

Glossary of Epidemiologic Terms

The definitions given are valid for the terms as they are used on this web site, but different definitions may be used in other contexts. *A Dictionary of Epidemiology*, edited by J.M. Last for the International Epidemiological Association and published by Oxford University Press, was helpful in providing a number of the definitions.

A

Active immunity. See [Immunity, active](#).

Age-adjusted mortality rate. See [Mortality rate, age-adjusted](#).

Agent. A factor that is essential for a disease, chronic conditions, or injury to occur. Examples of agents include microorganisms, chemical substances, forms of radiation, and, in the case of injury, physical force. Agents can cause a health problem by either being introduced, being present in excess, or being present at deficient levels.

Age-specific mortality rate. See [Mortality rate, age-specific](#).

Alternative hypothesis. See [Hypothesis, alternative](#).

Analytic epidemiology. See [Epidemiology, analytic](#).

Analytic study. See [Study, analytic](#).

Antibody. Any of a variety of proteins in the blood that are generated to produce immunity against microorganisms or their toxins.

Applied epidemiology. See [Epidemiology, applied](#).

Arithmetic mean. See [Mean, arithmetic](#).

Association. The statistical relationship between two or more events, characteristics, or other [variables](#).

Attack rate. A form of [incidence](#) that measures [frequency](#) of disease, chronic conditions, or injury in a particular population for a limited time, such as during an [outbreak](#). In calculating attack rates, the numerator is the number of new cases of a health problem during an outbreak, and the denominator is the population at the beginning of the period.

Attack rate, secondary. A measure of the [frequency](#) of new cases of a disease, chronic condition, or injury among the contacts of known [case-patients](#).

Attributable proportion. See [Proportion, attributable](#).

B

Bar chart. A visual display in which each category of a variable is represented by a bar. Bar charts are used to show variations in size among categories.

Bias. A systematic deviation from the truth; any trend in the collection, analysis, interpretation, publication, or review of data that can lead to conclusions that are systematically different from the truth.

Biologic transmission. See Transmission, biologic.

Birth cohort. See Cohort, birth.

Box plot. A visual display that summarizes data using a "box and whiskers" format to show the minimum and maximum values (ends of the whiskers), interquartile range (length of the box), and median (line through the box).

C

Carrier. A person or animal who harbors the infectious agent for a disease and can transmit it to others, but does not show signs of the disease. A carrier may be asymptomatic (never show signs of the disease) or may show signs of the disease only during the incubation period, convalescence, or postconvalescence. The period of being a carrier may be short (a transient carrier) or long (a chronic carrier).

Case. An instance of a particular disease, chronic condition, or type of injury. A variety of criteria may be used to identify cases (See Case definition), and the epidemiologic definition of a case is not necessarily the same as the ordinary clinical definition. (See also Case-patient)

Case-control study. See Study, case-control.

Case-patient. A person in a case-control study who has the disease or health condition under investigation. (See also Case)

Case definition. A set of standard criteria for determining whether a person has a particular disease or health condition. A case definition specifies clinical criteria and details of time, place, and person.

Case-fatality rate. The proportion of people with a particular condition (case-patients) who die from that condition. In calculating case-fatality rates, the numerator is the number of people who die from the condition, and the denominator is the total number of people with the condition.

Cause of disease. A factor (characteristic, behavior, event, etc.) that directly influences the occurrence of a disease. Reducing such a factor in a population should reduce occurrence of the disease.

Cause-specific mortality rate. See Mortality rate, cause-specific.

Census. The enumeration of an entire population, usually including details on residence, age, sex, occupation, ethnic group, marital status, birth history, and relationship to head of household.

Chain of infection. A process that begins when an agent leaves its source through a portal of exit, is conveyed by some mode of transmission, and then enters through an appropriate portal of entry to infect a susceptible host.

Class. A grouping of observations of values of a variable. Classes are created for convenience in analyzing frequency. (See also Class boundaries and Class interval)

Class boundaries. The values determining the upper and lower limits of a class.

Class interval. The span of values of a continuous variable that lies between the class boundaries.

Clinical criteria. The symptoms and features of a disease that would be detected by physician analysis.

Clinical disease. A disease that has been manifested by its symptoms and features.

Clinical trial. See Trial, clinical.

