Case Study 1

Definition and Use of Common Epidemiological Concepts and Terms

By Dr. Sam Tornquist through GOPA Worldwide Consultants
Specialist consultant, surveillance and information
Knowledge Support Technical Assistance (KSTA 6535)
Funded by the Asian Development Bank (ADB)

The aim of this document is to familiarize the reader with common epidemiological terms, which are used in communications from WHO, national and regional public health authorities and in research reports.

Familiarity with these terms and concepts enables the reader to follow news and to participate in discussions, and to use data needed to plan countermeasures. Building countermeasures on epidemic/pandemic data is vital to prevent and contain health threats such as COVID-19.

This document is accompanying elective Case Study reading materials which illustrates how these terms are used. Reading case study materials aims for the objective that participants shall familiarize with these key concepts and terms to be well prepared to use information provided in epidemiological reports.

Attack rate

- The number of cases of disease in a specific population, divided by the total population at risk for a defined period of time, usually expressed as %.

Attributable risk percentage

- A statistical measure that estimates the number of cases of a disease, attributable to the exposure of interest.
- *For example, what is the attributable risk for COVID-19, due to exposure at work, or exposure during commuting to/from work, as compared to exposure in the family home.*

Bias

- An error in a study design caused by the tendency of researchers to expect certain conclusions on the basis of own personal beliefs, that results in incorrect conclusions, for example regarding the association between potential risk factors and disease occurrence.

Case fatality rate

- Death from a certain disease in proportion to all cases.

Case reports

- Cases defined, such as by the WHO Standard case definition with or without diagnostic confirmation, which in the case of COVID should be done by PCR as best and most sensitive test, giving the least risk for false positive and false negative results

Case-series

- A compilation of case-reports

Cause-specific death rate

- Number of deaths from a specific cause, expressed as a number per 100,000 population

Cross-sectional survey

- A descriptive epidemiological study design that uses a representative sample of the population to collect information on current health status, personal characteristics and potential risk-factors at one point in time.

False-negative test result

- A test result that is negative although the individual actually has the disease or interest.

False positive test result

- A diagnostic test result which falsely show a positive test result, when the individual actually does not have the disease of interest.

Incidence rate

- The rate of NEW cases of a condition or disease in a population in a specific time period. This provides an estimate of the risk for the disease in the population

Measures of association

- Statistical analysis methods used to investigate the relationship between two or more variables or events

Morbidity rate

- A disease rate, specifically prevalence and incidence rates in a population over a specified time period

Mortality rate

- The number of deaths from all causes divided by the total population at a particular time and place.

Reproductive number or Reproductive rate R

- The Reproductive number for an infection, describes the estimated average number of susceptible individuals who will be infected by each infected individual.
- There is a potential ultimate reproductive rate which occurs when all individuals in society are susceptible and when there are no barriers against disease transmission, which is called R0.
- Once interventions are launched, and once vaccination programs are started, the effective reproduction number will be reduced, for example due to social distancing, use of face-masks etc
- In case the reproductive rate is high, then this will also affect how many people in society need to be immune, to stop spread of the infection. This threshold is called herd immunity threshold and is related to the reproductive rate.

Sensitivity

- The probability that an individual who has the disease of interest will receive a positive screening or diagnostic test result

Specificity

- The probability that an individual who does not have the disease of interest will have a negative screening test result

Surveillance

- The systematic collection, analysis and evaluation-analysis of all aspects of disease occurrence and spread, resulting in information that is useful for planning and launching interventions to mitigate and control the disease.

