
Welcome

Epidemiology

FIRST

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

DEFINITIONS

- ”the study of the distribution and determinants of health related states and events in specific populations and the application of the study of control health problems"

(John.M.Last 1988).

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 & promotion of health

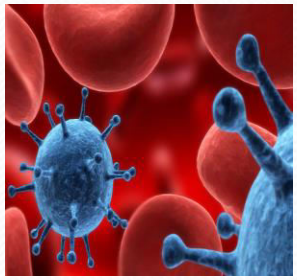
Epidemiology is a science and practice

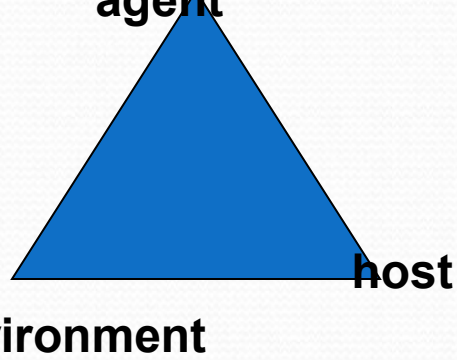
Epidemiology is an applied science

SECOND

BASIC CONCEPTS IN EPIDEMIOLOGY

- agent
- host
- environment





BASIC CONCEPTS IN EPIDEMIOLOGY

- **Agent:** an animate or inanimate factor that must be present or lacking for a disease or condition to develop.
- **Host:** a living species (human or animal) capable of being infected or affected by an agent.
- **Environment**: all that is internal or external to a given host or agent and that is influenced and influences the host and/or agent.

- Third

History…

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

Epidemiological thought emerged in 460 BC

Epidemiology flourished as a discipline in 1940s

Four

AIMS OF EPIDEMIOLOGY

According to International Epidemiological Association(IEA), epidemiology has three main aims,

- To describe the distribution and magnitude of health and disease problems in human populaions.
- To identify the etiological factors or risk factors in the pathogenicity of disease.
- To provide the data essential to the planning , implementation and evaluation of services.

- Five

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

contd

- It is a concerned with the systematized study of :
- 1 . Whole population in their living &working environment
- 2. Factors that determine a state of health & disease
- 3. Pattern of health as well as pattern of illness
- 4. Measures of prevention & control

Six

Purpose/use 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

seven

terminologies used in epidemiology

Frequency:

- relationship between the number of cases of disease and the size of the population

Determinants:

- causes and other factors that influence the occurrence of disease and other health-related events

Public Health Surveillance

- ongoing, systematic collection, analysis, interpretation, and dissemination of health data to help guide public health decision making and action

- Case Definition

  a set of standard criteria used for classifying whether a person has a particular disease, syndrome, or other health condition

- Descriptive Epidemiology

  Type of epidemiology which only covers time place and time to describe an outbreak rather than case definition, person, place, time, and causes/risk factors/modes of transmission

- Analytic Epidemiology

  The type of epidemiology turned to to test hypotheses formed with information acquired through descriptive epidemiology

- Experimental studies

  Form of analytic epidemiology which tests hypotheses in a very controlled environment

- Observational studies

  Type of analytic study in which the epidemiologist simply observes the exposure and disease status of each participant. (Cohort, case-control, and cross-sectional)

- Cohort Study

  Epidemiologist tracks whether or not the participant is exposed and then tracks them to see if they develop the disease. It is the most reliable form of analytic study, but most expensive and time-consuming.

- prospective study

  A study which monitors exposure first, then looks forward to see if disease is developed.

