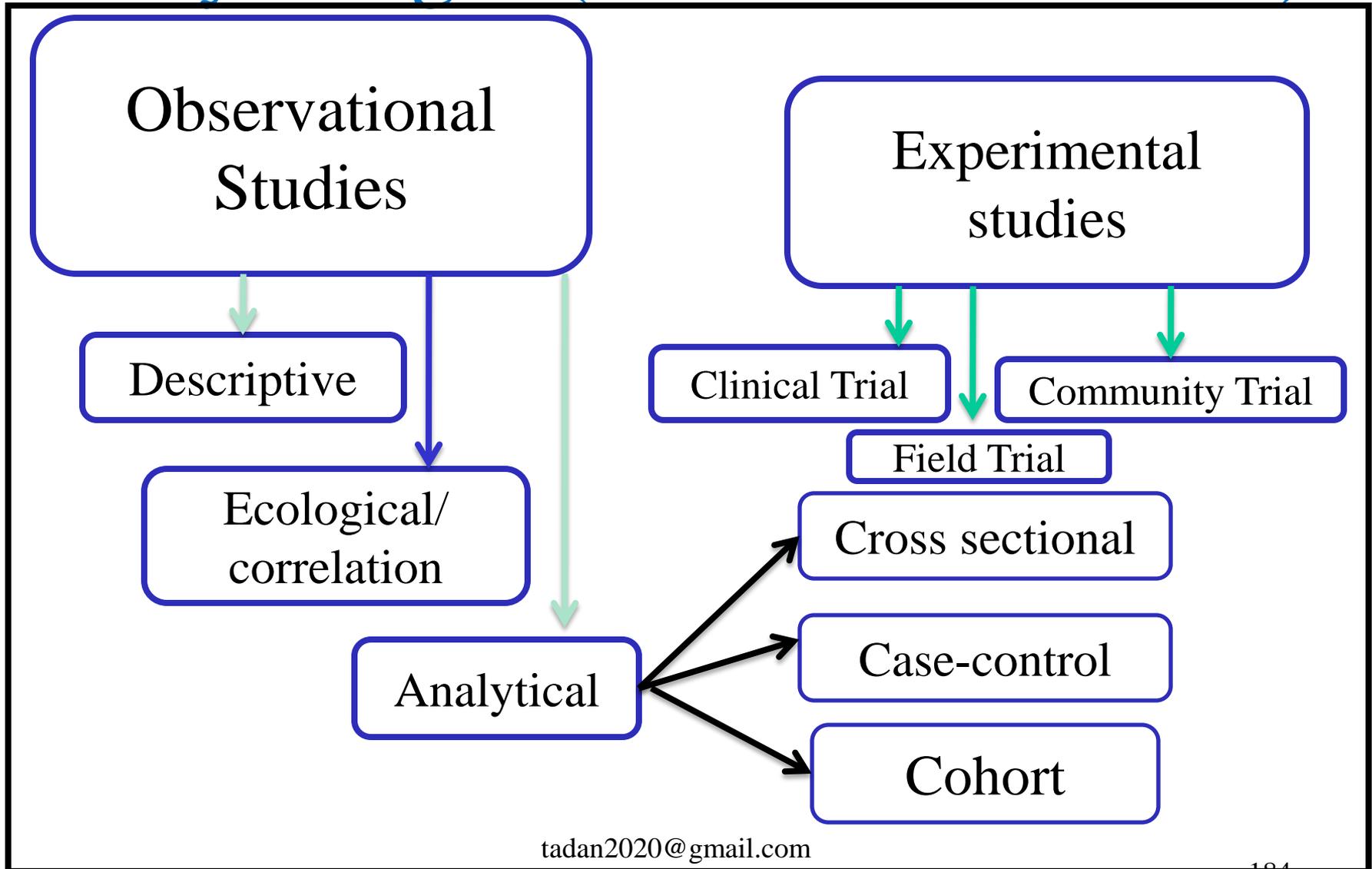
## VII. Standard classification

1. Cross-sectional studies
2. Case-control studies
3. Cohort studies
4. Experimental studies

# Classification of Study designs



# Study designs (other classification)



# Descriptive studies

- ❖ Relatively cheap in terms of time cost because it can use information already collected.
  
- ❖ It usually describes:
  1. Who gets a disease and who doesn't ( person)
  2. Where rates are highest and lowest ( place)
  3. Temporal patterns of disease (time)

# Descriptive.....

- ✍ It is the most common form of study in medical literatures
  
- ✍ There are three forms of descriptive study designs
  1. Correlational /ecological design  
( use population as study subject)
  2. Case report and/or case series studies
  3. Cross-sectional survey

# 1. Correlational/ Ecological study

- ✍ Use data from entire population (as a whole) to compare disease frequencies.
- ✍ Hence, it does not need data from individuals.
- ✍ Can be done quickly and inexpensively, often by using already available data.
- ✍ It has the following rationale
  1. Low cost and convenience
  2. Measurement limitations at individual level, Eg environmental contact.
  3. Other designs may be unable to measure interest on ecologic effect

# Ecological / correlational.....

- ❑ For initial investigation of causal hypothesis like a cross-sectional study design data on both the exposure and outcome of interest will be collected at one time (simultaneously).
- ❑ However the exposure and outcome of interest measured at group and not at individual level.

# Ecological / correlational.....

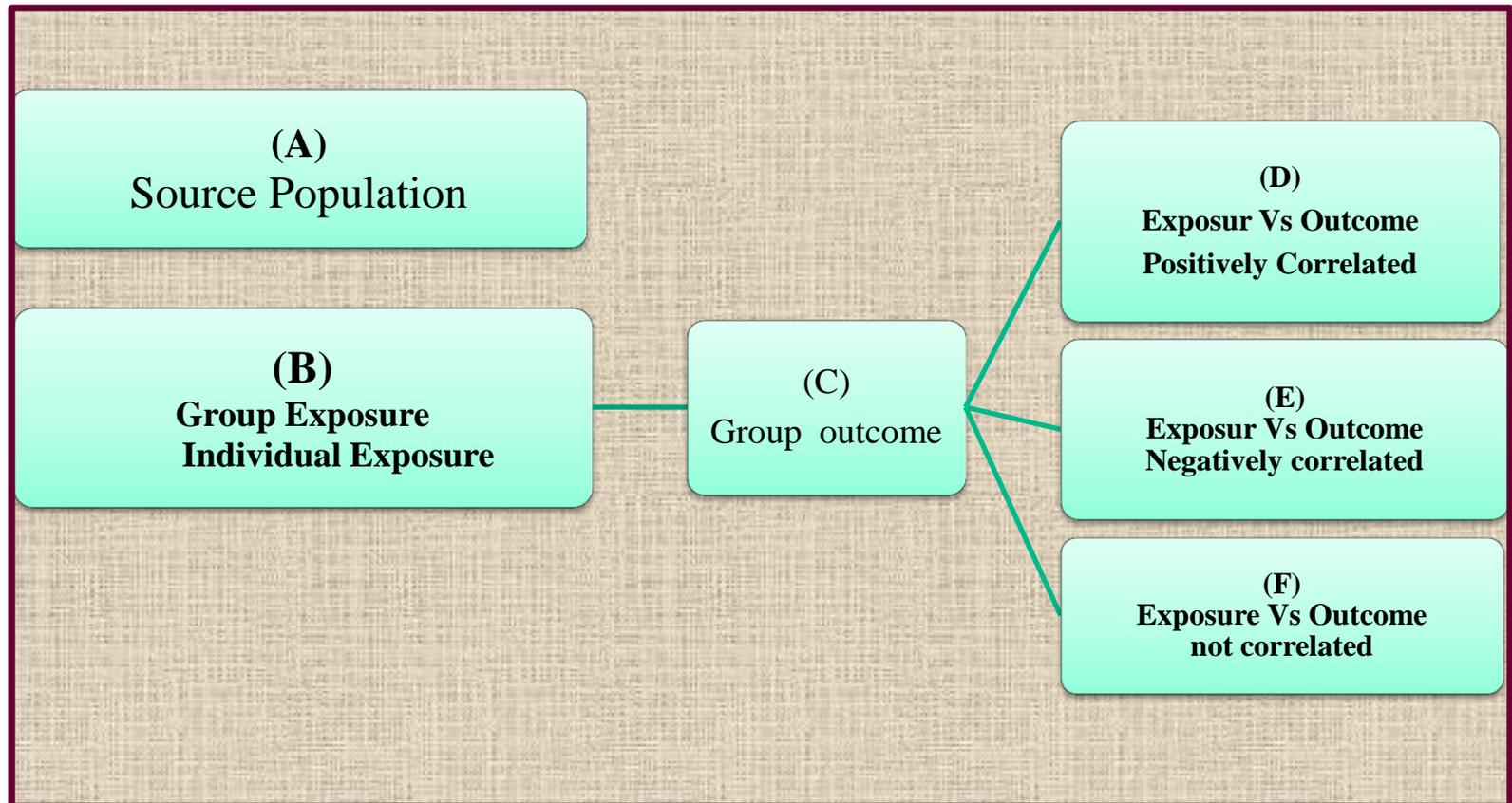
- ❑ Since ecological/Correlation studies take at least one group-level data (aggregate/average measurements of exposure and/ or outcome of interest),
- ❑ most often has no a comparative group and not used to test causal hypothesis rather help to generate causal hypothesis.

# Ecological / correlational.....

- ❑ However, comparison for the purpose of initial causal inference can be made during data analysis at three levels such as at:
  1. biological (by person),
  2. ecological (by group) or
  3. contextual (both by individual & group) characteristics.
  
- ❑ The design of ecological studies is completely similar to a cross-sectional survey type of design for descriptive and a comparative cross-sectional longitudinal design for analytic purposes.

# Ecological / correlational.....

The design is similar to cross-sectional study design



# Limitations

1. Unable to link an exposure to occurrence of disease in a single individual.
2. Lack of the ability to control for effect of confounders.
3. Data represent average exposure levels rather than actual individual values, ecological fallacy or bias.
4. Mask non-linear relationship between exposure and outcome

## 2. Case reports/ case series

- ❑ Useful for the recognition of new disease.
- ❑ Important for constructing the natural history of disease.
- ❑ Help to formulate a research hypothesis and detection of epidemics.

### 2.1. Case report

- ❖ It is the study of health profile of a single individual using a careful and detailed report by one or more clinicians.
- ❖ It is common form that is published in articles.
- ❖ It is made using a history, physical examination and lab investigation.

## 2. Case reports/ case series....

### 2.1. Case report...

- ❖ Report is usually documented if there is unusual medical occurrence, thus it may be first clue for identification of a new disease.
- ❖ It is useful in constructing a natural history of individual disease
- ❖ N.B. it was a single case report that formulated hypothesis of OC use increases venous thromboembolism.

