

Practical Epidemiology and Biostatistics in Research

Epidemiologic glossary

Pages with reference to book, From 245 To 247

Asma Fozia Qureshi, Melvyn A. Lobo, S.N. Bazmi Inam (Department of Community Health Sciences, Aga Khan University, Karachi.)

The glossary presents terms (in alphabetical order) used in scientific literature and Commonly encountered in articles, books and journals.

Association (Syn: relationship, correlation)

Statistical dependence between two or more events or variables. The association is described as positive when the occurrence of higher values of a variable is associated with the occurrence. of higher values of another variable. In a negative association the occurrence of higher values of one variable is associated with lower values of the other variable.

Attack Rate (Syn: cumulative incidence)

Used in studies of infectious diseases. It is calculated by dividing the total number of persons that got sick by the total number that were exposed to the risk (e.g., number of people who developed diarrhoea and vomiting after eating ice cream at a party for a 300 people).

Attributable Risk (AR)

Measures the excess risk of disease due to exposure to a risk factor. It is the absolute difference in incidence rates of disease in the exposed and unexposed populations. If there is no difference in the two incidence rates the AR will be zero. If there is causal association between the risk factor and the disease the AR will be more than zero, and its value will indicate the anticipated quantitative impact of a preventive intervention. This measure of risk can be calculated from cohort studies.

BIAS

Deviation of results from the truth. Any trend in the collection, analysis or interpretation of data that can lead to conclusions that are systematically different from the truth. "Selection Bias" results when the selection criteria for subjects differ between compared groups in relation to the exposure or outcome of interest, leading to non-comparability. "Recall Bias" may arise because individuals exposed to a particular hazard or disease remember their experiences differently or report subsequent events with a different degree of completeness or accuracy than those not exposed.

Case Control Study (Syn: retrospective study)

A study in which cases (persons with the disease) are compared with appropriate controls (persons without the disease) on the basis of exposure to aetiologic factors in the past histories. This is done to determine which of their past exposures/experiences is relevant to the disease. The study starts after the onset of disease and looks back to possible causal factors.

Case: Fatality Rate

The number of cases dying from a disease divided by the total cases during a specified interval of time.

Cause-specific Rates

The number of deaths or illness due to a specific disease divided by the total population.

Cohort Study (Syn: prospective study, follow-up study)

A study in which individuals are selected on the basis of being exposed or non-exposed to a potential risk factor. These individuals are followed up for a period of time to compare the development or incidence of disease in the two groups.

Confidence Interval (Syn: confidence limits)

The range likely to include the true value of the parameter of concern at a stated level of probability, such that the value of the parameter is expected to be included in a given percentage of repeated sampling estimation. It provides a useful device for arguing that the sample results are representative of

the population from which the sample was selected. For example, 95% confidence limits signify a 5% chance of error in representativeness of the sample result.

Confounding

A situation in which a measure of the effect of an exposure on outcome is distorted because of the association of exposure with other factors (confounders) that can influence the outcome.

Cross-sectional Study (Syn: prevalence study)

A study in which the study population is selected independent of their disease and exposure status which are then determined at the time of the study. It provides a 'snap-shot' of the population under study at a point in time. As temporal sequence between cause and effect generally cannot be determined, conclusions regarding causality cannot be drawn. Disease prevalence rather than incidence is recorded from this type of study.

Endemic

The usual prevalence or constant presence of a disease or infectious agent in a given geographic area or population group.

Epidemic

The occurrence in a community or region of a group of cases of an illness (or an outbreak) clearly in excess of normal expectancy. An epidemic occurring over a very wide area and usually affecting a large proportion of the population within a stated period of time or age. Also a non-technical term encompassing a variety of measures of the probability of an (generally un-favourable) outcome.

Incidence Rate

The number of net cases of a disease that develop in a population at risk during a specified period of time; this is also called "cumulative incidence". "Incidence density" is the number of new cases that develop in a population divided by the sum of each individual's time at risk; i.e., not all persons at risk are followed for equal periods of time, therefore, the time that each person is observed is added to give the total "person-time" of observation which is often used as the denominator.

Incubation Period

The interval between exposure to infection and the appearance of the first symptom.

Infective Period

The period of time that an individual is capable of transmitting disease with or without contact.

Mortality Rate

The Incidence rate of death

Odds Ratio (OR)

The measure of association generally used in a case-control study. It is a ratio of the odds of exposure among the cases to the odds of exposure among the controls. Generally the odds ratio is an appropriate estimate of the relative risk.

Population Attributable Rate

A measure of the amount of disease associated with an exposure within a population. It is the difference between the incidence of disease in the entire population and that among the non-exposed.

Prevalence Rate

The number of existing cases divided by the number of persons (both ill and well) in a population at a specified point in time.

P Value

A measure of the consistency of the data with the null hypothesis, i.e., it means was the probability that the difference between two sets of values are due to chance. Conventionally a p value of less than 0.05 is considered significant, indicating that the difference seen has a less than 5% probability of having occurred by chance.

Rate

The number of events (e.g. death or disease frequency) in a specified period divided by the population at risk of the events in the same time period.

Relative Risk

Measures the risk of disease occurrence. It is calculated by dividing the incidence of disease in the exposed by that in the unexposed. This measure of association is calculated from prospective or cohort studies.

Risk

A probability that an individual will become ill within a stated period of time or age. Also a non-technical term encompassing a variety of measures of the probability of an (generally unfavourable) outcome.

Screening

Screening is a method for presumptive identification of unrecognized disease or defect by use of tests that can be applied rapidly and economically. The purpose is to distinguish among apparently normal people those who probably have a disease from those who probably do not. Screening tests are not diagnostic; additional tests and examination are needed to establish definitive diagnosis.

Variable:

Any class of measurement (e.g., sex, age, haemoglobin, etc) on which individual observations are made is called a variable. Discrete variables are those for which the possible values are quite distinct and separated such as grades of students performance in a class. Continuous variables are those which can assume an uninterrupted or continuous range of values such as height, weight, blood pressure, etc. Studies often look at the relationship between two variables, and how one variable (X) changes with another (Y). The variable under study i.e. X is the "outcome/dependant" variable whereas Y is the "predictor/independent" variable as it predicts the values that variable X will assume for any given value of Y.

Standard Mortality Ratio (SMR)

"Standardization" is a procedure of adjustment used on crude rates which are influenced by differences in population structure (such as age or sex). The SMR is the ratio of the number of observed deaths in a population divided by the expected number of deaths in that population, commonly multiplied by 100. The denominator (i.e., the expected number of deaths) is calculated by applying age-specific rates of a standard population to the population of interest. Adjustment thus allows comparisons of rates of two populations.

References

1. Hennekens, C.H., Buring, J.E. Epidemiology in medicine. S.L. Myrent (Ed). 1st edition, Boston, Little Brown and Co., 1987.
2. Last, 3M. (Ed). A dictionary of epidemiology, (2nd ed), Oxford, Oxford University Press, 1988
3. Norman, G.R., Streiner, D.L. PDQ statistics, Asian Edition, Singapore, Maulygraphic Publishers Pte. Ltd., 1989.