- retrospective study

  starts with people already diseased, then traces back to see if they had been exposed

- Case-control study

  enroll people with disease and use people without the disease. they then compare previous exposures between the two groups

- cross-sectional study

  the least expensive and least effective analytic study

- Epidemiological triad

  agent host environment

- agent

  A microbe which causes disease

- Host
  the human who develops the disease

- environment
  extrinsic factors that affect the agent and the opportunity for exposure

- Incubation period
  the stage of subclinical disease extending from the time of exposure to onset of disease symptoms for infectious diseases

- Latency period
  the stage of subclinical disease extending from the time of exposure to onset of disease symptoms for chronic disease

- Spectrum of disease

  the disease process resulting in illness which is mild, severe, or fatal and eventually ends in recovery, disability, or death

- Infectivity

  the proportion of exposed persons who become infected

- pathogenicity

  the proportion of infected individuals who develop clinically apparent disease

- virulence

  the proportion of clinically apparent cases that are severe or fatal

- carriers
  persons who are infectious but have subclinical disease

- reservoir
  the habitat in which the agent normally lives, grows, and multiplies

- portal of exit
  the path by which a pathogen leaves its host

- mode of transmission
  the way through which an agent is transmitted to its host

- portal of entry
  the manner in which the pathogen enters a susceptible host

- direct transmission
  an infectious agent is transferred from a reservoir

- direct contact
  skin to skin contact

- droplet spread
  spray with relatively large, short-range aerosols produced by sneezing, coughing, or even talking

- indirect transmission
  the transfer of an infectious agent from a reservoir to a host by suspended air particles, inanimate objects, or vectors

- airborne
  transmission occurs when infectious agents are carried by dust or droplet nuclei suspended in air

- herd immunity
  suggests that if a high enough proportion of individuals in a population are resistant to an agent, then those few who are susceptible will be protected by the resistant majority since the pathogen will be unlikely to "find those few susceptible individuals

- endemic level
  the amount of a particular disease that is usually present in the community

- sporadic
  a disease that occurs infrequently and irregularly

- endemic
  the constant presence and or usual prevalence of a disease or infectious agent in a population within a geographic area

- common source outbreak
  an outbreak in which a group of persons are all exposed to an infectious agent or toxin from the same source

- point source outbreak
  if the group is exposed over a relatively brief period, so everyone who becomes ill does so in on incubation period, then the common source outbreak is further classified as a point source outbreak

- ratio

  the relative magnitude of two quantities or a comparison of any two values

- proportion

  the comparison of a part to the whole

- rate

  a measure of the frequency with which an event occurs in a defined population over a specified period of time

- incidence rate

  conveys a sense of speed with which disease occurs in a population, and seems to imply that this pattern has occurred and will continue to occur for the foreseeable future

- attack rate

  the proportion of the population that develops illness during an outbreak

- case fatality rate

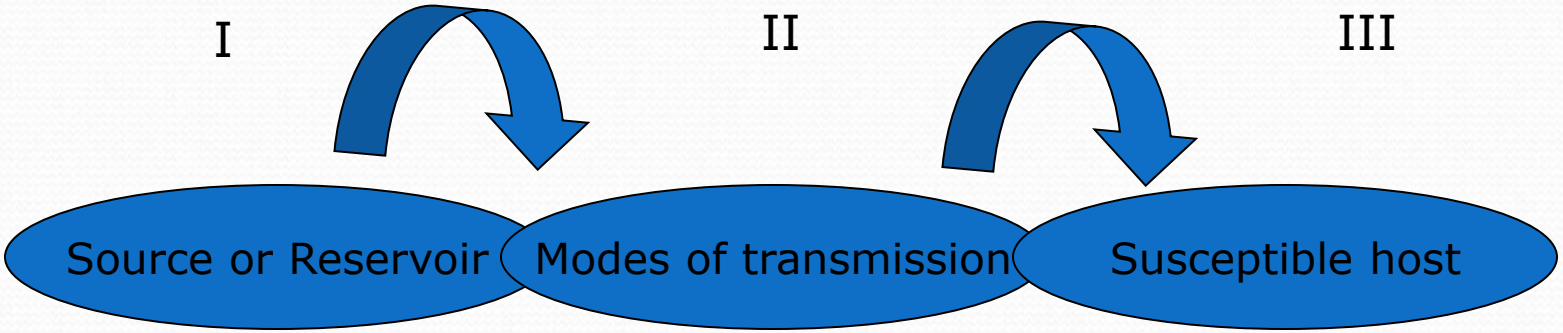
  the proportion of persons with the disease who die from it