## 2.2. Case series

- ❑ Individual case report can be expanded to a case series, which describes characteristics of a number of patients with a similar disease.
- ❑ It is often used to detect the emergency of new disease or epidemics. Eg the first five ADIS cases in USA.

### Limitation

- ❑ Both case report and case series are able to formulate a hypothesis but are not able to test for the presence of valid association.
- ❑ Fundamental limitation of case report is presence of a risk factor could be simply coincidental.
- ❑ You are not able to test for association because there is no relevant comparison group.

### 3. Cross-sectional studies

In this study design information about the status of an individual with respect to presence/absence of exposure and disease is assessed at a point in time.

Cross-sectional studies are useful to generate a hypothesis rather than to test it

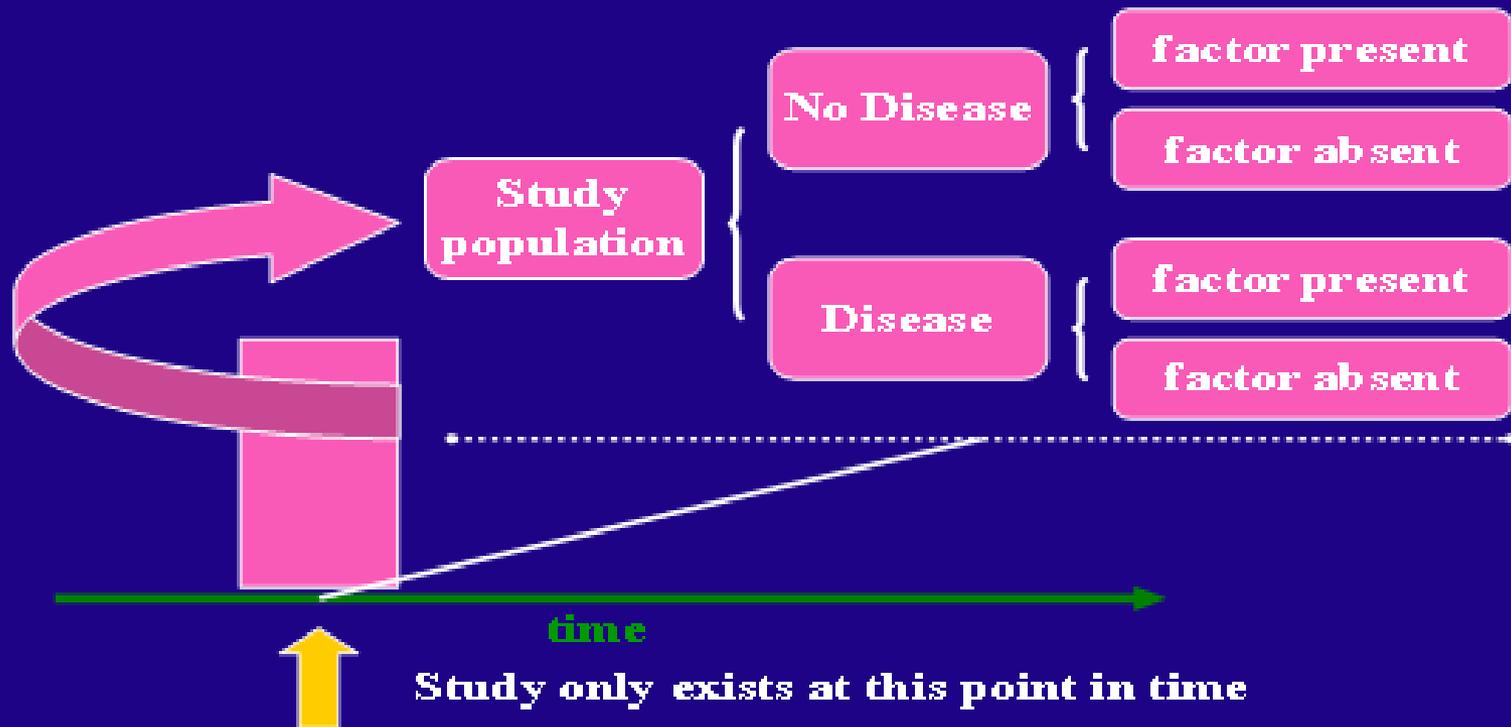
For factors that remain unaltered over time (e.g. sex, race, blood group) it can produce a valid association

## Cross-sectional...

- Comparison groups are formed after data collection
- The object of comparison are prevalence of exposure or disease
- Groups are compared either by exposure or disease status
- Cross-sectional studies are also called prevalence studies
- Cross-sectional studies are characterized by concurrent classification of groups

# Cross-sectional...

## Cross-sectional Design



# Cross-sectional...

## Advantages of cross-sectional studies

- Less time consuming
- Less expensive
- Provides more information
- Describes well
- Generates hypothesis

# Cross-sectional...

## Limitations of cross-sectional studies

- Antecedent-consequence uncertainty  
    “Chicken or egg dilemma”
- Data dredging leading to inappropriate comparison
- More vulnerable to bias

# Cross-sectional...

## Types of cross-sectional studies

### 1. Single cross-sectional studies

- Determine single proportion/mean in a single population at a single point in time

### 2. Comparative cross-sectional studies

- Determine two proportions/means in two populations at a single point in time

### 3. Time-series cross-sectional studies

- Determine a single proportion/mean in a single population at multiple points in time

# Cross-sectional...

In general, cross-sectional studies are

- Simplest to conduct
- Commonest to find
- Least useful to establish causation

# Analytical studies

- ❖ Purpose / aim of analytic epidemiologic study designs
  - ✚ To test hypothesis about causal relationship
  - ✚ To search for cause & effect
  - ✚ To compare treatment regimens/prevention programs
  - ✚ To assess diagnostic tests
  - ✚ To quantify the association b/n exposure & outcome (measure of association)
  - ✚ It focused on determinants of the diseases by testing hypothesis

# 1. Case-control studies

- Subjects are selected with respect to the presence (cases) or absence (controls) of disease, and then inquiries are made about past exposure
- We compare diseased (cases) and non-diseased (controls) to find out the level of exposure
- Exposure status is traced backward in time

# Case-control...

## Steps in conducting case-control studies

### I. Define who is a case anaemia

- Establish strict diagnostic criteria population
- All who fulfil the criteria will be “case population
- Those who don’t fulfil will be “control population”

### II. Select a sample of cases from case population

- This sample must be representative of the case population

# Case-control...

## Sources of cases

### 1. Hospitals (Health institution)

- Cost-less
- Bias-more

### 2. Population (Community)

- Cost-more
- Bias-less

# Case-control...

## III. Select controls from a control population

- Should be representative of control population
- Should be similar to cases except outcome
- Should be selected by the same method as cases

## Sources of controls

### 1. Hospital (Health institution) controls

- Readily available
- Low recall bias
- More cooperative

## Case-control...

However, hospital controls are

- Less representative
- More confounding

2. Population (community) controls

- More representative
- Less confounding
- Costly and time consuming
- More recall bias
- Less cooperative

# Case-control...

## IV. Measure the level of exposure in cases & controls

- Review or interview for exposure status
- Use same method for case and controls

## V. Compare the exposure between cases & controls

- Prepare 2X2 table
- Calculate OR
- Perform statistical tests

# Case-control...

## Types of case-control studies

### I. Based on case identification

#### 1. Retrospective case-control

- Uses prevalent cases
- Increased sample size
- Difficult to establish temporal sequence
- Useful for rare outcomes

# Case-control...

## 2. Prospective case-control

- Uses incident cases
- Establish temporal sequence
- Recall is not a serious problem
- Records are easily obtainable
- Common as nested with other study designs

# Case-control...

## II. Based on matching

Matching: Relating cases and controls with respect to certain variable  
(handling confounding variables at design)

1. Matched case-control studies
2. Unmatched case-control studies

# Case-control...

## Advantages of case-control studies

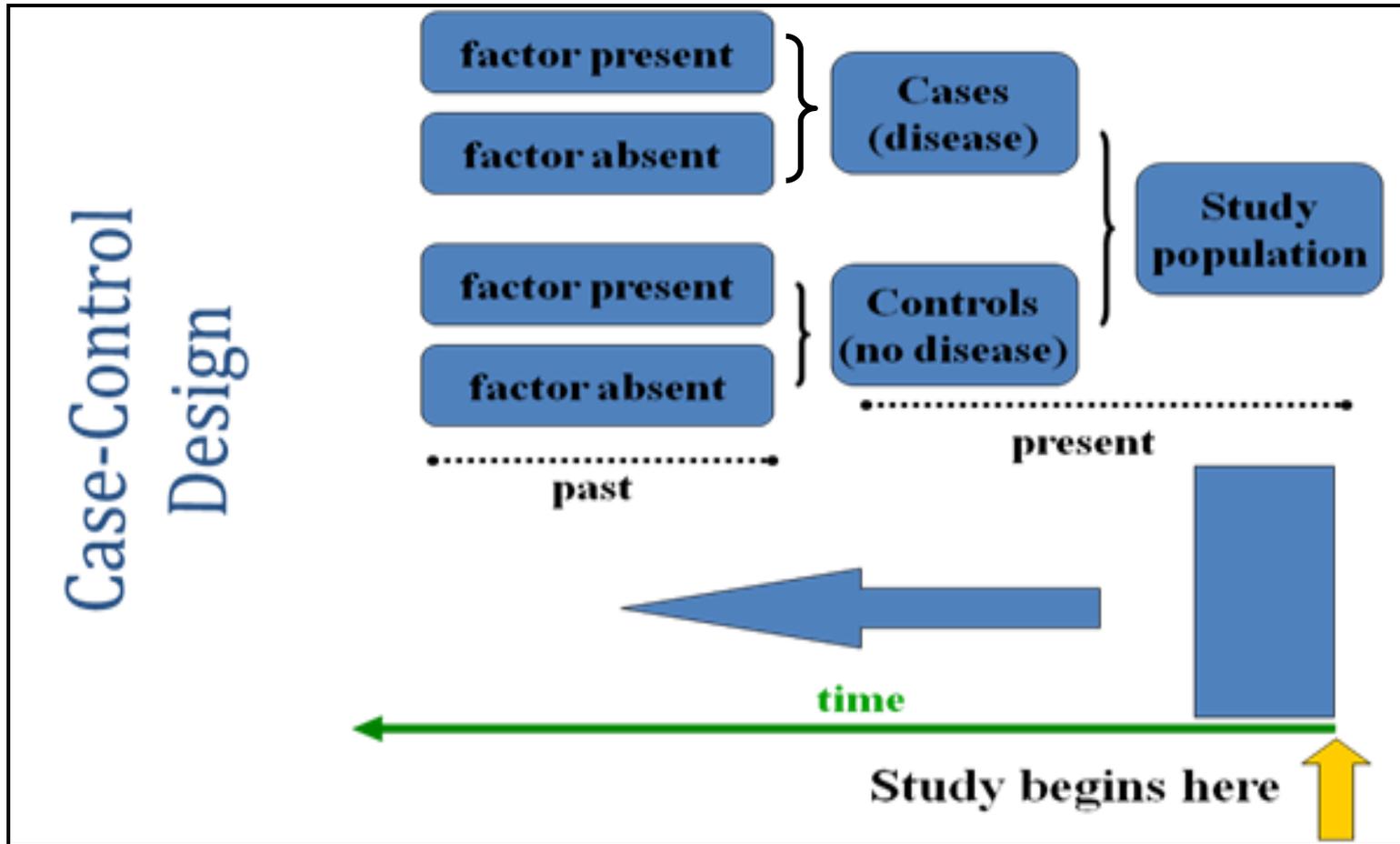
- Optimal for evaluation of rare diseases
- Examines multiple factors of a single disease
- Quick and inexpensive
- Relatively simple to carry out
- Guarantee the number of people with disease

