EPIDEMIOLOGIC STUDIES

Types of Epidemiologic Studies

- **Descriptive**
  - Correlational/ecological
  - Case report or case series
  - Cross-sectional
- **Analytic**
  - **Observational studies**
    - ✓ Cohort
    - ✓ Case control
  - **Interventional / Experimental**

Descriptive studies

- ✓ are mainly concerned with the distribution of diseases with respect to time, place and person
- ✓ provide useful information for health managers to allocate resource and to plan effective prevention programmes
- ✓ generate epidemiological hypothesis, an important first step in the search for disease determinants or risk factors
- ✓ can use information collected routinely which are readily available in many places. So generally descriptive studies are less expensive and less time-consuming than analytic studies.
- ✓ are the most common type of epidemiological design strategies in medical literature

Correlational or Ecological

- ✓ Uses data from entire population to compare disease frequencies - between different groups during the same period of time, or in the same population at different points in time
- ✓ Does not provide individual data, rather presents average exposure level in the community
- ✓ Cause could not be ascertained
- ✓ Correlation coefficient (r) is the measure of association in Correlational studies. It is important to note that positive association does not necessarily imply a valid statistical association.

  e.g. Hypertension rates and average per capita salt consumption compared between two communities

Correlational or Ecological cont'd….

Strength:

- can be done quickly and inexpensively, often using available data.

Limitation:

- Inability to link exposure with disease. Data on exposure and outcome are not linked at the individual level; association found with aggregate data may not apply to individuals (this is referred as ecological fallacy)
- Lack of ability to control for effects of potential confounding factors. There may be other things that are the true cause
- It may mask a non-linear relationship between exposure and disease

Case Report and Case Series

- Describes the experience of a single or a group of patients with similar diagnosis or health problem derived from either the practice of one or more health care professionals or a defined health care setting such as hospital, health center or specialized clinic
- The following information/data are important inputs in making better use of the case series study:
  - ✓ Defining the disease or health problem clearly

Case Report and Case Series cont'd…

- ✓Recording the date when the disease/death occurred (Time)
- ✓Recording where the person lived, worked,… (Place, information relevant to the study with regard to place)
- ✓Recording personal characteristics of the person such as age and sex (Person)
- ✓Explore the opportunities for collecting additional data from records or the person directly.
- ✓Estimating the size and the characteristics of the population at risk

Case Report and Case Series cont'd…

Strength:

- ✓ useful for studying signs and symptoms and creating case definitions for epidemiological studies
- ✓ case-series that include cases at various stages of an illness from mild cases to dead supplemented by investigation of the past medical history of these cases and observing them to death (doing autopsy as appropriate) can help build up a picture of the natural history of a disease
- ✓ very useful in providing critical information, for hypothesis generation, for sound analytical studies

Case Report and Case Series cont'd…

Limitations:

- ✓Report is based on single or few patients, which could happen just by coincidence.
- ✓Lack of an appropriate comparison group
- ✓Rates can not be calculated since the population corresponding to the source of cases can not be defined well
- ✓Detailed and complete risk factor information is difficult to obtain for all cases from records
- ✓Studies are prone to atomistic fallacy (the opposite of ecological fallacy); the forces that cause or prevent disease at an individual level are different from those that work at societal level

Cross Sectional Studies (Survey)

- investigate disease and risk factor (exposure) patterns in a representative sample of a population in a narrowly defined time period
- It can be useful to identify associations, generate and test hypothesis and, by repeating at different time periods, measure change and hence evaluate interventions
- Findings from cross-sectional study of a sample population can be generalized cautiously if basic characteristics of the populations are similar
- Cross sectional studies provide the most reliable estimate of burden of disease in a population by estimating prevalence

Cross Sectional Studies cont'd…

Strength:

- ✓ Easy to conduct
- ✓ Not time consuming
- ✓ Can be used to compare population with different characteristics as in comparative cross sectional studies

Limitation:

- ✓ "chicken or egg" dilemma - difficult to know which occurred first, the determinant/exposure or the outcome
- ✓ Survivor bias- people who died of the disease are missed in cross-sectional study. One way of correcting this problem is to supplement population studies with clinical studies