- prevalence rate

  the proportion of the population that has a health condition at a point in time

- Dynamics of disease transmission

Dynamics of disease Transmission (Chain of Infection)

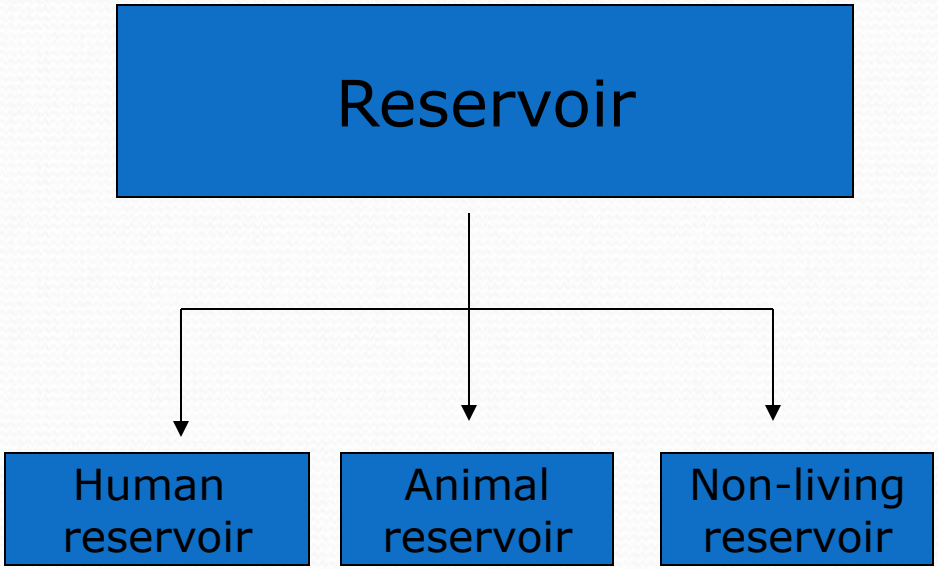


(I): Source or Reservoir

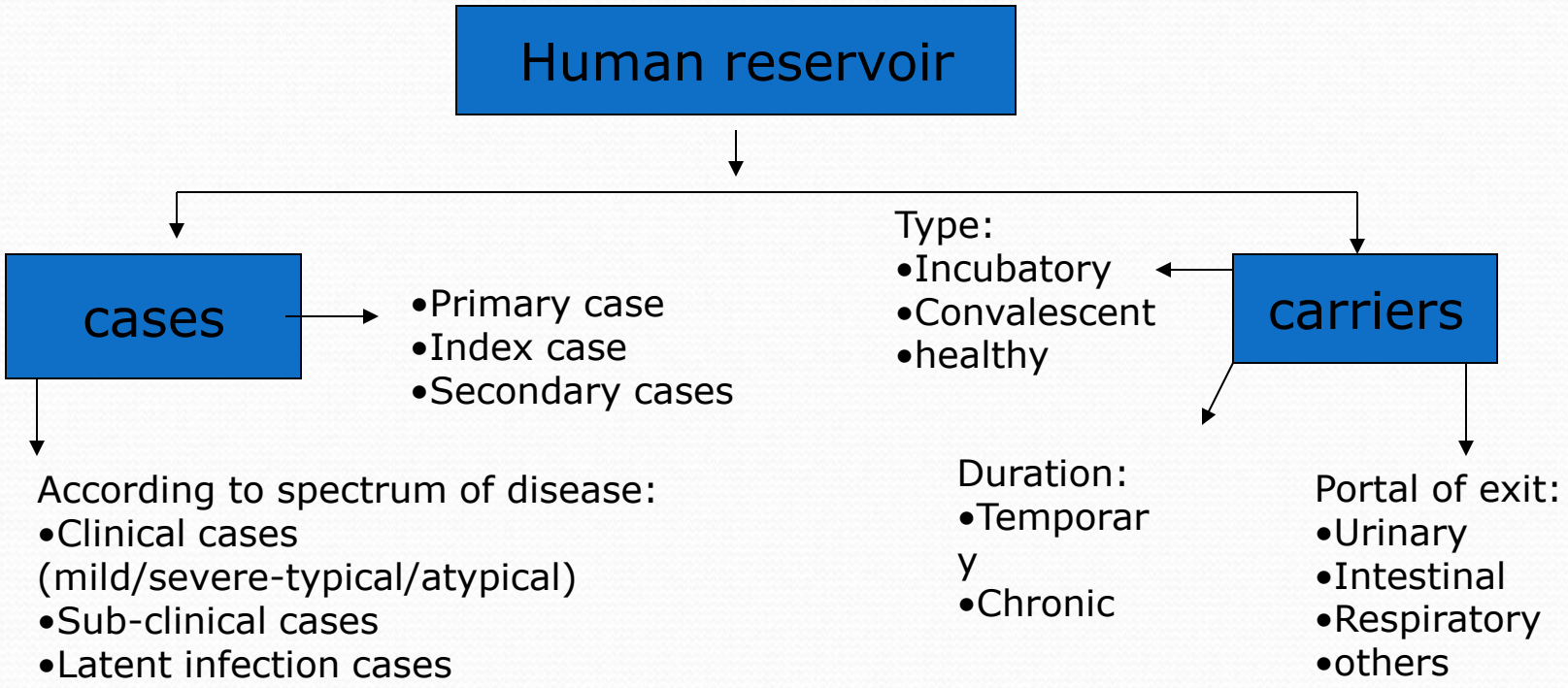
- The starting point for the occurrence of a communicable disease is the existence of a reservoir or source of infection.