# Case-control...

## Limitations of case-control studies

- Inefficient for evaluation of rare exposure
- Can't directly compute risk
- Difficult to establish temporal sequence
- Determining exposure will often rely on memory

# The design of Case-control Study



## 2. Cohort studies

- Subjects are selected by exposure and followed to see development of disease
- Two types of cohort studies
  1. Prospective (classical)
    - Outcome hasn't occurred at the beginning of the study
    - It is the commonest and more reliable

# Cohort...

## 2. Retrospective (Historical)

- Both exposure and disease has occurred before the beginning of the study
- Faster and more economical
- Data usually incomplete and in accurate

# Cohort...

## Steps in conducting cohort studies

1. Define exposure
2. Select exposed group
3. Select non-exposed group
4. Follow and collect data on outcome
5. Compare outcome b/n exposed & non-exposed

# Cohort...

## Advantages of cohort studies

- Valuable when exposure is rare
- Examines multiple effects of a single exposures
- Temporal relationship is known
- Allow direct measurement of risk
- Minimize bias in ascertainment of exosure

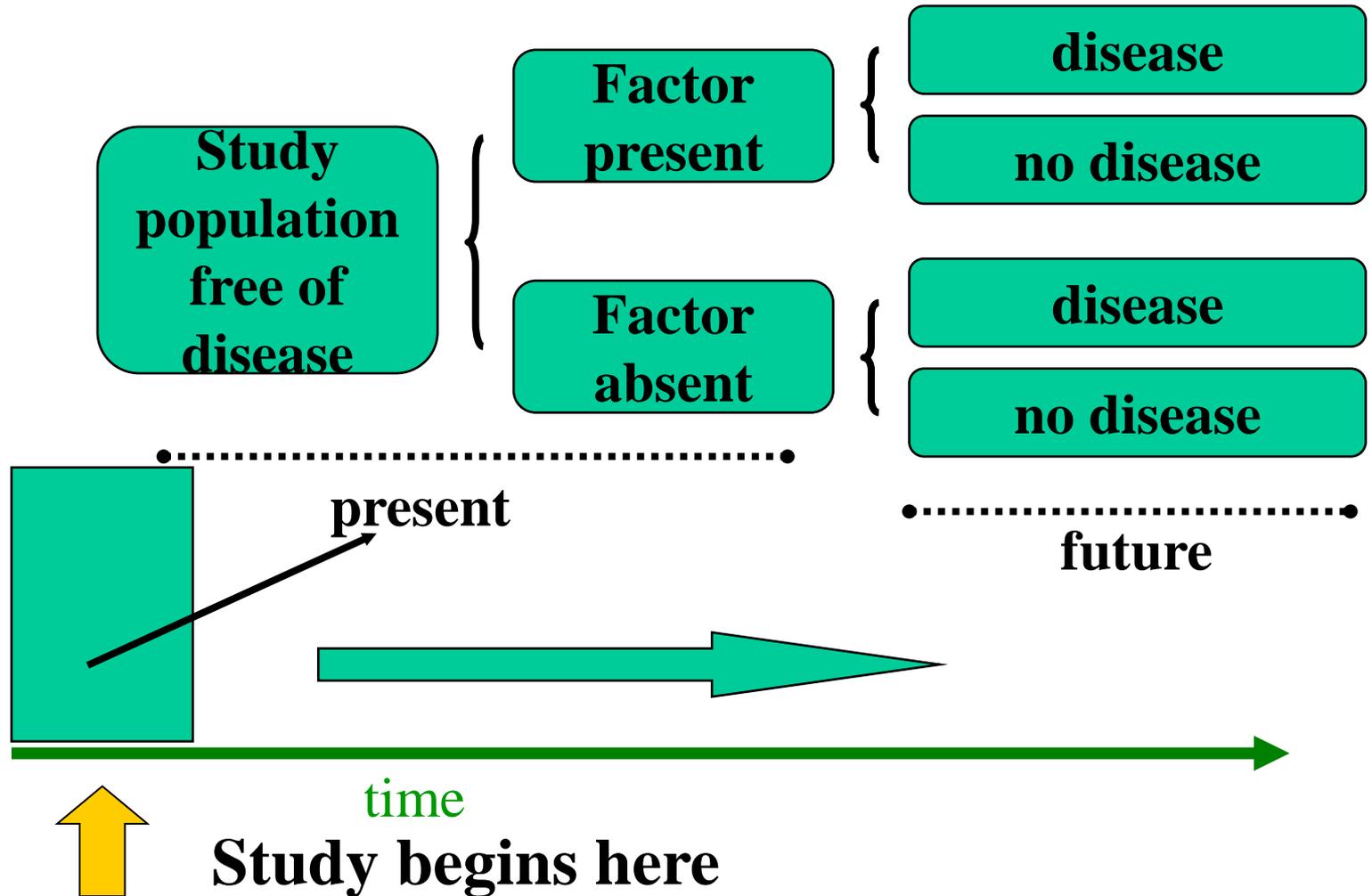
# Cohort...

## Limitations of cohort studies

- Inefficient for evaluation of rare disease
- Expensive
- Time-consuming
- Loss to follow-up creates a problem

# The design of Cohort Study

## Cohort Design



### 3. Experimental studies

- Individuals are allocated in to treatment and control groups by the investigator
- If properly done, experimental studies can produce high quality data
- They are the gold standard study design

# Experimental...

Experimental studies can be

## 1. Therapeutic trials

- Conducted on patients
- To determine the effect of treatment on disease

## 2. Preventive trials

- Conducted on healthy people
- To determine the effect of prevention on risk

# Classification of experimental studies

## 1. Based on the study subjects

- A. Clinical trial-usually performed in the clinical settings and the study subjects are patients.
- B. Field trial-used in testing medicines for preventive purpose and the study subjects are healthy people.
- C. Community trial- a field trial in which the unit of the study is group of people/community.

# Experimental....

## 2. Based on the study design

- A. Uncontrolled trial - no control group. Control will be past experience (history).
  
- B. Non-randomized controlled trial - there is control group but allocation into either group is not randomized. (quasi-experimental study design)
  
- C. Randomized controlled trial - there is control group and allocation into either group is randomized.

# Experimental....

## 3. Based on Trial Objective

- A. Phase I – clinical pharmacologic study - trial on small subjects to test a new drug with small dosage to determine the toxic effect, safe dose range. Scale: 20-80 healthy individuals.  
( Treatment Mechanism & Dose Finding study)
  
- B. Phase II–Safety and efficacy studies- trial on small group to determine the therapeutic effect. Scale: 100-200 patients.

# Experimental....

## 3. Based on Trial Objective

C. Phase III - study on large population to test effectiveness. Randomized Controlled Trial (RCT), often it is multi-centered and needs licensing. (Comparative SE study)

D. Phase IV – Post marketing surveillance – Long term prospective assessment of effects & side-effects. (ES study)

## Purposes of experimental studies

- ❑ When to choose an experimental design?  
Experimental studies are generally reserved for relatively "mature" research questions.
- ❑ A lot of ground work has to be done before embarking on a clinical trial.

## Purposes .....

- Experimental designs are chosen when one or more of the following conditions are fulfilled.
- ✗ The research question cannot be answered by observational studies
- ✗ Earlier observational studies have not answered the research question
- ✗ Existing knowledge is not sufficient to determine clinical or public health policy
- ✗ An experiment is likely to provide an important extension of this knowledge.

# Experimental....

- Interventions that can be evaluated are:-
  - ✍ New drugs and new treatment of diseases,
  - ✍ New medical and health care technology,
  - ✍ New methods of primary prevention,
  - ✍ New programs for screening,
  - ✍ New ways of organizing & delivering health services,
  - ✍ New community health programs and
  - ✍ New behavioral intervention programs.

# Experimental....

□ Generally, a good experimental design serves three purposes.

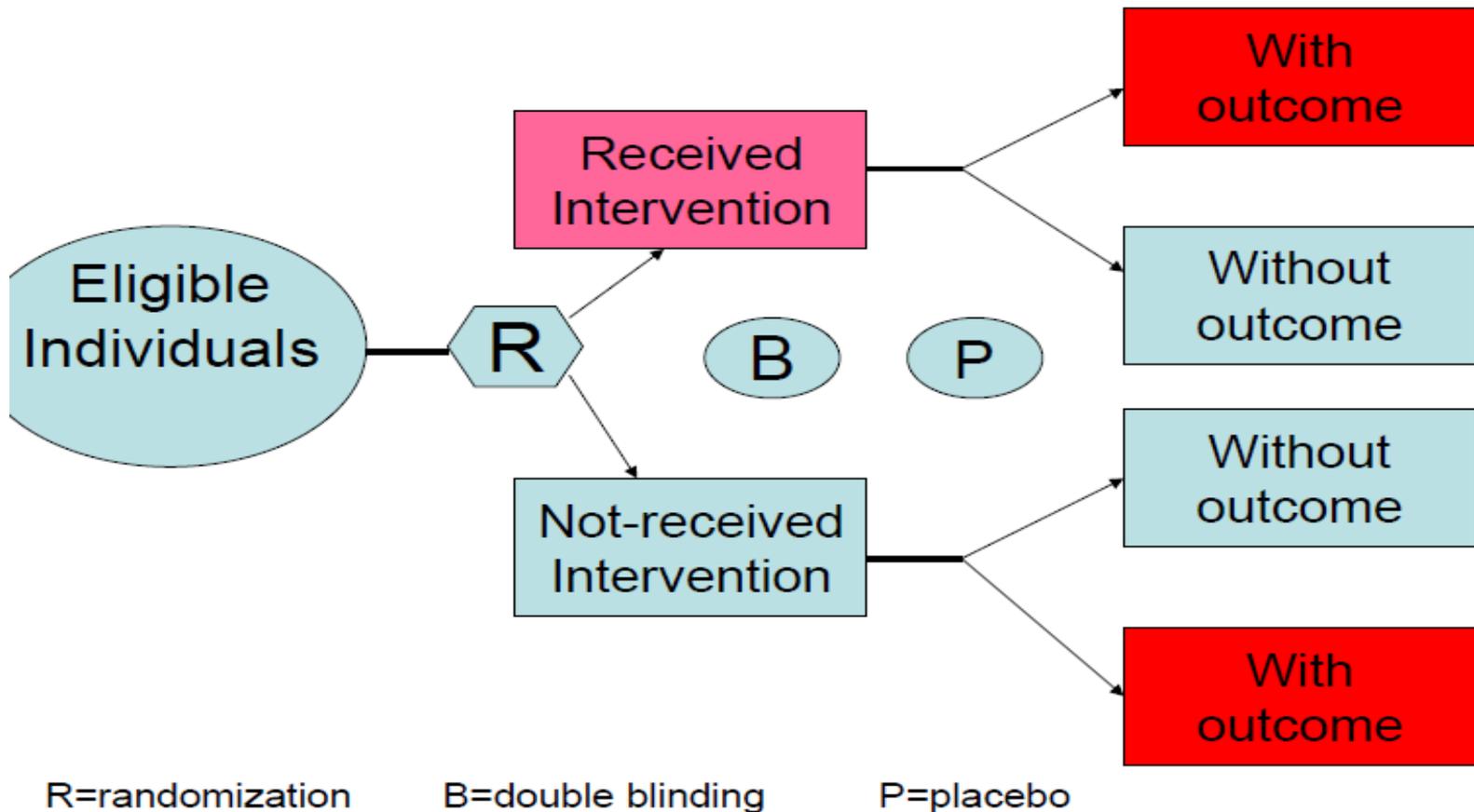
- ✂ **Causation:** It allows the experimenter to make causal inferences about the relationship between *independent variables* and a Dependent Variable.
- ✂ **Control:** It allows the experimenter to rule out alternative explanations due to the *Confounding* effects of extraneous variables (i.e., variables other than the independent variables).
- ✂ **Variability:** It reduces *variability within treatment* conditions, which makes it easier to detect differences in treatment outcomes.