Cluster. An aggregation of cases of a disease or other health condition that are closely grouped in time and place. The number of cases may or may not exceed the number expected, and frequently the expected number is not known. Cases of cancer and birth defects are often investigated as clusters.

Cohort. A well-defined group of people who have had a common experience or exposure and are then followed up, as in a cohort study or prospective study, to determine the incidence of new diseases or health conditions.

Cohort, birth. A group of people born during a particular period or year.

Cohort study. See Study, cohort.

Common source outbreak. See Outbreak, common source.

Community trial. See Trial, community.

Confidence interval. A range of values for a variable (e.g., a rate).

Confidence level. The proportion of similarly constructed confidence intervals that include the parameter of interest.

Confidence limits. The end points (i.e., the minimum and maximum values) of a confidence interval.

Contact. Exposure to a source of an infection; a person so exposed.

Contagious. Capable of being transmitted from one person to another by contact or close proximity.

Contingency table. A table of cross-tabulated data that allows for calculating associations. The 2-by-2 table, with cases tabulated by exposure and outcome, is the contingency table most commonly used in epidemiology.

Continuous variable. See Variable, continuous.

Control. The group of people without the health problem under study in a case-control study; a person in that group. For controls, investigators choose people who are as similar as possible to the cases, but without the health problem under study. In a case-control study, the control group is compared with the case group to determine associations between exposures and outcomes and to test hypotheses. (See also Study, case-control)

Crude mortality rate. (See Mortality rate, crude)

Cumulative frequency. In a frequency distribution, the number or proportion of cases with a particular value or less.

Cumulative frequency curve. A plot of the cumulative frequency rather than the actual frequency for each class interval of a variable. This type of graph is useful for identifying medians and quartiles and other percentiles.

D

Death-to-case ratio. The number of deaths attributed to a particular disease, chronic condition, or type of injury during a specified period divided by the number of new cases of that disease or injury identified during the same period.

Demographic information. The personal characteristics of age, sex, race, residence, and occupation. Demographic information is used in descriptive epidemiology to define the population at risk.

Denominator. The lower portion of a fraction. Epidemiologists use fractions to calculate rates, or ratios. The denominator is usually the population at risk, although it may also be a measure, such as person-time, that quantifies the population's exposure.

Dependent variable. See Variable, dependent.

Descriptive epidemiology. See Epidemiology, descriptive.

Determinant. Any factor that brings about change in a health condition or in other defined characteristics.

Direct transmission. See Transmission, direct.

Discrete variable (or data). See Variable (or data), discrete.

Distribution. The complete summary of the frequency and pattern of the values or categories of a measurement. In epidemiology, distribution is the frequency and pattern of health-related characteristics and events in a population.

Dot plot. A visual display of the specific data points of a variable.

Droplet nuclei. The residue of dried droplets that is easily inhaled and exhaled and may remain suspended in air for long periods and be blown over great distances.

Droplet spread. The direct transmission of an infectious agent by means of the aerosols produced in sneezing, coughing, or talking.

E, F, G

Endemic health condition. A disease, chronic condition, or type of injury that is constantly present in a given geographic area or population group; may also refer to the usual prevalence of a disease or condition.

Environmental factor. An extrinsic factor, such as geology, climate, insects, sanitation, or health services, that affects an agent and the opportunity for exposure.

Epidemic. (*Syn: outbreak*) The occurrence of more cases of a particular type of disease, chronic condition, or injury than expected in a given area, or among a specific group of people, over a particular period of time. (More at **Outbreak**)

Epidemic curve. A histogram that shows the course of an outbreak or epidemic by plotting the number of cases of a disease, chronic condition, or injury according to time of onset.

Epidemic period. The time span of an epidemic.

Epidemiologic triad. The traditional model of infectious disease causation, which has three components: an external agent, a susceptible host, and an environment that brings the host and agent together so that disease occurs.

Epidemiology. The study of the distribution and determinants of health conditions or events in populations, and the application of this study to control health problems.

Epidemiology, analytic. The aspect of epidemiology concerned with why and how a health problem occurs. Analytic epidemiology uses comparison groups to provide baseline data so that associations between exposures outcomes can be quantified and hypotheses about the cause of the problem can be tested. Examples include cohort studies and case-control studies.