Incidence

- Number of new cases occurring in a defined population during a defines period of time
- The **incidence** represents the **RATE of occurrence**, e.g. occurrence of **NEW cases** arising in a given period of time in a specified population
- Incidence = Number of new cases during a defined time period divided by population at risk
- Incidence can be presented as raw number, or per population in tranches of 1000, 10,000 or 100,000 population

Prevalence

- The **prevalence** of disease is the number of **existing cases = old + new cases**, in a defined population at a given point in time.
- Since this will also change over time, one may repeat the measurement, and present (old + new cases) at time 1, time 2 and time 3, etc. With this approach, presenting series of prevalence measurements, one talk about **point-prevalence**.
- Note the difference with incidence, which only includes NEW cases.
- Prevalence is the total number of existing disease cases (old + new) in a defined population at a particular point in time, or in a specified time-period
- Prevalence = Total number of cases at a given point in time per estimated population at time. This can be presented as a raw number, or expressed per 1000, per 10,000 or 100,000 population.

Association and Causation

- Association is defined as the **concurrence** of two variables more often than would be expected by chance. Association does not necessarily mean a causal relationship.

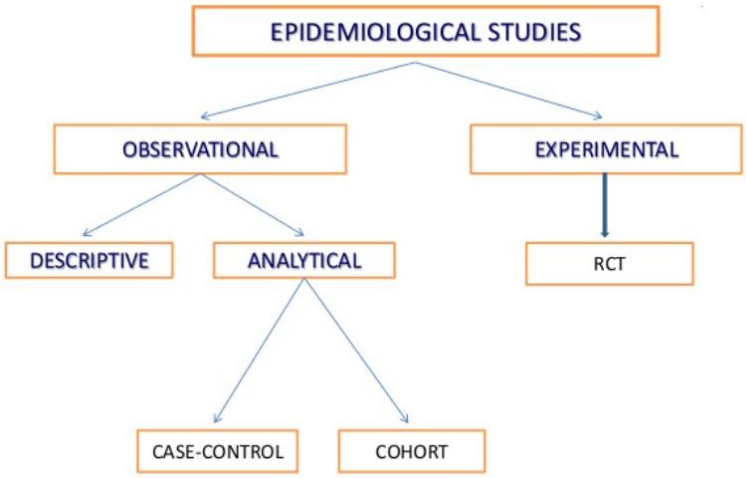
- Correlation describes the strength of association between two variables and is expressed with a correlation coefficient, ranging from -1 to +1
    - When the correlation coefficient measurement is +1, it describes a perfect linear positive relationship
    - When the correlation coefficient measurement is -1, it describes a perfect linear negative relationship
    - **Causation implies association <u>and</u> correlation.**
- NB that correlation and association **do not necessarily** imply causation !

Spurious association

- Example of a spurious relationship can be seen by examining a city's ice cream sales. The sales might be highest when the rate of drownings in city swimming pools is highest. To allege that ice cream sales cause drowning, or vice versa, would be to imply a spurious relationship between the two:
    - **Indirect association:** e.g. endemic goitre and altitude
    - **Direct and causal association:** e.g. (1) *one-to-one causal association*; Streptococcus infection and tonsillitis. *E.g. (2) multifactoral causation*; For cardiovascular disease – there are multiple risk factors

Descriptive Epidemiology

- Is used to characterize the distribution of disease within a population. It describes the person, place, and time characteristics of disease occurrence.
- Descriptive studied
    - Is often the first step in an epidemiological investigation
    - Is limited to description of the occurrence of a disease in a population
    - Provides foundation for formulation of hypotheses which need further research to be evaluated



Analytical Epidemiology

- Is used to test hypotheses to determine whether statistical associations exist between suspected causal factors and disease occurrence. It also is used to test the effectiveness and safety of therapeutic and medical interventions.
- The tests of analytical epidemiology are carried out through four major types of research study designs:
    - Cross-sectional studies,
    - Case-control studies
    - Cohort studies, and
    - Controlled clinical trials.

Experimental studies

- Interventional or experimental studies involves attempts to change variables influencing subjects under study.
-
-

- This could for example mean elimination of some dietary factor believed to cause an allergy, or testing a new treatment such as immunization of a selected groups of patients or populations to prevent spread or reduce mortality of COVID-19.
- The effect of an intervention is measured by comparing the outcome in the experimental group with the same measurement in a control group.
- Types of experimental studies
    - Randomized controlled trials, which preferentially also should be blinded and cross-over in design;
    - Field trials and community trials.