# Key features of experimental study design

- ❑ There are two key features of experimental of interventional study designs such as
  - 1) Investigator manipulates the condition under study
  - 2) Always prospective

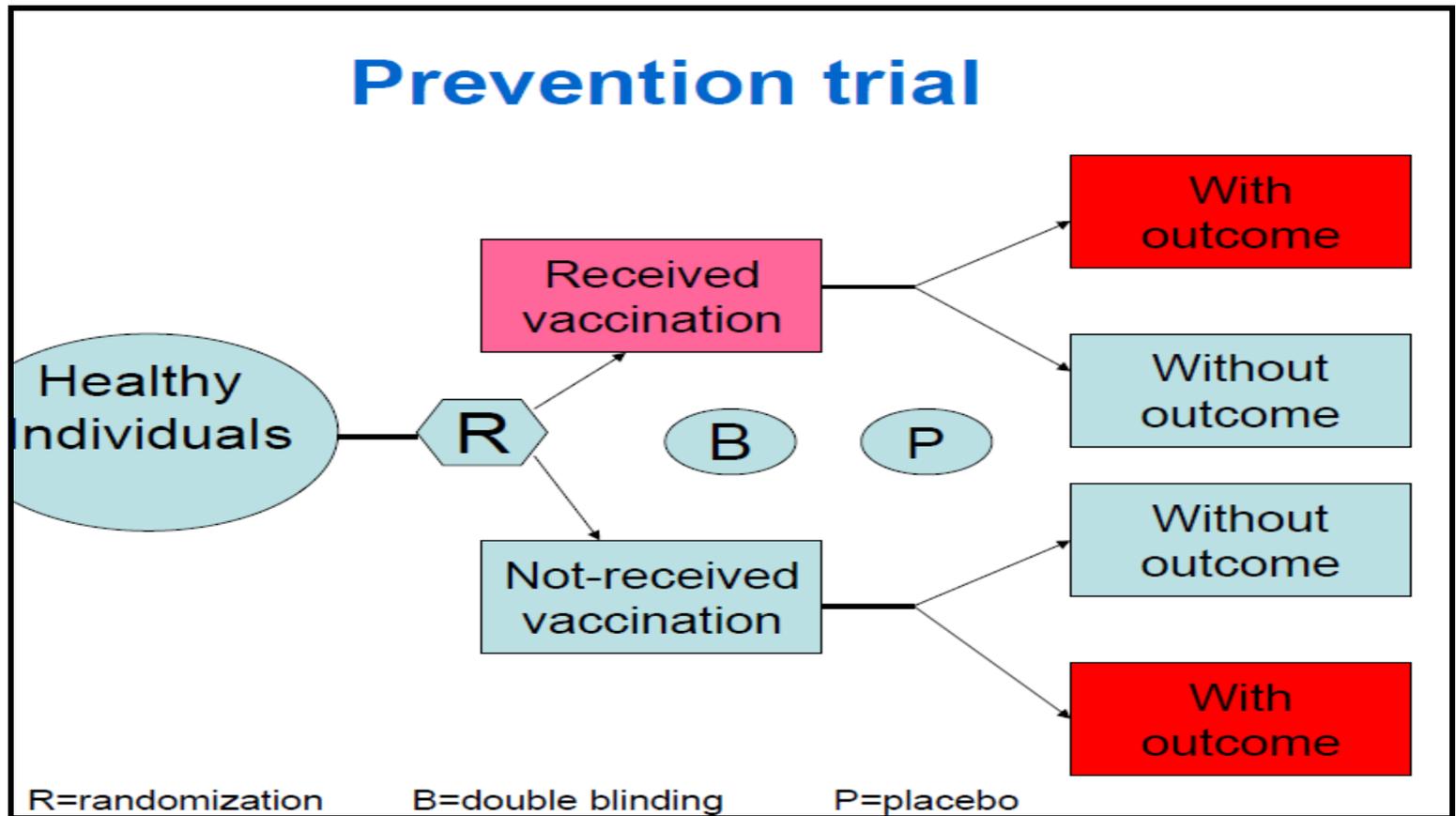
# Designs of Experimental Studies

## Design of Randomized Controlled trial (RCT)



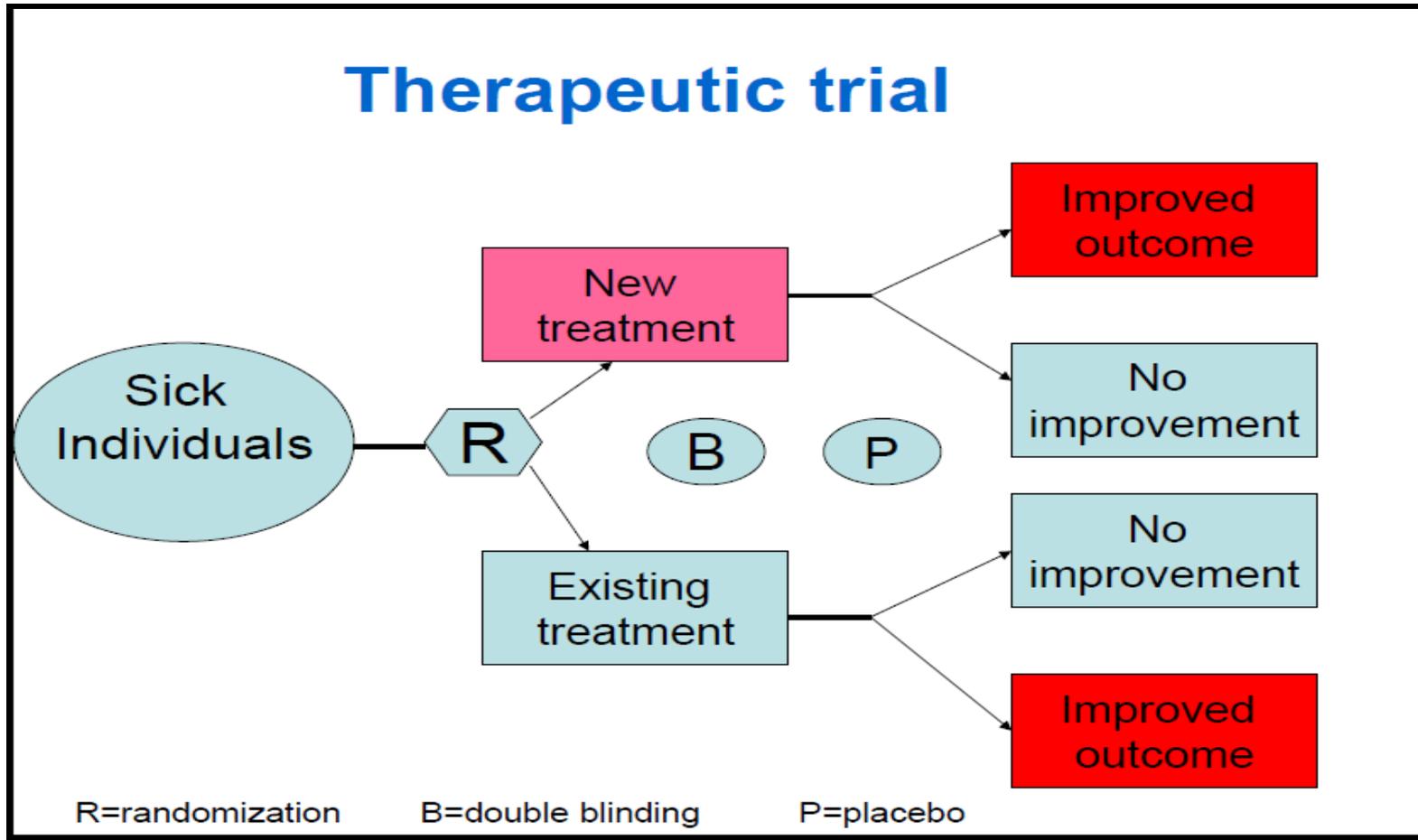
# Designs of experimental studies

## Design of Randomized Controlled Filed trial (RCFT)



# Designs of experimental studies

## Design of Randomized Controlled Clinical trial (RCCT)



# Experimental...

- Challenges in intervention studies
- **Ethical issues**
  - ✍ Harmful treatment shouldn't be given
  - ✍ Useful treatment shouldn't be denied
- **Feasibility issues**
  - ✍ Getting adequate subjects
  - ✍ Achieving satisfactory compliance
- **Cost issues**
  - ✍ Experimental studies are expensive

# Experimental...

- The quality of “Gold standard” in experimental studies can be achieved through
  - ❖ Randomization
  - ❖ Blinding
  - ❖ Placebo
  
- **Randomization:** random allocation of study subjects in to treatment & control groups
  - ❖ Advantage: Avoids bias & confounding
  - ❖ Increases confidence on results
  
- **Blinding:** Denying information on treatment/ control status

# Experimental...

- ❑ **Single blinding:** study subjects don't know to which group they belong
  
- ❑ **Double blinding:** Care givers also don't know to which group study subjects belong
  
