Appendix A
Glossary

The definitions given are valid as they are used in this publication but different definitions may be used in other contexts. A Dictionary of Epidemiology, Second Edition, edited by J.M. Last for the International Epidemiological Association and published by Oxford University Press, 1988, was helpful in providing a number of the definitions.

A

AGE-ADJUSTED MORTALITY RATE. A mortality rate statistically modified to eliminate the effect of different age distributions in the different populations.

AGENT. A factor, such as a microorganism, chemical substance, or form of radiation, whose presence, excessive presence, or (in deficiency diseases) relative absence is essential for the occurrence of a disease.

AGE-SPECIFIC MORTALITY RATE. A mortality rate limited to a particular age group. The numerator is the number of deaths in that age group; the denominator is the number of persons in that age group in the population.

ANALYTIC EPIDEMIOLOGY. The aspect of epidemiology concerned with the search for health-related causes and effects. Uses comparison groups, which provide baseline data, to quantify the association between exposures and outcomes, and test hypotheses about causal relationships.

ANALYTIC STUDY. A comparative study intended to identify and quantify associations, test hypotheses, and identify causes. Two common types are cohort study and case-control study.

APPLIED EPIDEMIOLOGY. The application or practice of epidemiology to address public health issues.

ASSOCIATION. Statistical relationship between two or more events, characteristics, or other variables.

ATTACK RATE. A variant of an incident rate, applied to a narrowly defined population observed for a limited period of time, such as during an epidemic.

ATTRIBUTABLE PROPORTION. A measure of the public health impact of a causative factor; proportion of a disease in a group that is exposed to a particular factor which can be attributed to their exposure to that factor.

B

BAR CHART. A visual display of the size of the different categories of a variable. Each category or value of the variable is represented by a bar.

BIAS. Deviation of results or inferences from the truth, or processes leading to such systematic deviation. 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. The indirect vector-borne transmission of an infectious agent in which the agent undergoes biologic changes within the vector before being transmitted to a new host.

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 without apparent disease who harbors a specific infectious agent and is capable of transmitting the agent to others. The carrier state may occur in an individual with an infection that is inapparent throughout its course (known as asymptomatic carrier), or during the incubation period, convalescence, and postconvalescence of an individual with a clinically recognizable disease. The carrier state may be of short or long duration (transient carrier or chronic carrier).

CASE. In epidemiology, a countable instance in the population or study group of a particular disease, health disorder, or condition under investigation. Sometimes, an individual with the particular disease.

CASE-CONTROL STUDY. A type of observational analytic study. Enrollment into the study is based on presence (``case'') or absence (``control'') of disease. Characteristics such as previous exposure are then compared between cases and controls.

CASE DEFINITION. A set of standard criteria for deciding whether a person has a particular disease or health-related condition, by specifying clinical criteria and limitations on time, place, and person.

CASE-FATALITY RATE. The proportion of persons with a particular condition (cases) who die from that condition. The denominator is the number of incident cases; the numerator is the number of cause-specific deaths among those cases.

CAUSE OF DISEASE. A factor (characteristic, behavior, event, etc.) that directly influences the occurrence of disease. A reduction of the factor in the population should lead to a reduction in the occurrence of disease.

CAUSE-SPECIFIC MORTALITY RATE. The mortality rate from a specified cause for a population. The numerator is the number of deaths attributed to a specific cause during a specified time interval; the denominator is the size of the population at the midpoint of the time interval.

CENSUS. The enumeration of an entire population, usually with details being recorded 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 reservoir or host through a portal of exit, and is conveyed by some mode of transmission, then enters through an appropriate portal of entry to infect a susceptible host.

CLASS INTERVAL. A span of values of a continuous variable which are grouped into a single category for a frequency distribution of that variable.

CLUSTER. An aggregation of cases of a disease or other health-related condition, particularly cancer and birth defects, which are closely grouped in time and place. The number of cases may or may not exceed the expected number; frequently the expected number is not known.

COHORT. A well-defined group of people who have had a common experience or exposure, who are then followed up for the incidence of new diseases or events, as in a cohort or prospective study. A group of people born during a particular period or year is called a birth cohort.

