

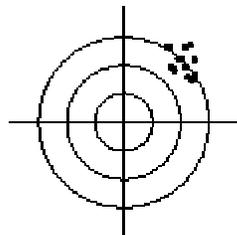
Glossary of Epidemiological Terms

A

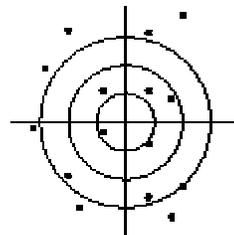
ACCURACY and PRECISION. Accuracy is the extent to which a measurement reflects the true value. It is the tendency of test measurement to center around the true value. "Precise" means sharply defined or measured. Data can be very precise, but inaccurate. It would be precise but inaccurate to say that a meter equals 29.3748 inches. It would actually be more accurate to say that a meter equals a little over one yard, although that may not sound as impressive.

An accurate measurement is one that is very close to the true value of the phenomenon we are observing. A precise measurement is one that has very little scatter: Repeat measurements give more or less the same value. If the measured data have high precision but poor accuracy, one may suspect that a systematic bias has been introduced, e.g., we are using an instrument where the zero position has not been set properly. If we do not know the expected value of a phenomenon but are trying to determine just that, it is obviously better to have accurate observations with poor precision than very precise, but inaccurate values, since the former will give a correct, but imprecise estimate while the latter will give a wrong, but very precise result.

Figure 1.2: Precision is a measure of repeatability while accuracy refers to how close the average value is to the "true" value.



PRECISION



ACCURACY

ACUTE CONDITION. A type of illness or injury that ordinarily lasts less than 3 months, was first noticed less than 3 months before the reference data of the interview, and was serious enough to have had an impact on behavior. (Pregnancy is also considered to be an acute condition despite lasting longer than 3 months).

AGE ADJUSTMENT, DIRECT METHOD. The application of age-specific rates in a population of interest to a standardized age distribution in order to eliminate differences in observed rates that result from age differences in population composition. This adjustment is usually done when comparing two or more populations at one point in time or one population at two or more points in time. *Age-adjusted rates* are calculated by the direct method as follows:

$$\sum_{i=1}^n r_i \times (p_i/P)$$

where r_i = age-specific rates for the population of interest

p_i = standard population in age group i

$P = \sum_{i=1}^n p_i$ for the age groups that comprise the age range of the rate being age adjusted

n = total number of age groups over the age range of the age-adjusted rate

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 RATE. A rate limited to a particular age group. The numerator is the number of occurrences in that age group; the denominator is the number of persons in that age group in the population.

AIRBORNE TRANSMISSION

Airborne transmission refers to situations where droplet nuclei (residue from evaporated droplets) or dust particles containing microorganisms can remain suspended in air for long periods of time. These organisms must be capable of surviving for long periods of time outside the body and must be resistant to drying. Airborne transmission allows organisms to enter the upper and lower respiratory tracts. Fortunately, only a limited number of diseases are capable of airborne transmission. Diseases capable of airborne transmission include: Influenza, Whooping cough, Pneumonia, Tuberculosis, Polio

AMBULATORY CARE. Health care provided to persons without their admission to a health facility. Ambulatory care is medical care, including diagnosis, observation, treatment and rehabilitation, provided on an outpatient basis. Ambulatory care is given to persons who are not confined to a hospital but rather are ambulatory and, literally, are able to ambulate or walk about. A well-baby visit is considered ambulatory care even though the baby is not walking.

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.

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

ATTACK RATE. The percentage of ill persons out of all persons reporting a risk behavior. It is 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.

AVERAGE LENGTH OF STAY. The average length of stay per discharged patient is computed by dividing the total number of hospital days for a specified group by the total number of discharges for that group.

B

BAR CHART. A graphical technique for presenting discrete data organized so that each observation can fall into one and only one category of the variable. Frequencies are listed along one axis and categories of the variable are listed along the other axis. The lengths of the bars represent frequencies of each group of observations.

A bar chart is a graphical comparison of several quantities in which the lengths of the horizontal or vertical bars represent the relative magnitude of the values. 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. Types of bias include: selection bias, detection bias, information (observation) bias, misclassification, and recall bias.

BIOLOGIC TRANSMISSION. A specific feature of a vector-borne infection. Indirect disease transmission in which an infectious organism undergoes some morphologic or physiologic change during its passage through the vector. For example, transmission of the infectious agent to susceptible host by bite of blood-feeding (arthropod) vector as in malaria, or by other inoculation, as in *Schistosoma* infection.

BIRTHWEIGHT. The first weight of the newborn obtained after birth.
Low birth weight is defined as less than 2,500 grams or 5 pounds 8 ounces.
Very low birth weight is defined as less than 1,500 grams or 3 pounds 4 ounces.
Before 1979 low birth weight was defined as 2,500 grams or less and very low birth weight as 1,500 grams or less.

BODY MASS INDEX (BMI). BMI is a measure that adjusts bodyweight for height. It is calculated as weight in kilograms divided by height in meters squared. Overweight for children and adolescents is defined as BMI at or above the sex- and age-specific 95th percentile BMI cut points from the 2000 CDC Growth Charts. *Healthy weight* for adults is defined as a BMI of 18.5 to less than 25; *overweight*, as

greater than or equal to a BMI of 25; and *obesity*, as greater than or equal to a BMI of 30.

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 that 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 unapparent 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. For example, all persons who attended the church bazaar on February 20 who exhibited signs of diarrhea, vomiting, and abdominal cramps within 72 hours after the bazaar.

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.