- ❑ **Triple blinding:** data collectors also don't know allocation status
  
- ❖ Advantage: Avoids observation bias

# Experimental...

- ❑ **Placebo:** an inert material indistinguishable from active treatment
  
- ❑ **Placebo effect:** tendency to report favourable response regardless of physiological efficacy
  
- ❖ Placebo is used as blinding procedure

# Experimental...

## □ Advantages

- ❖ Randomization minimize bias & confounding
- ❖ Blinding minimizes observer bias
- ❖ Placebo could blind patients

## □ Disadvantage

- ❖ Generalizability limited by selection of participants
- ❖ Rare/late adverse effects of intervention may not be detected
- ❖ High ethical problem

# 5. Measures of association

# Learning Objectives

At the end of this unit students will be able

1. To understand what is measure of association
2. To identify which measure of association is appropriate for a specific study design
3. To calculate measure of association for cause-effect relationship arises from each study design
4. To differentiate the types of measure of associations ( $\chi^2$ , RR, OR, AR, PAR)

# Epidemiological 2X2 table

- 2x2 tables can arise from each of the designs where there are two factors of interest (exposure status and disease) and the data are the numbers of individuals (frequencies) in each four cells.

|          |         | Disease |        | Total   |
|----------|---------|---------|--------|---------|
|          |         | Yes (+) | No (+) |         |
| Exposure | Yes (+) | a       | b      | a+b     |
|          | No (+)  | c       | d      | c+d     |
| Total    |         | a+c     | b+d    | a+b+c+d |

## 2X2 table Cells

- We will call the factor levels “exposed” (E) and “unexposed” (NE) and “diseased” (D) and “not-diseased” (ND).
  1. a= Exposed, and diseased
  2. b= Exposed, Not diseased
  3. c= Not exposed, diseased
  4. d= Not exposed, Not diseased

## 2X2 table Totals

### Marginal totals

$$a+b = \text{Exposed}$$

$$c+d = \text{Non-exposed}$$

$$a+c = \text{Diseased}$$

$$b+d = \text{Non-diseased}$$

### Grand total

$$n = a+b+c+d \quad (\text{sample size})$$

# Chi-square statistics

Chi-square tests whether there is an association between two categorical variables

Tests the hypothesis:

Ho: There is no association between row & column variables  
against

Ha: There is an association between row & column variables

Chi-square statistic has a degree of freedom  $(r-1)(c-1)$ ,  
where **r** is number of rows & **c** number of columns

## Chi-Square...

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

O: Observed cells

E: Expected cells

Expected value =  $\frac{(\text{Row total}) \times (\text{Column total})}{\text{Grand total}}$

For a 2X2, table

$$\chi^2 = \sum \frac{\left[ |ad - bc| - \frac{N}{2} \right]^2 N}{n_D n_{ND} n_E n_{NE}}$$

# Chi-Square...

## Calculation of Expected values

- Consider the following 2x2 table

| Head injury   | Wearing helmet |     | Row totals |
|---------------|----------------|-----|------------|
|               | Yes            | No  |            |
| Yes           | 17             | 218 | 235        |
| No            | 130            | 428 | 558        |
| Column totals | 147            | 646 | 793        |

# Chi-Square...

In the Example given above

- ❑ 235 out of 793 (29.6%) had head injury
- ❑ For those wearing helmet we expect 29.6% to have head injury =  $147 \times 0.296 = 43.6$
- ❑ 558 out of 793 (70.4%) had no head injury
- ❑ For those wearing helmet we expect 70.4% to have no head injury =  $147 \times 0.704 = 153.4$
- ❑ Similarly we can calculate expected frequencies for no wearing helmet.
- ❑ In general

Expected frequency of a cell =  $\text{Pr}(\text{event}) \times \text{number of trials}$

# Importance of Chi-square

If the calculated chi-square value is greater than the critical or  $P < 0.05$  we say that there is association

( $X^2$  Calculated  $>$   $X^2$  Tabulate/critical == Significant Association and vice versa )

Chi-square statistics tells only whether there is association. It doesn't tell us how much strong an association is.

# Measures of strength of association

□ Broadly classified in to two

1. Ratio measures: Relative Risk (RR/OR)
2. Difference measures: excess risk (AR/PAR)

# Measures of strength of association

□ Ratio or Relative Risk measures include: -

$$1 \text{ Rate Ratio} = \frac{\text{Incidence Rate in Exposed}}{\text{Incidence Rate in Unexposed}}$$

$$2 \text{ Risk Ratio} = \frac{\text{Commulative Incidence in Exposed}}{\text{Commulative Incidence in Unexposed}}$$

$$3 \text{ Prevalence Ratio} = \frac{\text{Prevalence in Exposed}}{\text{Prevalence in Unexposed}}$$

$$4 \text{ Prevalence Odds Ratio} = \frac{\text{Prevalence Odds in Exposed}}{\text{Prevalence Odds in Unexposed}}$$
$$= \frac{\text{Exposure Odds in Diseased}}{\text{Exposure Odds in non-diseased}}$$

## Relative risk (RR)

Expresses risk of developing a diseases in exposed group (a + b) as compared to non-exposed group (c + d)

$$\text{RR} = \frac{\text{Incidence (risk) among exposed}}{\text{Incidence (risk) among non-exposed}}$$

$$\text{RR} = \frac{a/(a+b)}{c/(c+d)}$$

## Interpretation of relative risk

What does a RR of 2 mean?

Risk in exposed =RRX Risk in non-exposed

RR of 2 means

Risk in exposed=2X Risk in non-exposed

- Thus a relative risk of 2 means the exposed group is two times at a higher risk when compared to non-exposed

## Strength of association

□ In general strength of association can be considered as:

1. High if  $RR \geq 3$
2. Moderate if RR is between 1.5 & 2.9
3. Weak if RR is between 1.2 & 1.4

# Relative Risk...

## **Example-1:** Rate Ratio [cohort study]

| HIV status | No dev.TB | p-Y observed | IR /1000 PYO | Rate Ratio |
|------------|-----------|--------------|--------------|------------|
| Negative   | 14        | 2054         | 0.0068       | 1          |
| positive   | 10        | 222          | 0.045        | 6.62       |

## **Interpretation**

Individuals who are HIV Positive have 6.62 fold risk of developing TB infection compared to those who are not positive in the cohort study.

# Relative Risk...

## Example-2

|                 | Develop<br>CHD | Do Not<br>Develop<br>CHD | Totals | Incidence<br>per<br>1000/yr |
|-----------------|----------------|--------------------------|--------|-----------------------------|
| Smokers         | 84             | 2916                     | 3000   | 28.0                        |
| Non-<br>smokers | 87             | 4913                     | 5000   | 17.4                        |

- ❖ Incidence in smokers =  $84/3000 = 28.0$
- ❖ Incidence in non-smokers =  $87/5000 = 17.4$
- ❖ Relative Risk =  $28.0/17.4 = 1.61$

## Interpretation

- ❖ Individuals who are smokers have 1.61 fold risk of developing CHD compared to those who are non-smokers in the cohort study.

# Relative Risk (RR)

Example-3: Prevalence Ratio (cross- Sectional study)

| <b>Breast feed for &gt;1 yr</b> | <b>Number with Asthma</b> | <b>Total</b> | <b>Prevalence of asthma</b> | <b>Prevalence ratio</b> |
|---------------------------------|---------------------------|--------------|-----------------------------|-------------------------|
| No                              | 14                        | 531          | 0.026                       | 1.37                    |
| Yes                             | 132                       | 6814         | 0.019                       | 1                       |

## Interpretation

Children > 1yr who do not have breastfeed had 37% increased risk of acquiring asthma compared to those who have breast feed. (N.B. it is for period prevalence study).

# Odds Ratio (OR)

- Odds Of Disease:
  - ❖ odds of having the disease if exposed
  - ❖ Ratio of diseased to non – diseased in exposed
  
- Odds Of Exposure:
  - ❖ Odds of having the exposure in disease
  - ❖ Ratio of exposed to non-exposed in those with disease

# Odds ratio...

## Odds Ratio

= Ratio of diseased to non-diseased in exposed =  
Ratio of diseased to non-diseased in unexposed

or

= Ratio of exposed to non-exposed in those with the disease  
Ratio of exposed to non-exposed in those without the disease

## Odds ratio.....

Odds ratio is the ratio of odds of exposure among diseased to odds of exposure among non-diseased

Odds of an event E is the ratio of probability of the event to its complement

$$\text{Odds} = P(E) / P(E') = P(E) / (1 - P(E))$$

# Odds ratio...

Odds of exposure among diseased= $a/c$

Odds of exposure among non-diseased= $b/d$

OR =  $\frac{\text{Odds of exposure among diseased}}{\text{Odds of exposure among non-diseased}}$

OR=  $(a/c)/(b/d)$

OR=  $ad/bc$  (it is also called cross-product ratio)

Interpretation of OR is the same as that of RR

# Odds ratio...

RR can be best estimated by OR if the following conditions are fulfilled

1. Controls are representative of general population
2. Selected cases are representative of all cases
3. The disease is rare

# Odds ratio...

Example-1: Prevalence odds ratio [cross- sectional]

| <b>Physical violence</b> | <b>Health poor</b> | <b>Health good</b> | <b>Odds</b> | <b>Odds ratio</b> |
|--------------------------|--------------------|--------------------|-------------|-------------------|
| No                       | 500                | 594                | 0.842       | 1                 |
| Yes                      | 677                | 424                | 1.597       | 1.9               |

## Interpretation

Individuals with physical violence have 1.9 fold risk of developing poor health than those without physical violence

N.B. (This is for point prevalence studies)

# Odds ratio...

Ex-2 prevalence odds in HIV behavioral survey, Asosa.

| Educational level | VCT willing | VCT unwilling | Odds ratio |
|-------------------|-------------|---------------|------------|
| Secondary         | 193         | 238           | 1          |
| Elementary        | 132         | 198           | 0.82       |
| Not literate      | 38          | 84            | 0.56       |

## Interpretation

1. Those elementary school students will have 18% reduction of VCT willing compared to those who are secondary school students.
2. Not literate individuals will have 44% Reduction of VCT willing compared to those who are secondary school students.

# Odds ratio...

## ❑ Advantages

- ✓ Estimated from all common study designs
- ✓ Can be compared from/between studies
- ✓ Used as a measure of strength of association.