Cross-sectional studies

- Are used to explore associations of disease with variables of interest. For example, a cross-sectional study designed to investigate whether residential exposure to the radioactive gas radon increases the risk of lung cancer may examine the level of radon gas in the homes of lung cancer patients.
- Cross-sectional studies have the advantage of being inexpensive and simple to conduct. Their main disadvantage is that they establish associations at most, not causality.

Case-control studies

- Start with people with a particular disease (cases) and a suitable control group without the disease. Then one compare the two groups for their exposure to the factor that is suspected of having caused the disease.
- Case-control studies are most useful for ascertaining the cause of rare events, such as rare cancers. Case-control studies have the advantages of being quick to conduct and inexpensive, and they require only a small number of cases and controls.
- Their main disadvantage is that they rely on recall, which may be biased, or on records to determine exposure status.

Cohort studies

- Are observational studies in which a defined group of people (the cohort) is followed over time and outcomes are compared for individuals who were exposed or not exposed to a factor at different levels. Cohorts can be assembled in the present and followed into the future or identified from past records (a historical cohort study).
- The main advantage of cohort studies is that they identify the timing and directionality of events. Their main disadvantages are that they require large sample sizes and long follow-up times. They also are not suitable for investigating rare diseases.

Community trials

- In this form of experiment, the treatment groups are communities rather than individuals
- This is particularly appropriate for diseases that are influenced by social conditions, and for which prevention efforts target group behaviour.
- Examples
    - Fortification of food and iron deficiency anemia
    - Social distancing, lockdown and spread of communicable diseases such as COVID-19

Controlled clinical trials

- Are studies that test various interventions, including medicines or other medical interventions to assess their effectiveness and safety. A controlled clinical compares the outcome of a test intervention given to an experimental group with a control group that does not receive the same drug or intervention.

- To minimize bias, individuals involved in clinical trials may be randomly assigned to the experimental and control groups.

Controlled field-trials

- Are in principle constructed as clinical trials, but conducted in community, for example to analyze effect of measures for controlling an epidemic problem, such as by social distancing, face-masks or other. There are particular difficulties with these designs for obvious reasons, but this is currently also done during the COVID-19 pandemic to assess efficacy of various containment initiatives.

Randomized Controlled Trials (RCT)

- RCT is a planned experiment designed to analyse the effficacy and safety of an intervention in humans, by comparing the effect of intervention in a study group compared to a control group.
- A critical aspects of a RCT design, is to measure a relevant endpoint (outcome) to assess efficacy, and to determine risk for adverse effects.
- Key to the relevance of the trial results, is how the design select participants in the intervention and control groups., and if the endpoint is relevant. Many trials fail on that, using ”surrogate endpoints” and biased selection of participants.
- The design of the RCT must be based on relevant criteria, representative of how the tested treatment will be used in daily clinical practice. Matching of control group and intervention group is also very important, as well as blinding and preferal cross-over
- **Basic steps in RCTs**
    - Design a protocol, with defined measurements for effect and protocol to detect adverse and side effects
    - Select reference control and experimental intervention group populations
    - Allocate individuals by randomization to intervention and control groups
    - Deliver manipulation or interventios to be tested
    - Follow-up measurement of relevant effect ”endpoint” (NB that many clinical trials use irrelevant endpoints, such as ”time to normalized blood pressure” when the needed effect parameter is reduced rate of stroke or survival as example.
    - Under-way continuously assess effect and adverse side effects, to allow breaking trial if risks are too high or if effect is so pronounced that all need to be given the treatment
    - Assessment of final outcome.

Reference population (Target population)

- This is the population among which the results of the study is applicable
- A reference population may be defined as humans of a certain age-group, gender, socioeconomic group or group with a specific pre-existing condition or risk-factor

Experimental population (Study population)

- This group is drawn from the above defined Reference population
    - The Experimental population must be representative of the Reference population
    - The selected Experimental population must be qualified for the study
    - The selected population must one by one, individually give INFORMED CONSENT