- The source of infection is defined as "the person, animal, object or substance from which an infectious agent passes or is disseminated to the host (immediate source). The reservoir is "any person, animal, arthropod, plant, soil, or substance, or a combination of these, in which an infectious agent normally lives and multiplies, on which it depends primarily for survival, and where it reproduces itself in such a manner that it can be transmitted to a susceptible host. It is the natural habitat of the infectious agent."

Types of reservoirs



Human reservoir



Cases

- A case is defined as “a person in the population or study group identified as having the particular disease, health disorder, or condition under investigation”

Carriers

- It occurs either due to inadequate treatment or immune response, the disease agent is not completely eliminated, leading to a carrier state.

- It is "an infected person or animal that harbors a specific infectious agent in the absence of discernible (visible) clinical disease and serves as a potential source of infection to others.

- Three elements have to occur to form a carrier state:

1. The presence in the body of the disease agent.
2. The absence of recognizable symptoms and signs of disease.
3. The shedding of disease agent in the discharge or excretions.

Animal reservoirs

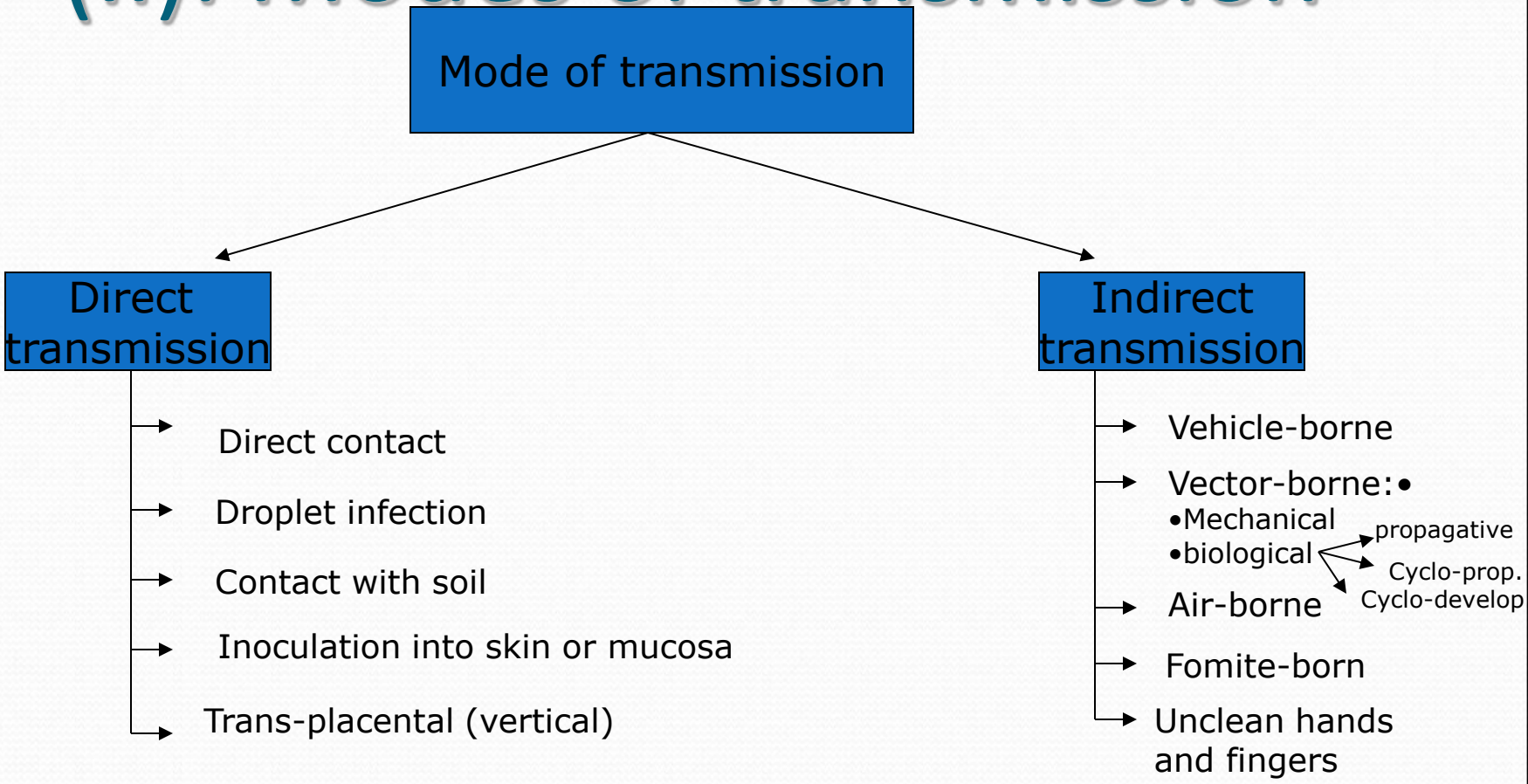
- Zoonosis is an infection that is transmissible under natural conditions from vertebrate animals to man, e.g. rabies, plague, bovine tuberculosis…..
- There are over a 100 zoonotic diseases that can be conveyed from animal to man.

Reservoir in non-living things

- Soil and inanimate matter can also act as reservoir of infection.

- For example, soil may harbor agents that causes tetanus, anthrax and coccidiodomycosis.

(II): Modes of transmission



(III): Susceptible host

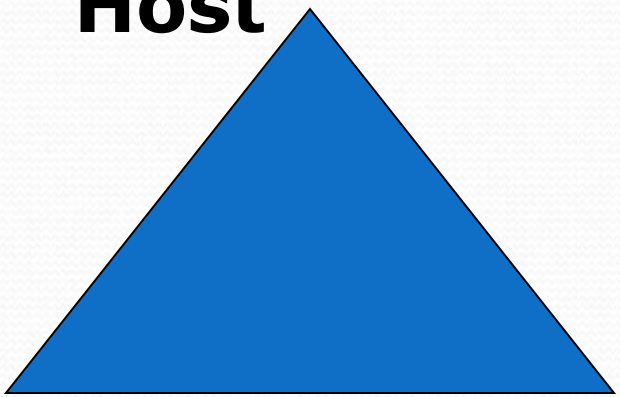
- An infectious agent seeks a susceptible host aiming "successful parasitism".