Epidemiology, applied (or field). The application or practice of epidemiology to control and prevent health problems.

Epidemiology, descriptive. The aspect of epidemiology concerned with gathering, organizing, and summarizing data on "person" (Who is ill?), "time" (When did they

become ill?), and "place" (Where could they have been exposed to the illness?). This information is then used to conduct analytic epidemiology.

Evaluation. Systematic and objective examination of activities to determine how relevant and effective they are.

Experimental study. See Study, experimental.

Exposed group. A group whose members have had contact with a cause of, or possess a characteristic that is a determinant of, a particular health problem.

Exposure. Coming into contact with a cause of, or possessing a characteristic that is a determinant of, a particular health problem.

Frequency. The amount, or number of occurrences, of a disease, chronic condition, injury, or other attribute or event in a population.

Frequency distribution. A complete summary of the frequencies of the values or categories of a variable. Frequency distribution is often displayed in a two-column table with the individual values or categories in the left column, and the number of observations in each category in the right column.

Frequency polygon. A graph of a frequency distribution in which values of the variable are plotted on the horizontal axis, and the number of observations are plotted on the vertical axis. Data points are plotted at the midpoints of the intervals and are connected with a straight line.

Geometric mean. See Mean, geometric .

Graph. A visual display of quantitative data arranged on a system of coordinates.

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

H

Health. A state of complete physical, mental, and social well-being and not merely the absence of disease or other infirmity.

Health indicator. Any of a variety of measures (e.g., mortality rate) that indicate the state of health of people in a defined population.

Health information system. A combination of health statistics from various sources. Data from these systems is used to learn about health status, health care, provision and use of services, and the impact of services and programs on health.

Herd immunity. See Immunity, herd.

High-risk group. A group of people whose risk for a particular disease, health condition, or type of injury is higher than that of the rest of their community or population.

Histogram. A visual representation of the frequency distribution of a continuous variable. The class intervals of the variable are grouped on a linear scale on the horizontal axis, and the class frequencies are on the vertical axis. Rectangles are drawn so that their bases equal the class intervals, and their heights correspond to the class frequencies.

Host. A person or other living organism that is susceptible to an infectious agent under natural conditions.

Host factor. An intrinsic factor (e.g., age, race, sex, behaviors) that influences an individual's exposure, susceptibility, or response to an agent.

Hyperendemic health problem. A disease, chronic condition, or type of injury that is constantly present at a high incidence and/or prevalence. Examples include diabetes among the Pima Indians in Arizona and Hepatitis B in China and elsewhere in Asia.

Hypothesis. A supposition, arrived at from observation or reflection, that leads to refutable predictions; any conjecture cast in a form that will allow it to be tested and refuted.

Hypothesis, alternative. The supposition that an exposure is associated with the health condition under study. The alternative is adopted if the null hypothesis proves implausible.

Hypothesis, null. The supposition that an exposure is not associated with the health condition under study. The null hypothesis is the basis for most parametric tests for statistical significance.

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

I, J, K, L

Immunity, active. Resistance developed in response to an antigen (infecting agent or vaccine) and usually characterized by the presence of antibody produced by the host.

Immunity, herd. The resistance of a group to an infectious agent. This group resistance exists because a high proportion of people in the group are immune to the agent. Herd immunity is based on the number of people who are susceptible and the probability that they will come into contact with an infected person. By vaccinating large numbers of people in a population to protect them from smallpox, health officials used herd immunity to control and eradicate the disease.

Immunity, passive. Immunity conferred by an antibody produced in another host. This type of immunity can be acquired naturally by an infant from its mother or artificially by administration of an antibody-containing preparation (antiserum or immune globulin).

Incidence. A rate that measures the frequency with which a health problem, such as a new injury or case of illness, occurs in a population. In calculating incidence, the numerator is the number of new cases occurring in the population during a given period of time, and the denominator is the total population at risk during that time.

Incubation period. The period following exposure, when pathologic changes are not apparent, and ending with the onset of symptoms of an infectious disease.

Independent variable. See Variable, independent.

Indirect Transmission. See Transmission, indirect.

Individual data. Data that have not been put into a frequency distribution.

Infant mortality rate. See Mortality rate, infant.

Infectivity. The proportion of people who are exposed to an agent and become infected.