ANALYTIC STUDIES

- Focuses on the determinants of a disease by testing the hypothesis formulated from descriptive studies, with the ultimate goal of judging whether a particular exposure causes or prevents disease.
- Analytic studies are broadly classified into two - observational and interventional studies.
- Both types use "control group", the use of control group is the main distinguishing feature of analytic studies

Observational studies

- Information are obtained by observation of events.
- No intervention is done, no deliberate interference with natural course of disease.
- Cohort and case-control are in this category

Observational studies cont'd…

Cohort

- Subjects are selected by exposure, or determinants of interest, and followed to see if they develop the disease or outcome of interest
- E.g. Follow 100 children who received BCG vaccination and another 100 who didn't get BCG vaccination and see how many of them get tuberculosis

Case Control

- Subjects are selected with respect to presence or absence of disease, or outcome of interest, and then inquiries are made about past exposure to the factor(s) of interest
- E.g. Take people with and without TB, ask them if they ever had BCG vaccination

Interventional / Experimental

- The researcher does something about the disease or exposure and observes the changes
- Investigator has control over who gets exposure and who don't. The key is that the investigator assign into either group, whether it is done randomly or not
- Always prospective
- E.g. Assign children randomly to get chloroquine or not, and see how many develop symptomatic malaria.

Measurement in Epidemiology

Measures of Disease Occurrence

The number of cases in a given community can give more epidemiologic sense if they are related to the size of the population

- **Ratio**: the value of *x* and *y* may be completely independent, or *x* may be included in *y*.
  - Example: Male: Female (male to female ratio)
- **Proportion**: is a ratio (expressed as a percent) in which *x* is included in *y*.
  - Example: Male/Both sexes (proportion of male in a community)

Measures of Disease Occurrence cont'd…

- **Rate**: measures the occurrence of an event in a population over time. The time component is important in the definition. Rates are often proportions
  - Example: $\frac{\text{Measles cases in under five in 1995}}{\text{Under five children in 1995}}$

- ✓When we call a measure a **ratio** we usually mean a non-proportional ratio.
- ✓When we call a measure a **proportion**, we usually mean a proportional ratio that doesn't measure an event over time.
- ✓When we use the term **rate**, we frequently refer to a proportional ratio that does measure an event in a population over time

Common Measures of Disease Frequency

Risk (cumulative incidence):

- Likelihood that an individual will contract a disease.
- The proportion of unaffected individuals who, on average, will contract the disease of interest over a specified period of time

$$\text{Risk} = \frac{\text{New cases occurring during a given time period}}{\text{Population at risk during the same period}}$$

Common Measures of Disease Frequency cont'd…

Prevalence:

- The amount of disease that is present already in a population
- Indicates the number of existing cases in a population

$$\text{Prevalence} = \frac{\text{all new and pre-existing cases during a given time period}}{\text{population during the same time period}}$$

Common Measures of Disease Frequency cont'd…

Incidence:

- Measures the rapidity with which newly diagnosed patients develop over time
- Most common way of measuring and comparing the frequency of disease in populations
- The period of time for the rate must be specified

$$\text{Incidence Rate} = \frac{\text{Number of new cases during observation period}}{\text{Person-time observed}}$$

Common Measures of Disease Frequency cont'd…

Case Fatality

- Case fatality is a measure of disease severity and is defined as the proportion of cases with a specified disease or condition who die within a specified time.
- It is usually expressed as a percentage.

Case fatality (%) =

$$\frac{\text{Number of deaths from diagnosed cases in a given period}}{\text{Number of diagnosed cases of the disease in the same period}} \times 100$$

Using available information to measure health and disease

Mortality

- Epidemiologists often investigate the health status of a population by starting with information that is routinely collected.
- In many high-income countries the fact and cause of death are recorded on a standard death certificate, which also carries information on age, sex, and place of residence.
- The International Statistical Classification of Diseases and Related Health Problems (ICD) provides guidelines on classifying deaths

Using available information cont'd…

Verbal autopsy

- A verbal autopsy is an indirect method of ascertaining biomedical causes of death from information on symptoms, signs and circumstances preceding death, obtained from the deceased person's family.
- In many middle- and low-income countries, verbal autopsy is the only method used to obtain estimates of the distribution of the causes of death
- Verbal autopsies are used mainly in the context of demographic surveillance and sample registration systems.