COHORT STUDY. A type of observational analytic study. Enrollment into the study is based on exposure characteristics or membership in a group. Disease, death, or other health-related outcomes are then ascertained and compared.

COMMON SOURCE OUTBREAK. An outbreak that results from a group of persons being exposed to a common noxious influence, such as an infectious agent or toxin. If the group is exposed over a relatively brief period of time, so that all cases occur within one incubation period, then the common source outbreak is further classified as a point source outbreak. In some common source outbreaks, persons may be exposed over a period of days, weeks, or longer, with the exposure being either intermittent or continuous.

CONFIDENCE INTERVAL. A range of values for a variable of interest, e.g., a rate, constructed so that this range has a specified probability of including the true value of the variable. The specified probability is called the confidence level, and the end points of the confidence interval are called the confidence limits.

CONFIDENCE LIMIT. The minimum or maximum value of a confidence interval.

CONTACT. Exposure to a source of an infection, or a person so exposed.

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

CONTINGENCY TABLE. A two-variable table with cross-tabulated data.

CONTROL. In a case-control study, comparison group of persons without disease.

CRUDE MORTALITY RATE. The mortality rate from all causes of death for a population.

CUMULATIVE FREQUENCY. In a frequency distribution, the number or proportion of cases or events with a particular value or in a particular class interval, plus the total number or proportion of cases or events with smaller values of the variable.

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, quartiles, and other percentiles.

D

DEATH-TO-CASE RATIO. The number of deaths attributed to a particular disease during a specified time period divided by the number of new cases of that disease identified during the same time period.

DEMOGRAPHIC INFORMATION. The ``person'' characteristics--age, sex, race, and occupation--of descriptive epidemiology used to characterize the populations at risk.

DENOMINATOR. The lower portion of a fraction used to calculate a rate or ratio. In a rate, the denominator is usually the population (or population experience, as in person-years, etc.) at risk.

DEPENDENT VARIABLE. In a statistical analysis, the outcome variable(s) or the variable(s) whose values are a function of other variable(s) (called independent variable(s) in the relationship under study).

DESCRIPTIVE EPIDEMIOLOGY. The aspect of epidemiology concerned with organizing and summarizing health-related data according to time, place, and person.

DETERMINANT. Any factor, whether event, characteristic, or other definable entity, that brings about change in a health condition, or in other defined characteristics.

DIRECT TRANSMISSION. The immediate transfer of an agent from a reservoir to a susceptible host by direct contact or droplet spread.

DISTRIBUTION. In epidemiology, the frequency and pattern of health-related characteristics and events in a population. In statistics, the observed or theoretical frequency of values of a variable.

DOT PLOT. A visual display of the actual data points of a noncontinuous variable.

DROPLET NUCLEI. The residue of dried droplets that may remain suspended in the air for long periods, may be blown over great distances, and are easily inhaled into the lungs and exhaled.

DROPLET SPREAD. The direct transmission of an infectious agent from a reservoir to a susceptible host by spray with relatively large, short-ranged aerosols produced by sneezing, coughing, or talking.

E

ENDEMIC DISEASE. The constant presence of a disease or infectious agent within a given geographic area or population group; may also refer to the usual prevalence of a given disease within such area or group.

ENVIRONMENTAL FACTOR. An extrinsic factor (geology, climate, insects, sanitation, health services, etc.) which affects the agent and the opportunity for exposure.

EPIDEMIC. The occurrence of more cases of disease than expected in a given area or among a specific group of people over a particular period of time.

EPIDEMIC CURVE. A histogram that shows the course of a disease outbreak or epidemic by plotting the number of cases by time of onset.

EPIDEMIC PERIOD. A time period when the number of cases of disease reported is greater than expected.

EPIDEMIOLOGIC TRIAD. The traditional model of infectious disease causation. Includes 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-related states or events in specified populations, and the application of this study to the control of health problems.

EVALUATION. A process that attempts to determine as systematically and objectively as possible the relevance, effectiveness, and impact of activities in the light of their objectives.

EXPERIMENTAL STUDY. A study in which the investigator specifies the exposure category for each individual (clinical trial) or community (community trial), then follows the individuals or community to detect the effects of the exposure.