- Four stages are required for successful parasitism:
  1. Portal of entry
  2. Site of election inside the body
  3. Portal of exit
  4. Survival in external environment

Epidemiologic triad

Host

- **Demographic characteristics**
- **Biological characteristics**
- **Socioeconomic characteristics**



Agent

- **Biological agents**
- **Physical agents**
- **Chemical agents**
- **Nutrient agents**
- **Mechanical agents**
- **Social agents**

Environment

- **Physical environment**
- **Biological environment**
- **Social environment**

Types of Epidemiology

Two major categories of Epidemiology

1. *Descriptive Epidemiology*

Defines ***frequency*** and ***distribution*** of diseases and other health related events

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

Types…

2. *Analytic Epidemiology*

Analyses ***determinants*** of health problems

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

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 science

2. Communicable disease Epidemiology

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?

Ninetieth century theories

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

Theories…

Twentieth century theories

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

Necessary Vs Sufficient

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 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

3. ***Wheel model***

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

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 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***

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

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

Possible outcomes of exposure to an infectious agent

Infection: invasion & multiplication in the host

Infectivity: the proportion of exposed who becomes infected

Infection rate= Infected/exposed

Disease: A clinically apparent infection

The agent…

Pathogenicity:the proportion of infected who develop clinical disease

Clinical-to-Subclinical ratio

Virulence: the proportion clinical cases resulting in severe clinical disease

Case fatality & hospitalization rate

Immunogenecity:the infection's ability to produce specific immunity

Reservoir Vs Carrier

Reservoir

An organism or habitat in which an infectious agent normally lives, transforms, develops and/or multiplies

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

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

Modes of disease…

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

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

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

- Pre-patent period
  - When should we investigate?
- Incubation period
  - When was time of exposure?
- Communicable period
  - When should we take care of infectiousness?
- Latent period
  - When would relapse occur?
- Convalescent period
  - When,after recovery an individual becomes non-infectious?
- Generation time
  - When is the maximum risk for contacts?

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

3. Measures of disease occurrence

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

- Numbers
- Ratios
- Proportions
- Rates

Descriptors

Numbers: Use of actual number of events

e.g 100 cases of TB in community A

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

e.g Ratio of TB cases in community A to B is 1:10

Descriptors

Proportions: a ratio in which the numerator is included in the denominator

e.g proportion of TB cases in community A is 10%

Rates: a proportion with time element

It measure the occurrence of an event overtime

e.g U5 measles cases in 2000/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?

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

Morbidity rates

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

Two common morbidity rates

Incidence rates(Cumulative incidence, incidence density)

Prevalence (Period prevalence, point prevalence)

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

Cumulative incidence

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

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

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

Point prevalence rate

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

Point prevalence rate=All persons with a disease at a point in time/Total population

It is not a rate, but a true proportion

Period prevalence rate

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

Period prevalence rate=All persons with a disease overtime period/Average(mid-year)population in the same period

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

Relationship between prevalence & incidence rates

Prevalence α incidence

Prevalence α Average duration

Prevalence α Incidence X Average duration

An increase in prevalence rate may not necessarily be due to an increase in incidence rate, it could due to an increase in average duration of a disease due to decrease in death and/or recovery rates

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 ration

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

4. Measures of association

2X2 table

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

Cells

A= Exposed, and diseased

B= Exposed, Not diseased

C= Not exposed, diseased

D= Not exposed, Not diseased

Totals

Marginal totals

a+b= Exposed

c+d= Non-exposed

a+c= Diseased

b+d= Non-diseased

Grand total

n = a+b+c+d

Chi-square statistics

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

Ho: There is no association between row & column variables

Ha: There is an association between row and 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…

$$X^2 = \Sigma \frac{(O - E)^2}{E}$$

O: Observed cells

E: Expected cells

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

For a 2X2, table

$$X^2 = \frac{(/ad-bc/-n/2)^2 n}{(a+b)(a+c)(c+d)(b+d)}$$

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

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

Relative risk (RR)