## ❑ Disadvantage

- ✓ Poor guide for public health policy

# Attributable Risk (AR)

- ❑ AR indicates how much of the risk is due to /attributable/ to the exposure
- ❑ Quantifies the excess risk in the exposed that can be attributable to the exposure by removing the risk of the disease occurred due to other causes
- ❑  $AR = \text{Risk (incidence) in exposed} - \text{Risk (incidence) in non-exposed}$
- ❑  $AR = \{a/(a+b)\} - \{c/(c+d)\}$
- ❑ Attributable risk is also called risk difference/excess risk

## Interpreting AR

- ❑ What does attributable risk of 10 mean?
- ❑ 10 of the exposed cases are attributable to the exposure
- ❑ By removing the exposure one can prevent 10 cases from getting the disease

## Attributable risk percent (AR%)

Estimates the proportion of disease among the exposed that is attributable to the exposure

The proportion of the disease in the exposed that can be eliminated by eliminating the exposure

$$AR\% = \frac{\text{Risk in exposed} - \text{Risk in non-exposed}}{\text{Risk in exposed}} \times 100$$

## Interpretation of AR%

- ❖ What does AR% of 10% mean?
- ❖ 10% of the disease can be attributed to the exposure
- ❖ 10% of the disease can be eliminated if we avoid the exposure

# Example of AR

## Fast driving and Automobile Deaths

|       | Dead | Not dead | Total | Risk | RD   |
|-------|------|----------|-------|------|------|
| Fast  | 100  | 1900     | 2000  | 0.05 |      |
| Slow  | 80   | 7920     | 8000  | 0.01 | 0.04 |
| Total | 180  | 9820     | 10000 |      |      |

## **Population Attributable Risk (PAR)**

- ❖ Estimates the rate of disease in total population that is attributable to the exposure
- ❖  $PAR = \text{Risk in population} - \text{Risk in unexposed}$
- ❖  $PAR = AR \times \text{prevalence rate of exposure}$

# Population attributable risk percent (PAR%)

- ❑ Estimates the proportion of disease in the study population that is attributable to exposure and thus could be eliminated if the exposure were eliminated
- ❑  $PAR\% = \frac{\text{Risk in population} - \text{Risk in unexposed}}{\text{Risk in population}}$

# Example of PAR

## Fast driving

|      | Dead | Not dead |       | Risk  |
|------|------|----------|-------|-------|
| Fast | 100  | 1900     | 2000  | 0.050 |
| Slow | 80   | 7920     | 8000  | 0.010 |
|      | 180  | 9820     | 10000 | 0.018 |

$$\text{PAR} = 0.018 - 0.010 = 0.008$$

$$\text{PAR}\% = \frac{0.018 - 0.010}{0.018} \times 100 = 44\%$$

# Conclude

- ❑ 44% of driving-related deaths in population were presumably due to fast driving

# Possible outcomes in studying the relationship between exposure & disease

1. No association

$$RR=1$$

$$AR=0$$

2. Positive association

$$RR>1$$

$$AR>0$$

3. Negative association

$$RR<1 \text{ (fraction)}$$

$$AR<0 \text{ (Negative)}$$

## Risk Vs Preventive factors

- ❑ A *risk factor* is any factor positively associated with a disease ( $RR > 1$ ,  $AR > 0$ )
- ❑ It is associated with an increased occurrence of a disease
  
- ❑ A *preventive factor* is any factor negatively associated with a disease ( $RR < 1$ ,  $AR < 0$ )
- ❑ It is associated with a decreased occurrence of a disease
  
- ❑ Risk and preventive factors *may (not)* amenable to change (e.g. Smoking, age)

# 6. Epidemiological Surveillance

# LEARNING OBJECTIVES/ OUTCOMES

- ❑ At the end of this lesson you will be able to
  1. Define and differentiate surveillance Vs Survey
  2. Understand and use case definition
  3. State at least 5 uses of disease surveillance information.
  4. Identify some public sources of disease surveillance data.
  5. Evaluate the effectiveness of a disease surveillance system.

## What is epidemiologic surveillance?

- ❑ Surveillance is the continuous monitoring of disease and other health-related events relevant to the prevention and control of disease within the population.
- ❑ The ongoing systematic collection, analysis, and interpretation of data essential to the planning, implementation, and evaluation of public health practice, closely integrated with the timely dissemination of these data to those responsible for prevention and control.(CDC)

# Surveillance and Survey

## Surveillance

- It is relatively cheap (for health department). Can often use existing system and personnel
- Allow monitoring of trends of diseases over time
- Ongoing collection allows to use enough cases for study
- Quality control may be the major problem
- May not provide representative data

## Survey

- It is costly, needs to use by hiring trained once
- Represents only single point in time and informs little of anything about change in time
- More in-depth data could be collected
- More accurate assessment of true incidence and prevalence
- Can identify those which don't warrant medical care

# Purposes of surveillance

- The main purposes of surveillance are:
  - ❖ To enable the early recognition, investigation, and control of outbreaks.
  - ❖ To provide baseline information for priority setting, planning and evaluating disease control programs.
  - ❖ To provide information for understanding the distribution of disease by time, place, and person, so as to provide clues for the investigation of disease etiology, again for the purpose of more effective control.

# Uses of surveillance

- Disease surveillance information is useful for:
  - Estimating the magnitude of a problem
  - Determining the geographic distribution of illness
  - Portraying the natural history of a disease
  - Detecting epidemics or defining a problem
  - Generating hypotheses, stimulating research
  - Evaluating control measures
  - To identify changes in infectious agents, host and environmental factors
  - Detecting changes in health practices
  - Facilitating planning

# Types of surveillance

- Surveillance is based on two mechanisms for the detection of the disease
  - Passive case detection- this involves cases detected in the course of the normal operation of the health service, through the self-reporting of patients to the health institutions.
  - Active case detection- this involves an active search for cases, by special surveys or other methods outside of the routine health service activities.

## Types of surveillance...

- ❑ Based on these two mechanisms for the detection of diseases, we have three forms of surveillances:
  1. Passive surveillance- this is the mechanism for routine surveillance, based on passive case detection, and on the routine recording and reporting system.
  2. Active surveillance- this refers to any surveillance activity outside of the routine system, undertaken for the further investigation of diseases outbreaks and other occurrences of public health importance.
  3. Sentinel surveillance-using a pre-arranged sample of reporting sources to represent all conditions in a specified community

# Selection of Diseases for surveillance

- ❑ All diseases may not be included in surveillances:
- ❑ The importance of a health event to be included in surveillance system, it should be assessed through certain criteria

# Criteria of inclusion in surveillance

## 1. The current impact of the health event

- Having high incidence/ prevalence
- Mortality(overall and age-specific)
- Morbidity (hospitalization, disability, top morbidity)
- Severity (case fatality)
- Health care cost
- E.g. Malaria, pneumonia, TB, HIV/AIDS

# Criteria of inclusion in surveillance

2. Having Epidemic Potential  
(Eg. measles, cholera, meningitis)
3. Surveillance internationally required  
(Eg. plague, yellow fever, cholera)
4. Having available and effective control and prevention intervention  
(Eg. Schistosomiasis, onchocerciasis, trypanosomiasis)
5. Can easily be identified using simple case definitions
6. Having intervention program already found in the country (Eg. IMCI, EPI)

# Lists of priority diseases in Ethiopia (22)

## A. Epidemic Prone diseases

1. Cholera
2. Bloody diarrhea
3. Measles
4. Meningitis
5. Plague
6. Viral hemorrhagic fever
7. Yellow fever
8. Typhoid fever
9. Relapsing fever
10. Endemic typhus
11. malaria

## B. Diseases eradication targeted

1. Polio (AFP)
2. Dracunculosis (Guinea worm)
3. Leprosy
4. Neonatal tetanus

## C. Public health important diseases

1. Child pneumonia
2. Diarrhea in children
3. New AIDS cases
4. Oncocerciasis
5. STDs
6. TB
7. Rabies

## **Once diseases for surveillance are selected we should do the following items**

1. Case definition of diseases included in the surveillance
2. Determine the population under surveillance
3. Determine time period of data collection  
(immediately, weekly, monthly)
4. Determine source of data, who should report
5. Determine how data are handled  
(confidentiality)

# Case Definition

Epidemiologists ask,

1. Who gets the disease?
2. In what frequency is?
3. Is the frequency changing over time?
4. How does the frequency compare between population?"

□ Answering such questions requires:

- a case definition and
- estimating comparable disease frequency measures

## Case Definition....

- ❑ Just as a clinical diagnosis for an individual requires meeting specific clinical and laboratory criteria,
- ❑ Measuring disease frequency in populations requires prior stipulation of which clinical, laboratory, epidemiologic or quantitative criteria indicate the presence of the disease.
- ❑ Case definitions can include a degree of certainty (e.g. probable or confirmed, etc.) or specify the method used in assessing whether or not criteria met.

## Case Definition....

- ❑ Suppose you are asked to estimate the population prevalence of AIDS among Ethiopian adults or adolescent (aged  $>12$  years).
- ❑ How will you identify the adults/or adolescent who should be counted as cases of AIDS?
- ❑ What defines a ‘case’? The definition of a ‘case’ is critical in planning an epidemiologic investigation.

## Case Definition....

- ❑ The case definition must carefully formulate to:
  - meet objectives of the investigation,
  - allow valid comparisons with other study results
  
- ❑ In this example, it may be of interest to consider whether the proportion of adults or adolescent (aged >12 years) with AIDS has changed over a period.
  
- ❑ If the case definition changes significantly from one period to the next, comparisons with previous years are meaningless.

## Case Definition....

- ❑ The case definition must carefully formulate to meet objectives of the investigation, while also permitting valid comparisons with results from other studies.
- ❑ In this example, it may be of interest to consider whether the proportion of adults or adolescent (aged >12 years) with AIDS has changed over a period.
- ❑ If the case definition changes significantly from one period to the next, comparisons with previous years are meaningless.

## Case Definition....

- ❑ Suppose instead of estimating prevalence, the task is to define cases for a case-control study that is examining the risk from an exposure.
- ❑ If the case definition is broad, it will be easier to include prospective cases, hastening the study results.
- ❑ However, variability among the cases will be greater than if the case definition was narrower.
- ❑ A narrow case definition can slow the identification of sufficient numbers of cases but has the potential to increase precision.

# Case Definitions...

Cases of infectious disease categorized as follows:

- 1. Clinically compatible case:** is a clinical syndrome generally compatible with the disease, as described in the clinical description. It is a general clinical impression that this is a case of disease.
- 2. Confirmed case:** A case classified as confirmed for reporting purposes. The case meets established criteria.

## Case Definitions ...

- 3. Epidemiologically linked case:** It is a case with either of the following conditions.
- (a) the patient had contact with one or more persons, has or had the disease or has exposed to a point source of infection,  
  
i.e., a single source of infection, leading to a food borne-disease outbreak to which all confirmed case-patients were exposed,  
  
*and*
  - (b) transmission of the agent by the usual modes of transmission is plausible.

## Case Definitions ...

*A case considered epidemiologically linked to a laboratory-confirmed case;*

*“ if at least one case  
....in the chain of transmission  
.....is laboratory confirmed.*

## Case Definitions ...

### **4. Laboratory-confirmed case:**

a case that confirmed by one or more of the laboratory methods listed in the case definition under laboratory criteria for diagnosis.