EXPOSED (GROUP). A group whose members have been exposed to a supposed cause of disease or health state of interest, or possess a characteristic that is a determinant of the health outcome of interest.

F

FREQUENCY DISTRIBUTION. A complete summary of the frequencies of the values or categories of a variable; often displayed in a two column table: the left column lists the individual values or categories, the right column indicates the number of observations in each category.

FREQUENCY POLYGON. A graph of a frequency distribution with values of the variable on the x-axis and the number of observations on the y-axis; data points are plotted at the midpoints of the intervals and are connected with a straight line.

G

GRAPH. A way to show quantitative data visually, using a system of coordinates.

H

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

HEALTH INDICATOR. A measure that reflects, or indicates, the state of health of persons in a defined population, e.g., the infant mortality rate.

HEALTH INFORMATION SYSTEM. A combination of health statistics from various sources, used to derive information about health status, health care, provision and use of services, and impact on health.

HIGH-RISK GROUP. A group in the community with an elevated risk of disease.

HISTOGRAM. A graphic representation of the frequency distribution of a continuous variable. Rectangles are drawn in such a way that their bases lie on a linear scale representing different intervals, and their heights are proportional to the frequencies of the values within each of the intervals.

HOST. A person or other living organism that can be infected by an infectious agent under natural conditions.

HOST FACTOR. An intrinsic factor (age, race, sex, behaviors, etc.) which influences an individual's exposure, susceptibility, or response to a causative agent.

HYPERENDEMIC DISEASE. A disease that is constantly present at a high incidence and/or prevalence rate.

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, NULL. The first step in testing for statistical significance in which it is assumed that the exposure is not related to disease.

HYPOTHESIS, ALTERNATIVE. The hypothesis, to be adopted if the null hypothesis proves implausible, in which exposure is associated with disease.

I

IMMUNITY, ACTIVE. Resistance developed in response to stimulus by 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 invasion and spread of an infectious agent, based on the resistance to infection of a high proportion of individual members of the group. The resistance is a product of the number susceptible and the probability that those who are susceptible will come into contact with an infected person.

IMMUNITY, PASSIVE. Immunity conferred by an antibody produced in another host and acquired naturally by an infant from its mother or artificially by administration of an antibody-containing preparation (antiserum or immune globulin).

INCIDENCE RATE. A measure of the frequency with which an event, such as a new case of illness, occurs in a population over a period of time. The denominator is the population at risk; the numerator is the number of new cases occurring during a given time period.

INCUBATION PERIOD. A period of subclinical or inapparent pathologic changes following exposure, ending with the onset of symptoms of infectious disease.

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

INDIRECT TRANSMISSION. The transmission of an agent carried from a reservoir to a susceptible host by suspended air particles or by animate (vector) or inanimate (vehicle) intermediaries.

INDIVIDUAL DATA. Data that have not been put into a frequency distribution or rank ordered.

INFECTIVITY. The proportion of persons exposed to a causative agent who become infected by an infectious disease.

INFERENCE, STATISTICAL. In statistics, the development of generalizations from sample data, usually with calculated degrees of uncertainty.

INTERQUARTILE RANGE. The central portion of a distribution, calculated as the difference between the third quartile and the first quartile; this range includes about one-half of the observations in the set, leaving one-quarter of the observations on each side.

L

LATENCY PERIOD. A period of subclinical or inapparent pathologic changes following exposure, ending with the onset of symptoms of chronic disease.

M

MEAN, ARITHMETIC. The measure of central location commonly called the average. It is calculated by adding together all the individual 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 disease; includes relative risk, rate ratio, odds ratio.

MEASURE OF CENTRAL LOCATION. A central value that best represents a distribution of data. Measures of central location include 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 in epidemiology include the interquartile range, variance, and the standard deviation.

MEDIAN. The measure of central location which divides a set of data into two equal parts.

MEDICAL SURVEILLANCE. The monitoring of potentially exposed individuals to detect early symptoms of disease.

MIDRANGE. The halfway point or midpoint in a set of observations. For most types of data, it is calculated as the sum of the smallest observation and the largest observation, divided by two. For age data, one is added to the numerator. The midrange is usually calculated as an intermediate step in determining other measures.