Death rates

- The death rate (or crude mortality rate) for all deaths or a specific cause of death is calculated as follows:

Crude mortality rate =

$$\frac{\text{Number of deaths during a specified period}}{\text{Number of persons at risk of dying during the same period}} \times 1000$$

- The main disadvantage of the crude mortality rate is that it does not take into account the fact that the chance of dying varies according to age, sex, race, socioeconomic class and other factors

Death rates cont'd….

Age-specific death rates

- Death rates can be expressed for specific groups in a population which are defined by age, race, sex, occupation or geographical location, or for specific causes of death.

Age- and sex-specific death rate =

Total number of deaths occurring in a specific age and sex group of the population in a defined area during a specified period X 1000

Estimated total population of the same age and sex group of the population in the same area during the same period

Infant mortality

- The infant mortality rate is commonly used as an indicator of the level of health in a community
- It measures the rate of death in children during the first year of life, the denominator being the number of live births in the same year

Infant mortality rate =

$$\frac{\text{Number of deaths in a year of children less than 1 year of age}}{\text{Number of live births in the same year}} \times 1000$$

Infant mortality cont’d …

- The use of infant mortality rates as a measure of overall health status for a given population is based on the assumption that it is particularly sensitive to socioeconomic changes and to health care interventions
- Infant mortality has declined in all regions of the world, but wide differences persist between and within countries

Child mortality rate

- The child mortality rate (under-5 mortality rate) is based on deaths of children aged 1–4 years, and is frequently used as a basic health indicator.
- Injuries, malnutrition and infectious diseases are common causes of death in this age group. The under-5 mortality rate describes the probability (expressed per 1000 live births) of a child dying before reaching 5 years of age

Child mortality rate cont'd …

$$<5MR = \frac{\text{Number of deaths of children <5 yrs in a yr}}{\text{Total number of children < 5 years of age}} \times 1000$$

OR

$$< 5MR = \frac{\text{Number of deaths of children <5yrs in a yr}}{\text{Total live births in the same year}} \times 1000$$

Maternal mortality rate

- The maternal mortality rate refers to the risk of mothers dying from causes associated with delivering babies, complications of pregnancy or childbirth.
- This important statistic is often neglected because it is difficult to calculate accurately

Maternal mortal it y rate=

$$\frac{\text{Number of maternal deaths in a given geographic area in a given year}}{\text{Number of live births that occurred among the population of the given geographic area during the same year}} \times 100{,}000$$

Morbidity

- Death rates are particularly useful for investigating diseases with a high case-fatality
- many diseases have low case-fatality, for example, most mental illnesses, musculoskeletal diseases, rheumatoid arthritis, chickenpox and mumps. In this situation, data on morbidity (illness) are more useful than mortality rates
- Morbidity data are often helpful in clarifying the reasons for particular trends in mortality
- Changes in death rates could be due to changes in morbidity rates or in case-fatality

Morbidity cont’d ….

sources of morbidity data include:

- ✓hospital admissions and discharges
- ✓outpatient and primary health care consultations
- ✓specialist services (such as injury treatment)
- ✓registers of disease events (such as cancer and congenital malformations)

Disability

- Epidemiologists are concerned not only with the occurrence of disease, but also with the consequences of disease: impairments, disabilities and handicaps.
- These have been defined by the WHO International Classification of Functioning, Disability and Health (ICF)
- ICF is a useful tool for understanding and measuring health outcomes
- It can be used in clinical settings, health services or surveys, at the individual or population level

Disability cont’d …

key parameters of ICF are:

- **impairment**: any loss or abnormality of psychological, physiological or anatomical structure or function
- **disability**: any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being
- **handicap**: a disadvantage for a given individual, resulting from an impairment or a disability, that limits or prevents the fulfilment of a role that is normal (depending on age, sex, and social and cultural factors) for that individual