Although other laboratory methods used in clinical diagnosis, only accept those listed as laboratory confirmation for national reporting purposes.

## Case Definitions ...

### 5. Probable case:

a case that classified as probable for reporting purposes.

Supportive or presumptive laboratory results:

*“specified laboratory results that are consistent with the diagnosis, yet do not meet the criteria for laboratory confirmation”.*

# Case Definitions ...

## 6. Suspected case:

a case that has a lower certainty; classified as suspected for reporting purposes.

*(Definitions originally published in MMWR 1997: 46(RR10); 1-55).*

# Case Definitions ...

## *Example 1: Measles (Rubella) case definition*

The CDC has established case definitions for public health reporting purposes.

The CDC clinical case definition requires at least 3 characteristics to be present in a clinical case,

*...rash, fever and cough, coryza or conjunctivitis*

A patient characterized clinically without any laboratory testing

## Case Definitions ...

### *Example 1: Measles (Rubella) case definition*

With laboratory results, if any of these four criteria met, the case of measles is laboratory confirmed.

1. positive serological test,
2. a rise in the measles antibody level,
3. detection of measles-virus-specific nucleic acid
4. isolation of the virus from the actual specimen.

## 2. Determine the population under surveillance

- ❑ A surveillance system remains effective when it is continuously assessed
  
- ❑ Periodically updating information about the catchment population is necessary
  
- ❑ Important target groups usually include:
  - Under-5 years children
  - Women of child bearing age
  - People living in refuge etc..
  - (prepare detailed demographic data)

### 3. Time period of data collection

- ❑ It is useful to identify problems and solve timely
  
- ❑ There are three periods of reporting
  1. Immediate reporting:
    - A. For disease that are not considered as endemic and are considered epidemic  
  
Eg, cholera, plague, viral hemorrhagic fever, polio, yellow fever
  
    - B. Suspected epidemic when a threshold is crossed

## 3. Time period of data collection

2. Weekly reporting:  
For epidemic-prone diseases

Eg, malaria, meningitis,

3. On monthly bases  
For routine surveillance

Eg, TB, leprosy, AIDS cases

# Steps of surveillance

## 1. Data collection:

Sources include

Mortality Report ( vital statistics)

Morbidity Report ( notify able diseases report, Hospital report, Lab report etc)

Epidemic Report

Reports of lab utilization

Reports of individual case investigation

Special surveys

Information on animal reservoir and vectors

Demographic data

Environmental data

# Steps of surveillance

## 2. Analysis:

include

1. Analyzing by time
2. Analyzing by place
3. Analyzing by person

## 3. Interpretation

## 4. Dissemination of surveillance data

## 5. Link to public health action

# **10. Evaluation of Evidence (Judgment of causality)**

# Learning Objectives

At the end of this unit students will be able

1. To differentiate association vs causation
2. To identify Possible explanations for observed association
3. To identify internal and external validity
4. To understand role of chance bias, & confounding
5. To understand controlling/ minimizing role of chance, bias, and confounding
6. To understand the Criteria to assess the strength of evidence for cause and effect relationship

# Association Vs Causation

The existence of an association doesn't itself constitute a proof of causation.

An observed association could be a fact or an artifact.

Hence, an association is a necessary but not a sufficient condition for causation.

# Possible explanations for observed association

1. Chance
2. Bias
3. Confounding
4. Reverse causation
5. Reciprocal causation
6. Cause-effect relationship

# Accuracy of measurement

Accuracy = Validity + Precision

Validity is the extent to which a measured value actually reflects truth

There are two types of validity

- Internal validity
- External validity

# Types of validity

## **Internal validity:**

Is the degree to which a measured value is true within the sample

## **External validity:**

Is the extent to which a measured value apply beyond the sample

This is related to generalizability

# Precision

Precision is the extent to which random error alters the measurement of effects

Threats to validity of study:

Random error (chance): is sampling error

Systematic error (bias): is error in the conduct of the study

# Judgment of causality

Judgment of causality has two steps

1. Check whether the observed association between exposure and disease is Valid (Rule out chance, bias and confounding)
2. Check whether the observed association is causal (Does the totality of evidence supports the findings)

# Role of chance

The role of chance as an alternative explanation for an association emerges from sampling variability

Evaluation of the role of chance is mainly the domain of statistics and involves

1. Test of statistical significance
2. Estimation of confidence interval

# 1. Test of statistical significance

P-value quantifies the degree to which chance accounts for observed association

P-value is the probability of obtaining a result at least as extreme as the observed by chance alone

$P < 0.05$  indicates statistical significance for medical research

## Test of statistical...

A very small difference may be significant if you have large sample

A large difference may not achieve statistical significance if you have small sample

One can't make a definite decision based on p-value only

## 2. Estimation of confidence interval

Confidence interval represents the range within which true magnitude of effect lies within a certain degree of assurance

It is more informative than p-value because it reflects on both the size of the sample and the magnitude of effect

## Role of bias

Bias is any systematic error in the design, conduct or analysis of an epidemiologic study that results in an incorrect estimate of association between exposure and disease

Unlike chance bias can't be statistically evaluated

There are two major types of bias

1. Selection bias
2. Information bias

# Selection bias

- Any systematic error that arises in the process of identifying the study population
- It affects the representativeness of the study
- It occurs when there is a difference between sample and population with respect to a variable
- Examples of selection bias:
  1. Diagnostic bias/misclassification bias
  2. Volunteer bias
  3. Berkson's bias
  4. Non-response bias
  5. Loss to follow-up bias

# Information/observation/bias

Any systematic error in the measurement of information on exposure or disease

Examples of information bias:

1. Interviewer bias/observer bias
2. Recall bias / Response bias
3. Social desirability bias
4. Placebo effect
5. Hawthorn/healthy workers effect

## Ways to minimize bias

1. Choose study design carefully
2. Choose objective rather than subjective outcomes
3. Blind interviewers whenever possible
4. Use close ended questions whenever possible
5. Collect data on variables you don't expect to differ between two groups

# Role of confounding

Confounding refers to the mixing of the effect of an extraneous variable with the effect of the exposure and disease of interest

## Characteristics of a confounding variable

1. Associated with disease in absence of exposure
2. Associated with exposure but not as a consequence of exposure
3. The frequency of the confounding variable vary between the groups that are compared

Example: In association between smoking and lung cancer alcohol drinking suspected as a confounding

# Effect of confounding

Totally or partially account for the apparent effect

Mask an underlying true association

Reverse the actual direction of association

# Control of confounding variables

During designing stage:

- Randomization
- Restriction
- Matching

During analysis stage

- Standardization
- Stratification/pooling
- Multivariate analysis

# Criteria to assess the strength of evidence for cause and effect relationship

Observational studies have many biases and confounding. Experimental studies if properly done can show cause-effect relationship. But they are not usually feasible due to ethical issues

In the absence of an experimental trail, the following criteria (Bradford Hill criteria) are used to assess the strength of evidence for a cause and effect relationship

# Criteria to assess the strength...

1. *Strength of association*: The stronger the association the more likely it is causal
2. *Consistency of association*: The more consistent, the more likely it is causal
3. *Specificity of association*: If single exposure linked to single disease more likely is causal
4. *Temporal relationship*: The exposure must come before the disease to be causal

# Criteria to assess the strength...

5. ***Dose-response relationship***: risk of disease increase with increasing exposure with factor
6. ***Biological plausibility***: Knowledge of association coherent with biology and descriptive epidemiology of the disease
7. ***Reversibility***: Eliminating the exposure should be followed by a decrease in the incidence rate of the disease

# 11. Screening

# LEARNING OBJECTIVES/OUTCOMES

**When you complete this unit, you will be able to:**

1. Define disease Screening
2. calculate and differentiate between sensitivity, specificity, positive and negative predictive values of a diagnostic test
3. Recognize considerations in the Establishment of Screening Recommendations and Programs
4. Recognize how to evaluate a screening program

# Screening

## **Definition:**

Screening refers to the presumptive identification of a disease/defect by application of tests, examinations or other procedures in apparently healthy people.

Screening is an initial examination

Screening is not intended to be diagnostic

## Aims of screening program

- Changing disease progression efficiently
- Altering natural course of disease
- Protecting society from contagious disease
- Allocating resources rationally
- Selection of healthy people for job
- Studying the natural history of disease

# Criteria for selecting diseases for screening

Severity- The disease should be serious

Treatment- Early treatment should be more beneficial

Prevalence- Pre-clinical Prevalence should be high

# Criteria for establishing screening program

1. The *problem* should have public health importance
2. There should be accepted *treatment* for positives
3. Diagnostic & treatment *facilities* should be available
4. Recognized *latent stage* in the time course
5. *Test* is acceptable, reliable & valid
6. *Natural history* of the disease well understood
7. *Case-finding* is economical and continuous

## Screening tests

The performance of a screening test is evaluated against a diagnostic test in 2X2 table

|                |                | Diagnostic test |                | Total |
|----------------|----------------|-----------------|----------------|-------|
|                |                | D <sup>+</sup>  | D <sup>-</sup> |       |
| Screening test | T <sup>+</sup> | a               | b              | a+b   |
|                | T <sup>-</sup> | c               | d              | c+d   |
| Total          |                | a+c             | b+d            | n     |

## Definitions of cells

True positives (a): Diseased identified by test as diseased

False positives (b): Disease free falsely labelled as disease

False negatives (c): Diseased falsely labelled as disease free

True negatives (d): Disease free identified as free

## Definition of totals

$D^+$  (a+c): total subjects with a disease

$D^-$  (b+d): total subjects without disease

$T^+$  (a+b): total test positives

$T^-$  (c+d): total test negatives

# Test performance

## Validity of a test

The ability of a test to differentiate correctly those who have the disease and those who don't

- It is a function of sensitivity and specificity

# Test performance

## A. *Sensitivity of a test*

- The ability of a test to correctly identify those who have the disease
- The probability that a diseased individual will have a positive test result
- The proportion of people with a disease who have a positive test result
- True positive rate (TPR)

# Test performance

## A. *Sensitivity of a test (TPR)*

$$\text{Sensitivity} = \frac{\text{True positives}}{\text{True positives} + \text{false negatives}} = \frac{\text{True positives}}{\text{All persons with the disease}} \times 100$$

$$= \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{a}{(a + c)}$$

$$\text{False negative rate (FNR)} = P(T^- | D^+)$$

$$= \frac{\text{FN}}{\text{FN} + \text{TP}} = \frac{c}{(a + c)}$$

$$\text{Sensitivity} = 1 - \text{FNR}$$

# Test performance

## *B. Specificity of a test*

- The ability of a test to correctly identify those who don't have the disease
- The probability that a disease-free individual will have a negative test result
- The proportion of people without the disease who have a negative test result
- True negative rate (TNR)