MODE. A measure of central location, the most frequently occurring value in a set of observations.

MORBIDITY. Any departure, subjective or objective, from a state of physiological or psychological well-being.

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

MORTALITY RATE, INFANT. A ratio expressing the number of deaths among children under one year of age reported during a given time period divided by the number of births reported during the same time period. The infant mortality rate is usually expressed per 1,000 live births.

MORTALITY RATE, NEONATAL. A ratio expressing the number of deaths among children from birth up to but not including 28 days of age divided by the number of live births reported during the same time period. The neonatal mortality rate is usually expressed per 1,000 live births.

MORTALITY RATE, POSTNEONATAL. A ratio expressing the number of deaths among children from 28 days up to but not including 1 year of age during a given time period divided by the number of lives births reported during the same time period. The postneonatal mortality rate is usually expressed per 1,000 live births.

N

NATURAL HISTORY OF DISEASE. The temporal course of disease from onset (inception) to resolution.

NECESSARY CAUSE. A causal factor whose presence is required for the occurrence of the effect (of disease).

NOMINAL SCALE. Classification into unordered qualitative categories; e.g., race, religion, and country of birth as measurements of individual attributes are purely nominal scales, as there is no inherent order to their categories.

NORMAL CURVE. A bell-shaped curve that results when a normal distribution is graphed.

NORMAL DISTRIBUTION. The symmetrical clustering of values around a central location. The properties of a normal distribution include the following: (1) It is a continuous, symmetrical distribution; both tails extend to infinity; (2) the arithmetic mean, mode, and median are identical; and, (3) its shape is completely determined by the mean and standard deviation.

NUMERATOR. The upper portion of a fraction.

O

OBSERVATIONAL STUDY. Epidemiological study in situations where nature is allowed to take its course. Changes or differences in one characteristic are studied in relation to changes or differences in others, without the intervention of the investigator.

ODDS RATIO. A measure of association which quantifies the relationship between an exposure and health outcome from a comparative study; also known as the cross-product ratio.

ORDINAL SCALE. Classification into ordered qualitative categories; e.g., social class (I, II, III, etc.), where the values have a distinct order, but their categories are qualitative in that there is no natural (numerical) distance between their positive values.

OUTBREAK. Synonymous with epidemic. Sometimes the preferred word, as it may escape sensationalism associated with the word epidemic. Alternatively, a localized as opposed to generalized epidemic.

P

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

PATHOGENICITY. The proportion of persons infected, after exposure to a causative agent, who then develop clinical disease.

PERCENTILE. The set of numbers from 0 to 100 that divide a distribution into 100 parts of equal area, or divide a set of ranked data into 100 class intervals with each interval containing 1/100 of the observations. A particular percentile, say the 5th percentile, is a cut point with 5 percent of the observations below it and the remaining 95% of the observations above it.

PERIOD PREVALENCE. The amount a particular disease present in a population over a period of time.

PERSON-TIME RATE. A measure of the incidence rate of an event, e.g., a disease or death, in a population at risk over an observed period to time, that directly incorporates time into the denominator.

PIE CHART. A circular chart in which the size of each ``slice'' is proportional to the frequency of each category of a variable.

POINT PREVALENCE. The amount of a particular disease present in a population at a single point in time.

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.

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 which are true cases.

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

PREVALENCE RATE. The proportion of persons in a population who have a particular disease or attribute at a specified point in time or over a specified period of time.

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

PROPORTION. A type of 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), as a fraction (1/5), or, loosely, as a percentage (20%).

PROPORTIONATE MORTALITY. The proportion of deaths in a specified population over a period of time attributable to different causes. Each cause is expressed as a percentage of all deaths, and the sum of the causes must add to 100%. These proportions are not mortality rates, since the denominator is all deaths, not the population in which the deaths occurred.

PUBLIC HEALTH SURVEILLANCE. The systematic collection, analysis, interpretation, and dissemination of health data on an ongoing basis, to gain knowledge of the pattern of disease occurrence and potential in a community, in order to control and prevent disease in the community.

R

RACE-SPECIFIC MORTALITY RATE. A mortality rate limited to a specified racial group. Both numerator and denominator are limited to the specified group.