Measures of Association

Rate Ratio

- Measures of association between risk factors and disease are often calculated from data presented in a two by two table

Exposure	Disease	
	YES (+)	NO (-)
YES (+)	A	B
NO (-)	C	D

Measures of Association cont'd …

- The **relative risk** or **risk ratio** compares the risk of some health-related event (often disease or death) in two groups, typically in persons exposed to the disease to those not exposed:

$$\frac{A}{A+B} \div \frac{C}{C+D}$$

Measures of Association cont'd …

- **Odds ratio**, or **cross-product ratio**, is another measure of association which quantifies the relationship between an exposure and health outcome from a comparative study

$$OddsRatio = \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{ad}{bc}$$

Measures of Association cont'd …

- Attributable risk (AR) or Risk difference (RD) indicate how much of the risk is due to (or attributable to) the exposure. Quantify the excess risk in the exposed that can be attributable to the exposure by removing the risk of disease that could have occurred anyway due to other causes.

  AR = Risk in exposed - Risk in non-exposed

Measures of Association cont'd …

- **Relative risk (RR):** estimates the magnitude of the association between exposure and disease and indicates the likelihood of developing the disease in the exposed group relative to those who are not exposed.

$$RR = \frac{\text{Risk in exposed}}{\text{Risk in unexposed}}$$

Measures of Association cont'd …

Odds of exposure: is a simple ratio, not a proportion. Indicates odds of exposure relative to the disease status.

Odds of exposure in diseased = a/c or a:c

Odds of exposure in not-diseased = b/d or b:d

Odds of disease: is a simple ratio, not a proportion. Indicates odds of diseased relative to the exposure status.

Odds of disease in exposed = a/b or a:b

Odds of disease in unexposed = c/d or c:d

Measures of Association cont'd …

Odds Ratio (OR): is the chance of being exposed (or diseased) as opposed to not being exposed (or diseased). It is possible to calculate either exposure or disease odds ratio, which are exactly the same.

The epidemiological thinking behind odds ratio is that if a disease is casually associated with an exposure, then the odds of exposure in the diseased group will be higher than the corresponding odds in the non-diseased group. It is also called cross product ratio.

$$\text{Exposure OR} = a/c \div b/d = a/c \times d/b = ad/cb$$

$$\text{Disease OR} = a/b \div c/d = a/b \times d/c = ad/bc$$

Measures of Association cont'd …

Attributable Risk Percent (AR%) among exposed: estimate the proportion of disease among the exposed that is attributable to the exposure, or the proportion of the disease that could be prevented by eliminating the exposure.

$$AR\% = \frac{\text{Risk in the exposed - Risk in unexposed}}{\text{Risk in exposed}}$$

$$= \frac{RR - 1}{RR} \times 100 \quad \text{or} \quad OR\text{-}1/OR \times 100$$

Measures of Association cont'd …

Population Attributable Risk (PAR) is the risk in total population minus risk in the non-exposed. Estimate the excess rate of disease in the total study population that is attributable to the exposure.

PAR = Risk in population - Risk in unexposed

Population Attributable Risk Percent (PAR%) Estimate the proportion of disease in the study population that is attributable to the exposure and thus could be eliminated if the exposure were eliminated.

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

Possible Outcomes in studying the relationship between disease and exposure

- No association between exposure and disease

  Attributable risk = 0

  Relative risk/odds ratio = 1

- Positive association between the exposure and the disease (i.e., more exposure, more disease)

  Attributable risk > 0

  Relative risk/odds ratio > 1

- Negative association between the exposure and the disease (i.e., more exposure, less disease)

  Attributable risk < 0 (negative)

  Relative risk/odds ratio < 1 (a fraction)

Interpretation of Measures of Association

- **Rate Ratio**: measures the strength of association between an exposure and disease and provide information that can be used to judge whether a valid observed association is likely to be causal

- **Attributable Risk**: measures the public health impact of an exposure, assuming that the association is one of cause and effect More explanation about interpretation.docx