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

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

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

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:

High if RR≥3

Moderate if RR is between 1.5 & 2.9

Weak if RR is between 1.2 & 1.4

Odds ratio (OR)

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

$$Odds = P(E)/P(E') = P(E)/(1-P(E))$$

Odds ratio…

Odds of exposure among exposed=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

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= Risk (incidence) in exposed- Risk (incidence) in non-exposed

AR= {a/(a+b)} / {c/(c+d)))}

Attributable risk is also called risk difference

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

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

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

Population Attributable Risk (PAR)

Estimates the rate of disease in total population that is attributable to the exposure

PAR = Risk in population – Risk in unexposed

PAR = ARX 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

$$PAR\% = \frac{\text{Risk in population} - \text{Risk in unexposed}}{\text{Risk in population}}$$

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 (fraction)

   AR<0 (Negative)

Risk Vs Preventive factors

A ***risk factor*** is any factor positively associated with a disease (RR>1)

It is associated with an increased occurrence of a disease

A ***preventive factor*** is any factor negatively associated with a disease (RR<1)

It is associated with a decreased occurrence of a disease

Risk and preventive factors ***may (not)*** amenable to change (e.g. Smoking, age)

5. Evaluation of Evidence (Judgment of causality)

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 is 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
2. Volunteer bias
3. Non-response bias
4. 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

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 asses 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 asses the strength of evidence for a cause and effect relationship

Criteria to asses 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

4. ***Temporal relationship***: The exposure must come before the disease

Criteria to asses 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

6. Epidemiologic Study Designs

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 communities
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

Cross-sectional studies

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

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

For factors that remain unaltered overtime (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…

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

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
   - Establish strict diagnostic criteria
   - 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

Case-control…

II. Based on matching

Matching: Relating cases and controls with respect to certain variable

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

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

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

Experimental…

Three different ways of classifying intervention studies

I. Based on population studies
- Clinical trial: on patients in clinical settings
- Field trial: on healthy people in the field
- Community trial: on the community as a whole

II. Based on design
- Uncontrolled trial: no control (self-control)
- Non-randomized controlled: allocation not random
- Randomized control: Allocation random

Experimental…

III. Based on objective

- Phase I: to determine toxic effect
- Phase II: to determine toxic effect
- Phase III: to determine applicability

Challenges in intervention studies

- Ethical issues
  - Harmful treatment shouldn't be given
  - Useful treatment shouldn't be denied

Experimental…

- 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

Experimental…

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

7. Screening

Screening

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

The ***problem*** should have public health importance

There should be accepted ***treatment*** for positives

Diagnostic & treatment ***facilities*** should be available

Recognized ***latent stage*** in the time course

Test is acceptable, reliable & valid

Natural history of the disease well understood

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		
		D^+	D^-	
Screening test	T^+	a	b	a+b
	T^-	c	d	c+d
		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 negatves

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

$$\text{Sensitivity (TPR)} = P(T^+ | D^+)$$

$$= TP/(TP + FN)$$

$$= \mathbf{a/(a+c)}$$

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

$$= \mathbf{c/(a+c)}$$

Sensitivity= 1- 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

Specificity (TNR)= $P(T^- | D^-)$

$= TN/(TN + FP)$

$= \mathbf{d/(d+b)}$

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

$= \mathbf{b/(b+d)}$

Specificity= 1- FPR

Test performance

Predictive value of a test

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

Two type: 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

$$PVP = P(D^+ \mid T^+)$$

$$= \mathbf{a/(a+b)}$$

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

$$PVN = P(D^- \mid T^-)$$

$$= d/(\mathbf{c+d})$$

Test performance

Prevalence of a disease

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

Prevalence = $P(D^+)$

$= (a+c)/n$

Yield of a test

Proportion of cases detected by the screening program

Yield = a/n

Multiple testing

Parallel testing:

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

Multiple testing

Serial testing:

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

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

Thank you!

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)