RANDOM SAMPLE. A sample derived by selecting individuals such that each individual has the same probability of selection.

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 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 value obtained by dividing one quantity by another.

RELATIVE RISK. A comparison of the risk of some health-related event such as disease or death in two groups.

REPRESENTATIVE SAMPLE. A sample whose characteristics correspond to those of the original population or reference population.

RESERVOIR. The habitat in which an infectious agent normally lives, grows and multiplies; reservoirs include human reservoirs, animals reservoirs, and environmental reservoirs.

RISK. The probability that an event will occur, e.g. that an individual will become ill or die within a stated period of time or age.

RISK FACTOR. An aspect of personal behavior or lifestyle, an environmental exposure, or an inborn or inherited characteristic that is associated with an increased occurrence of disease or other health-related event or condition.

RISK RATIO. A comparison of the risk of some health-related event such as disease or death in two groups.

S

SAMPLE. A selected subset of a population. A sample may be random or non-random and it may be representative or non-representative.

SCATTER DIAGRAM. A graph in which each dot represents paired values for two continuous variables, with the x-axis representing one variable and the y-axis representing the other; used to display the relationship between the two variables; also called a scattergram.

SEASONALITY. Change in physiological status or in disease occurrence that conforms to a regular seasonal pattern.

SECONDARY ATTACK RATE. A measure of the frequency of new cases of a disease among the contacts of known cases.

SECULAR TREND. Changes over a long period of time, generally years or decades.

SENSITIVITY. The ability of a system to detect epidemics and other changes in disease occurrence. The proportion of persons with disease who are correctly identified by a screening test or case definition as having disease.

SENTINEL SURVEILLANCE. A surveillance system in which a pre-arranged sample of reporting sources agrees to report all cases of one or more notifiable conditions.

SEX-SPECIFIC MORTALITY RATE. A mortality rate among either males or females.

SKEWED. A distribution that is asymmetrical.

SPECIFICITY. The proportion of persons without disease who are correctly identified by a screening test or case definition as not having disease.

SPORADIC. A disease that occurs infrequently and irregularly.

SPOT MAP. A map that indicates the location of each case of a rare disease or outbreak by a place that is potentially relevant to the health event being investigated, such as where each case lived or worked.

STANDARD DEVIATION. The most widely used measure of dispersion of a frequency distribution, equal to the positive square root of the variance.

STANDARD ERROR (OF THE MEAN). The standard deviation of a theoretical distribution of sample means about the true population mean.

SUFFICIENT CAUSE. A causal factor or collection of factors whose presence is always followed by the occurrence of the effect (of disease).

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. May be applied not only to survival as such, but also to the persistence of freedom from a disease, or complication or some other endpoint.

T

TABLE. A set of data arranged in rows and columns.

TABLE SHELL. A table that is complete except for the data.

TRANSMISSION OF INFECTION. Any mode or mechanism by which an infectious agent is spread through the environment or to another person.

TREND. A long-term movement or change in frequency, usually upwards or downwards.

U

UNIVERSAL PRECAUTIONS. Recommendations issued by CDC to minimize the risk of transmission of bloodborne pathogens, particularly HIV and HBV, by health care and public safety workers. Barrier precautions are to be used to prevent exposure to blood and certain body fluids of all patients.

V

VALIDITY. The degree to which a measurement actually measures or detects what it is supposed to measure.

VARIABLE. Any characteristic or attribute that can be measured.

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.

VECTOR. An animate intermediary in the indirect transmission of an agent that carries the agent from a reservoir to a susceptible host.

VEHICLE. An inanimate intermediary in the indirect transmission of an agent that carries the agent from a reservoir to a susceptible host.

VIRULENCE. The proportion of persons with clinical disease, who after becoming infected, become severely ill or die.

VITAL STATISTICS. Systematically tabulated information about births, marriages, divorces, and deaths, based on registration of these vital events.

Y

YEARS OF POTENTIAL LIFE LOST. A measure of the impact of premature mortality on a population, calculated as the sum of the differences between some predetermined minimum or desired life span and the age of death for individuals who died earlier than that predetermined age.

Z

ZOONOSES. An infectious disease that is transmissible under normal conditions from animals to humans.