# Test performance

## *B. Specificity of a test*

$$\text{Specificity} = \frac{\text{True negatives}}{\text{True Negatives} + \text{false Positives}} = \frac{\text{True negatives}}{\text{All persons with out the disease}} \times 100$$

$$= \frac{\text{TN}}{\text{TN} + \text{FP}} = \frac{d}{(d + b)}$$

False positive rate (FPR) =  $P(T^+ | D^-)$

$$= \frac{\text{FP}}{\text{FP} + \text{TN}} = \frac{b}{(b + d)}$$

*Specificity = 1 - FPR*

# Test performance

## Predictive value of a test

The ability of a test to predict the presence or absence of disease

Two types: Positive & negative predictive values

# Test performance

## A. *Predictive value positive (PVP)*

- The ability of a test to predict the presence of disease among who test positive
- The probability that a person with a positive test result has a disease
- The proportion of diseased individuals in a population with a positive test result
- Prior/post-test probability for true positive test

$$\begin{aligned} \text{PVP} &= P(D^+ | T^+) \\ &= \mathbf{a/(a+b)} \end{aligned}$$

# Test performance

## A. *Predictive value negative (PVN)*

- The ability of a test to predict the absence of disease among who test negative
- The probability that a person with a negative test result is disease-free
- The proportion of disease-free individuals in a population with a negative test result
- Prior/post-test probability for true negative test

$$\begin{aligned} \text{PVN} &= P(D^- | T^-) \\ &= d / (c + d) \end{aligned}$$

# Test performance

- **Bayes' theorem**

$$PVP = \frac{\text{sensitivity} \times \text{prevalence}}{(\text{Sensitivity} \times \text{prevalence}) + (1 - \text{specificity}) \times (1 - \text{prevalence})}$$

$$PVN = \frac{\text{sensitivity} \times (1 - \text{prevalence})}{[(\text{Specificity} \times (1 - \text{prevalence})) + [(1 - \text{sensitivity}) \times \text{prevalence}]}$$

# Test performance

## Prevalence of a disease

- The proportion of individuals with a disease
- Prior/pre-test probability of a disease

$$\begin{aligned}\text{Prevalence} &= P(D^+) \\ &= (a+c)/n\end{aligned}$$

## Yield of a test

- Proportion of cases detected by the screening program
- $$\text{Yield} = a/n$$
- Important for program evaluation (cost-effectiveness)

# Multiple testing

- Multiple tests are commonly done in medical practice.
- The choices depend on
  - cost,
  - invasiveness,
  - volume of test,
  - presence and capability of lab infrastructure,
  - urgency, etc.
- Can be done sequentially or simultaneously.

# Multiple testing

## Serial testing:

- Tests are administered sequentially
- All positive indicates disease
- Results in
  - Lower sensitivity
  - Increased specificity
  - Increased PVP

## Multiple testing

### Serial testing:

- Net Sensitivity in a Two-Stage Screening when Test + in the First Test are Re-Screened,

$$P(A \cap B) = P(A) \times P(B)$$

$$\text{Net sensitivity} = \text{Sensitivity 1} \times \text{Sensitivity 2}$$

- Net Specificity in a Two-Stage Screening when Test + in the First Test are Re-Screened

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$\text{Net specificity} = \text{Spec1} + \text{Spec2} - (\text{Spec1} \times \text{Spec2})$$

# Multiple testing

## Parallel testing:

- Tests are given concurrently
- At least one positive indicates disease
- Results in
  - Greater sensitivity
  - Increased PVN
  - Decreased specificity

# Multiple testing

## Parallel testing:

### – Net Sensitivity:

- When two tests are used simultaneously, disease positives are defined as those who test positive by either one test or by both tests.
- We use the addition rule of probability to calculate the net sensitivity.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$\text{Net sensitivity} = \text{Sens 1} + \text{Sens 2} - (\text{Sens 1} \times \text{Sens 2})$$

# Multiple testing

## Parallel testing:

### – Net Specificity:

- When two tests are used simultaneously, disease negatives are defined as those who test negative by both tests.
- We use the multiplication rule of probability to calculate the net specificity.

$$P (A \cap B) = P (A) \times P (B)$$

$$\text{Net specificity} = \text{specificity test 1} \times \text{specificity test 2}$$

# Hypothetical Two-Stage Screening

Assume: Disease Prevalence = 5%,  
Population = 10,000

TEST 1 (Blood Sugar)

Sensitivity = 70%  
Specificity = 80%

|              |   | DIABETES |      |        |
|--------------|---|----------|------|--------|
|              |   | +        | -    |        |
| TEST RESULTS | + | 350      | 1900 | 2250   |
|              | - | 150      | 7600 | 7750   |
|              |   | 500      | 9500 | 10,000 |

Only Pos. Test 1  
are given Test 2

# Hypothetical Two-Stage Screening (cont.)

TEST 2 (Glucose Tolerance Test)

Sensitivity = 90%

Specificity = 90%

|              |   | DIABETES |      |      |
|--------------|---|----------|------|------|
|              |   | +        | -    |      |
| TEST RESULTS | + | 315      | 190  | 505  |
|              | - | 35       | 1710 | 1745 |
|              |   | 350      | 1900 | 2250 |

$$\text{Net Sensitivity} = 315/500 = 63\%$$

$$\text{Net Specificity} = \frac{7600 + 1710}{9500} = 98\%$$

# Reliability of a test

- Ability of a test to give consistent results up on repeated measurements
- Two major factors affect reliability
  - Method variation
  - Observer variation
    - Inter-observer variation
    - Intra-observer variation

# Reliability of a test

Reliability can be classified as:

- Internal reliability
  - Internal consistency reliability
- External reliability
  - Alternate test reliability
  - Test-retest reliability

# Evaluation of a screening program

Evaluation of a screening program involves consideration of two issues

1. **Feasibility:** Determined by acceptability of the screening program
2. **Effectiveness:** Determined by the outcome of the screening program

## In general

In general, a screening test should be

- Reliable & valid
- Sensitive & specific
- Simple & acceptable
- Effective & efficient

# Example

|                | Reference test |          |          |       |
|----------------|----------------|----------|----------|-------|
|                |                | positive | negative | total |
| Screening test | +ve            | 1555     | 988      | 2543  |
|                | -ve            | 514      | 1394     | 1908  |
|                | Total          | 2069     | 2382     | 4451  |

1. What is the sensitivity of the test?
2. What is the specificity of the test?
3. What is the percent false positive?
4. What is the percent false negative?
5. What is the PVP of the test?
6. What is the PVN of the test?
7. What is the Yield of the test?

# **12. OUTBREAK INVESTIGATION AND MANAGEMENT**

# LEARNING OBJECTIVES/ OUTCOMES

- ❑ At the end of this lesson you will be able to
  1. use common terms in outbreak investigation appropriately
  2. develop an outbreak investigation plan
  3. describe a potential outbreak with regard to person, place and time.
  4. construct and interpret an epi-curve to describe the course of an outbreak
  5. practically use an outbreak management principles

# DEFINITION

- ❑ *Epidemic* is “The unusual occurrence in a community of disease, specific health related behavior, or other health related events clearly in excess of expected occurrence”
- ❑ **Outbreak** is an epidemic with a brief occurrence in terms of short duration and limited geographical population

# Types of Outbreak

- ❑ Two principal types are well known; these are the
  - 1.common source and
  - 2.propagated/progressive.
- ❑ The two types can be distinguished by plotting an **epidemic curve**.
- ❑ An epidemic, which shows the features of both types, is referred as **mixed**.

# Common source epidemic

- ❑ Caused by exposure of a group of people to a common noxious influence, such as an infectious agent or a toxin.

## Three Types of Common Source Epidemics

### ❑ **point source epidemic**

- ❖ simultaneous exposure to agent
- ❖ All exposed diseased within one incubation period
- ❖ Suggested by Rapid rise & fall of an epidemic curve
- ❖ This is called **long normal distribution**.

# Common source ...

## Three Types of Common Source Epidemics

### □ Continuous common source epidemic

- ❖ Has mainly wide peak in the epidemic curve,
- ❖ has a range of exposures &
- ❖ A range of incubation periods.

### □ Intermittent common source

- ❖ results in particular patterns of the epidemic curve that reflects the intermittent nature of the epidemic.

## propagated/progressive epidemics.

- ❑ Outbreak of this type can occur through direct person-to-person transmission or the transmission could pass through a vector from infected to healthy person.
- ❑ This kind of epidemic may have a chance to traverse a geographic boundary from one place to other
- ❑ More difficult both to investigate and manage/control

# Steps in epidemic investigation

## 1. Prepare for fieldwork

### A. Investigation related preparation.

- Specific knowledge
- Supplies and equipment
- Literature review

### B. Administration related

- Observe all administrative procedures such as transportation and personal.

### C. Consultation

- Clarify your and your team role in the field.
- Identify local contacts at the site.

## Steps ....

### 2. Verify the existence of an epidemic

#### 2.1. Verify the diagnosis

##### □ Case definition

- Is a standard set of criteria for deciding whether an individual should be classified as having the health condition of the interest.
- **Types of the case definition:-**
  - I. Confirmed/definite-a case with lab. Verification.
  - II. Probable –a case with typical clinical features of the disease without lab. confirmation.
  - III. Possible-a case presented with fewer of the typical clinical features.

## Steps ...

### 3. Describe the epidemic with respect to time, place and person.

- **Epidemic curve**-plots the cases by the time of onset and provides a time frame for the outbreak investigation.
- **Spot map**-plots cases by location and shows the geographic spreads of cases.
- **Attack rates**-calculate rates of illness in population at risk

# Steps ....

## 4. Formulate and test hypothesis

- Formulate hypothesis based on your characterization of the epidemic by time, place and person.
  
- The hypothesis should address:-
  - ✓ The source of the agent
  - ✓ The mode of transmission
  - ✓ The exposure that cause the disease.

## Steps ...

- 5. Determine the type of epidemic**
- 6. Define the population at risk.**
- 7. Search for additional case.**
- 8. Analyze the data.**
- 9. Make a decision on the hypothesis tested**
- 10. Intervention and follow up.**
- 11. Report of the investigation**

# MANAGING OUTBREAK/EPIDEMIC

## 1. Measures directed against the reservoir

- Domestic animals
    - Immunization
    - Distraction of infected animals
    - Testing of herds.
  - Wild animals
    - Post-exposure prophylaxis
  - Humans
    - Removal of the focus of infection.
    - Isolation of infected person
    - Treatment
    - Disinfection of contaminated object.
- ☐ **Quarantine** –is the limitation of freedom of movement of apparently healthy persons or animals who have been exposing to a case of infectious diseases.

# MANAGING ...

## **2. Measures that interrupt the transmission of organisms**

- ❖ Purification of water.
- ❖ Pasteurization of milk.
- ❖ Inspection procedures designed to ensure safe food supply.
- ❖ Improve housing condition.

## **3. Measures that reduce host susceptibility.**

- ❖ Active immunization.
- ❖ Passive immunization.
- ❖ Chemoprophylaxis.

**Thank you!**