Interquartile range. The central portion of a distribution, calculated as the difference between the third quartile and the first quartile. This range includes the middle one-half of the observations in the set, leaving one-quarter of the observations on each side.

Latency period. The period following exposure, when pathologic changes are not apparent, and ending with the onset of symptoms of a chronic disease.

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

M

Mean, arithmetic. The measure of central location commonly called the average. The arithmetic mean is calculated by adding all the values in a group of measurements and dividing by the number of values in the group.

Mean, geometric. The mean, or average, of a set of data measured on a logarithmic scale.

Measure of association. A quantified relationship between exposure and a particular health problem. Commonly used measures of association include relative risk, rate ratio, and odds ratio.

Measure of central location. A central value that best represents a distribution of data. Common measures of central location are the mean, median, and mode. Also called the measure of central tendency.

Measure of dispersion. A measure of the spread of a distribution out from its central value. Measures of dispersion used often in epidemiology are the interquartile range, variance, and the standard deviation.

Measurement scale. The complete range of possible values for a measurement. An example is the set of possible answers to a question in a survey.

Mechanical transmission. See Transmission, mechanical.

Median. The middle value in a set of numbers (or the average of two middle numbers) above and below which lie an equal number of values. (See also Measure of central location)

Medical surveillance. Monitoring individuals who may have been exposed to a health problem, for the purpose of detecting early symptoms.

Midrange. The halfway point, or midpoint, in a set of observations. For most types of data, the midrange is calculated by adding the smallest observation and the largest observation and dividing by two. The midrange is usually calculated as an intermediate step in determining other measures.

Mode. The most frequently occurring value in a set of observations. (See also Measure of central location)

Morbidity. Disease; any departure, subjective or objective, from a state of physiological or psychological health and well-being.

Mortality rate. A measure of the frequency of occurrence of death in a defined population during a specified time interval.

Mortality rate, age-adjusted. A mortality rate that has been statistically modified to account for the effect of different age distributions in different populations in a study.

Mortality rate, age-specific. A mortality rate limited to a particular age group. In calculating age-specific mortality rates, the numerator is the number of deaths in the age group, and the denominator is the number of people in that age group.

Mortality rate, cause-specific. The mortality rate from a specified cause. In calculating cause-specific mortality rates, the numerator is the number of deaths attributed to a specific cause during a specified time interval in a population, and the denominator is the size of the population at the midpoint of the time interval.

Mortality rate, crude. A population's mortality rate from all causes of death.

Mortality rate, infant. The mortality rate for children less than one year of age. In calculating infant mortality rates, the numerator is the number of deaths reported

among this age group during a given time period, and the denominator is the number of live births reported during the same period. The infant mortality rate is usually expressed per 1,000 live births. See also **Mortality rate, neonatal** and **Mortality rate, postneonatal**.

Mortality rate, neonatal. The mortality rate for children from birth up to, but not including, 28 days of age. In calculating neonatal mortality rates, the numerator is the number of deaths in this age group during a given time period, and the denominator is the number of live births reported during the period. The neonatal mortality rate is usually expressed per 1,000 live births.

Mortality rate, postneonatal. The mortality rate for children from 28 days up to, but not including, 1 year of age. In calculating postneonatal mortality rates, the numerator is the number of deaths among this age group during a given time period, and the denominator is the number of live births during the same period. The postneonatal mortality rate is usually expressed per 1,000 live births.

Mortality rate, race-specific. A mortality rate limited to a specified racial group. Both numerator and denominator are limited to that group.

Mortality rate, sex-specific. A mortality rate among either males or females.

N

Natural history of disease. The course of a disease from the time it begins until it is resolved.

Necessary cause. A factor that must be present for a disease or other health problem to occur.

Neonatal mortality rate. See **Mortality rate, neonatal**.

Nominal scale. A type of measurement scale. Nominal scales consist of qualitative categories whose values have no inherent order or rank. Information regarding the race, religion, and country of birth of people in a population is a good example.

Normal curve. A bell-shaped curve, which results when a normal distribution is graphed.

Normal distribution. The symmetrical clustering of values around a central location. A normal distribution 1) is a continuous, symmetrical distribution with both tails extending to infinity; 2) has an identical arithmetic mean, mode, and median; and 3) has a shape that is completely determined by the mean and standard deviation.

Notifiable disease. A disease that, by law, must be reported to public health authorities upon diagnosis.

Null hypothesis. See **Hypothesis, null**.

Numerator. The upper portion of a fraction.

O

Observational study. An epidemiologic study in which there is no intervention and nature is allowed to take its course. Changes or differences in one characteristic are studied in relation to changes or differences in others.

Odds ratio. A measure of association used in comparative studies to quantify the relationship between an exposure and a health outcome; also known as the cross-product ratio.

Ordinal scale. A type of measurement scale. Ordinal scales consist of qualitative categories whose values have a distinct order. The categories are qualitative in that there is no natural distance to be measured between their possible values. An example is social class (I, II, III).

Outbreak. (*Syn: epidemic*) Because the public sometimes perceives "outbreak" as less sensational than "epidemic," it is sometimes the preferred word. Sometimes the two words are sometimes differentiated, with "outbreak" referring to a localized health problem, and "epidemic," to one that takes in a more general area. (More at **Epidemic**)

Outbreak, common source. An outbreak in which people are exposed to a common harmful influence, such as an infectious agent or toxin. The exposure period may be brief, or people may be exposed over a period of days, weeks, or longer, with the exposure being either intermittent or continuous.

Outbreak, point source. A common source outbreak in which the exposure period is relatively brief so that all cases occur within one incubation period.

Outbreak, propagated. An outbreak that does not have a common source, but instead spreads from person to person.

Outcome(s). Any or all of the possible results that may stem from exposure to a causal factor or from preventive or therapeutic interventions; all identified changes in health status that result from the handling of a health problem.

P, Q, R

Pandemic. An epidemic occurring over a very wide area (several countries or continents) and usually affecting a large proportion of the population.

Passive immunity. See **Immunity, passive**.

Pathogenicity. The proportion of people who are infected by an agent and then develop clinical disease.

Percentile. A set of cut points used to divide a distribution or a set of ranked data into 100 parts of equal area with each interval between the points containing 1/100 of the observations. For example, the 5th percentile is a cut point with 5% of the observations below it and the remaining 95% above it.

Period prevalence. See Prevalence, period.

Person-time. The total of the units of time, whether weeks, months, or years, that people were exposed to a condition or were actively involved in a study. One person-year can represent a single person who was exposed for one year or an accumulation, such as two people who were each exposed for half a year.

Person-time rate. A measure of the incidence rate of an event (e.g., disease, injury, or death) in a population over an observed period. The person-time rate directly incorporates time into the denominator.

Pie chart. A circular chart depicting observed data and divided into "slices" that are proportional to the frequency of the categories of the variable assigned to them.

Point prevalence. See Prevalence, point.

Point source outbreak. See Outbreak, point source.

Population. The total number of inhabitants of a given area or country. In sampling, the population may refer to the units from which the sample is drawn, not necessarily the total population of people. A population can also be a particular group at risk, such as everyone who is engaged in a certain occupation.

Portal of entry. A pathway into the **Portal of entry**. A pathway into the host that gives an agent access to tissue that will allow it to multiply or act.

Portal of exit. A pathway by which an agent can leave its source.

Postneonatal mortality rate. See Mortality rate, postneonatal.

Predictive value positive. A measure of the predictive value of a reported case or epidemic; the proportion of cases reported by a surveillance system or classified by a case definition that are true cases.

Prevalence. The number or proportion of cases or events or conditions in a given population.

Prevalence, period. The amount of a particular disease, chronic condition, or type of injury present in a population over a period of time. (See also Prevalence, point)

Prevalence, point. The amount of a particular disease, chronic condition, or type of injury present in a population at a single point in time. (See also Prevalence, period)

Prevalence rate. The proportion of people in a population who have a particular disease, chronic condition, injury, or attribute at a specified point in time or over a specified period of time.

Propagated outbreak. See Outbreak, propagated.

Proportion. A ratio in which the numerator is included in the denominator; the ratio of a part to the whole, expressed as a "decimal fraction" (e.g., 0.2), a fraction (1/5), or a percentage (20%).

Proportion, attributable. A measure of the impact of a causative factor on the public health; the proportion of a disease, chronic condition, or injury that can be attributed to exposure to a particular factor.

Proportionate mortality. The **Proportionate mortality**. The proportion of deaths in a population attributable to a particular cause over a period of time. Each cause of death is expressed as a percentage of all deaths, and the sum the proportionate mortality for all causes must equal 100%. These proportions are not mortality rates because the denominator is all deaths instead of the population in which the deaths occurred.

Public health surveillance. The systematic, ongoing collection, analysis, interpretation, and dissemination of health data. The purpose of public health surveillance is to gain knowledge of the patterns of disease, injury, and other health problems in a community so that we can work toward controlling and preventing them.

Race-specific mortality rate. See Mortality rate, race-specific.

Random sample. See Sample, random.

Range. In statistics, the difference between the largest and smallest values in a distribution; in common use, the span of values from smallest to largest.

Rate. An expression of the relative frequency with which an event occurs in a defined population.

Rate ratio. A comparison of two groups in terms of incidence rates, person-time rates, or mortality rates.

Ratio. The relative size of two quantities. A ratio is expressed by dividing one quantity by the other.

Relative risk. A comparison of the risk of a health problem in two groups.

Representative sample. See Sample, representative.

Reservoir. The habitat in which an infectious agent normally lives, grows, and multiplies. Humans, animals, and the environment can serve as reservoirs.

Risk. The probability that an individual will be affected by, or die from, an illness or injury within a stated time or age span.

Risk factor. An aspect of personal behavior or lifestyle, an environmental exposure, or a hereditary characteristic that is associated with an increase in the occurrence of a particular disease, chronic condition, or injury.

Risk ratio. A comparison of the risk of a particular health problem in two groups.

S

Sample. A selected subset of a population. A sample may be random or nonrandom and representative or nonrepresentative.

Sample, random. A sample of individuals chosen in such a way that each one has the same (and known) probability of being selected.

Sample, representative. A sample whose characteristics correspond to those of the original or reference population.

Scatter diagram (or Scattergram). A graphic display of the relationship between two variables. A dot is plotted on the graph for each set of paired values for two continuous variables, with one variable plotted on the horizontal axis, and the other plotted on the vertical axis.

Seasonality. Change in physiological status or in the occurrence of a disease, chronic condition, or type of injury that conforms to a regular seasonal pattern.

Secondary attack rate. See Attack rate, secondary.

Secular trend. See Trend, secular.

Sensitivity. The ability of a system to detect epidemics and other changes in the occurrence of health problems; the proportion of people with a health problem who are correctly identified by a screening test or case definition. (See also **Specificity**)

Sentinel event. A case of one of the notifiable diseases that has been included in a sentinel surveillance system.

Sentinel surveillance. A surveillance system using a prearranged sample of sources (e.g., physicians, hospitals, clinics) who have agreed to report all cases of one or more notifiable diseases.

Sex-specific mortality rate. See Mortality rate, sex-specific.

Skewed. A distribution that is asymmetrical.

Specificity. The proportion of people without a particular disease, chronic condition, or type of injury who are correctly identified by a screening test or case definition. (See also, **Sensitivity**)

Sporadic illness. An illness that occurs infrequently and irregularly.

Spot map. A visual display of the geographic pattern of a health problem. On a map of the area, a marker is placed to indicate where each affected person lives, works, or may have been exposed. Spot maps can reveal clusters or patterns that provide clues to the identity and origins of the problem.

Standard deviation. A statistical summary of how dispersed the values of a variable are around its mean. Standard deviation is equal to the positive square root of the variance.

Standard error (of the mean). The standard deviation of a theoretical distribution of sample means of a variable around the true population mean of that variable. Standard error is computed as the standard deviation of the variable divided by the square root of the sample size.

Statistical inference. Generalizations developed from sample data, usually with calculated degrees of uncertainty.

Statistical significance. The measure of how likely it is that a set of study results could have occurred by chance alone. Statistical significance is based on an estimate of the probability of the observed or a greater degree of association between independent and dependent variables occurring under the null hypothesis. The level of statistical significance is usually expressed by the P value.

Study, analytic. A study in which groups are compared to identify and quantify associations, test hypotheses, and identify causes. Two common types are cohort studies and case-control studies.

Study, case-control. An analytic study that compares a group of people with a certain disease, chronic condition, or type of injury (case-patients) with a group of people without the health problem (controls) to detect differences in characteristics such as exposure to an agent.

Study, cohort. (*Syn: follow-up, longitudinal, and prospective study*) An observational analytic study in which enrollment is based on status of exposure to a certain factor or membership in a certain group. Populations are followed and disease, death, or other health-related outcomes are determined and compared.

Study, experimental. A study in which investigators identify the type of exposure that each individual (clinical trial) or community (community trial) has had and then follows the individuals' or communities' health status to determine the effects of the exposure.

Study, prospective. See Study, cohort.

Sufficient cause. A causal factor or collection of factors whose presence is always followed by the occurrence of a particular health problem.

Surveillance. See **Public health surveillance**.

Survival curve. A curve that starts at 100% of the study population and shows the percentage of the population still surviving at successive times for as long as information is available. A survival curve may also be used to depict freedom from a health problem, complication, or some other endpoint.

T

Table shell. A table that is complete except for the data; the outline for a planned analysis.

Transmission (of infection). Any mode or mechanism by which an infectious agent is spread to a susceptible host. (See also **Transmission, direct** and **Transmission, indirect**)

Transmission, biologic. Indirect transmission by a vector in which the infectious agent undergoes part of its life cycle inside the vector before it is transmitted to the host. (See also **Transmission, mechanical**)

Transmission, direct. Immediate transfer of an agent from a reservoir to a host by direct contact or droplet spread.

Transmission, indirect. Transfer of an agent from a reservoir to a host either by being suspended in air particles (airborne), carried by an inanimate intermediary (vehicleborne), or carried by an animate intermediary (vectorborne).

Transmission, mechanical. Indirect transmission by a vector in which the infectious agent does not undergo physiologic changes inside the vector. (See also **Transmission, biologic**)

Trend. Movement or change in frequency over time, usually upwards or downwards. (See **Trend, secular**)

Trend, secular. Changes over a long period of time, generally years or decades. (See **Trend**)

Trial, clinical. An experimental study using data from individuals. Investigators identify the type of exposure that each individual has had and then follow the individuals' health status to determine the effects of the exposure.

Trial, randomized clinical. A clinical trial in which individuals are randomly assigned to exposure or treatment groups.

Trial, community. An experimental study that uses data from communities. Investigators identify the type of exposure that each community has had and then follow the communities' health status to determine the effects of the exposure.

Trial, randomized community. A community trial in which communities are randomly assigned to exposure or intervention groups.

U, V, W, X, Y, Z

Validity. The degree of accuracy of a measurement. For survey instruments, validity refers to what the questions actually measure in practice, as compared with what they are intended to measure.

Variable. Any characteristic or attribute that can be measured and can have different values.

Variable, continuous. A variable that has the potential for having an infinite number of values along a continuum. Common examples are height and weight. (See also **Discrete variable**)

Variable, dependent. In a statistical analysis, a variable whose values are a function of other variable.

Variable (or data), discrete . A variable that is limited to a finite number of values; data for such a variable. One example would be the number of cigarettes smoked per day by people in a study of smoking and lung cancer. (See also **Variable, continuous**)

Variable, independent. An exposure, risk factor, or other characteristic being observed or measured that is hypothesized to influence an event or manifestation (the dependent variable).

Variance. A measure of the dispersion shown by a set of observations, defined by the sum of the squares of deviations from the mean , divided by the number of degrees of freedom in the set of observations. (See **Standard deviation**)

Vector. In epidemiology, an animate intermediary in the **Vector**. In epidemiology, an animate intermediary in the indirect transmission of an agent that carries the agent from a reservoir to a susceptible host.

Vehicle. In epidemiology, an inanimate intermediary in the indirect transmission of an agent that carries the agent from a reservoir to a susceptible host.

Virulence. The measure of severity of a disease, expressed as the proportion of people with the disease who become extremely ill or die.

Vital statistics. Systematically tabulated data on registered information about births, marriages, divorces, and deaths.

Years of potential life lost (YPLL). A measure of the impact of premature death on a population. YPLL is calculated using a predetermined expected life span (usually 65 years) and is the sum of the differences between this age and the age at death for everyone who died early.

Zoonoses. An infectious disease that is transmissible from animals to humans.