CASE REPORT STUDY. A type of descriptive study that consists of a careful, detailed profile of an individual patient.

CASE SERIES STUDY. A type of descriptive study that describes characteristics of number of patients with a given disease.

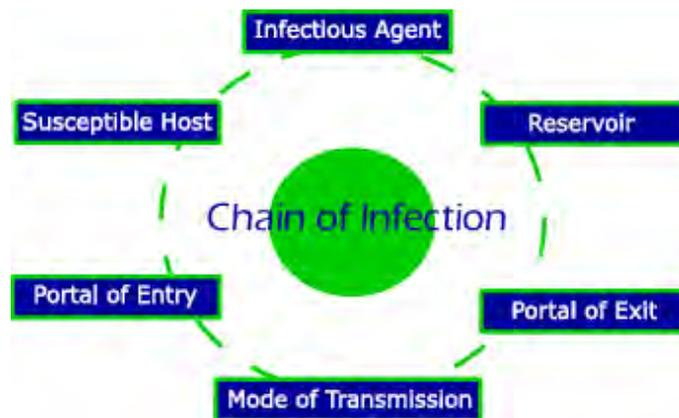
CAUSALITY. The relating of causes to the effects they produce. Some of the criteria for inferring a causal relationship between an implicated food and illness include: strength of association, consistency of the observed association, temporal sequence of events biological plausibility of the observed association, effect of removing the exposure, dose-response relationships, and the exclusion of alternate explanations.

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 infectious agent (e.g., bacteria, virus, parasite, or fungus) leaves its reservoir or host (humans, animals, plants, or the general environment) through a portal of exit (body orifices, bodily fluids, tissue, excreta, secretions), and is conveyed by some mode of transmission (direct, indirect, airborne), then enters through an appropriate portal of entry to infect a susceptible host (compromised individuals. Susceptibility factors include age, pre-existing disease, lifestyle factors, stress, medication, etc.).



CHI SQUARE TEST. A test of statistical significance. A Chi Square (χ^2) test looks at the difference between what we observe in the data and what we would expect if the exposure was not associated with the illness.

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

CLINICAL TRIAL. A clinical trial is an experimental study with patients as subjects. The goal either is to evaluate a potential cure to prevent disease sequelae such as death or disability.

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. Cases are more likely to represent a cluster if they involve (1) a specific disease or condition, (2) a rare type of disease or condition, or (3) a disease or condition in a group not usually affected by that particular disease or condition, such as a cancer in children that is normally seen in adults.

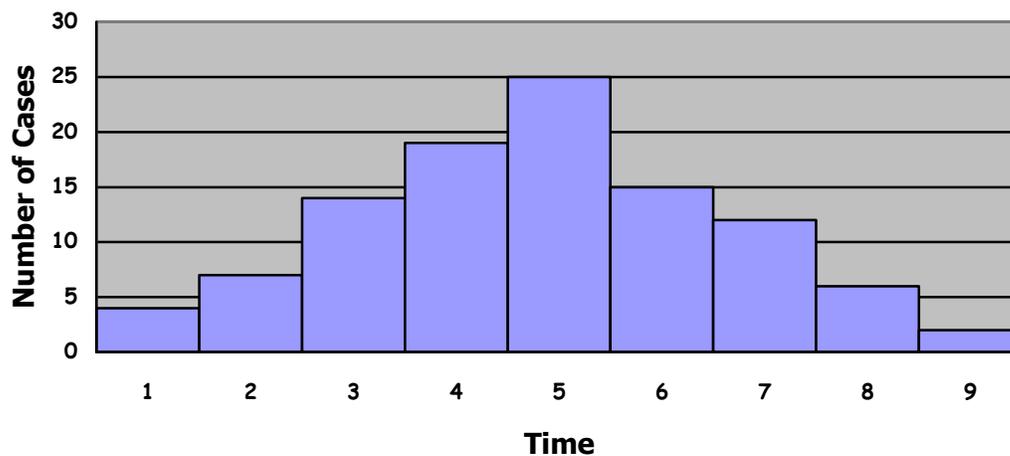
COHORT. A well-defined group of people who have had a common experience or exposure, who are then followed up for the incidence of new diseases or events. For example, a group of people born during a particular period or year is called a birth cohort.

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.

COLONIZED. A carrier state that occurs when a person is not infected with a pathogen, but simply has it on the skin or mucous membrane.

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.

Common Source Epidemic Curve

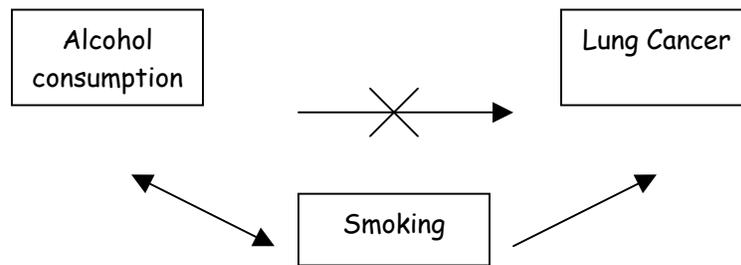


COMMUNICABLE DISEASE. An infectious disease transmitted from an infected person, animal, or reservoir to a susceptible host through an intermediate plant, animal, or the inanimate environment.

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.

CONFOUNDING. This can occur when an association between a risk factor and disease can be explained by a factor associated with both disease and risk factor. For example, an association between smoking, alcohol consumption, and lung cancer is observed.



However, because alcohol consumption does not affect lung cancer rate in non-smokers and there are more smokers among alcohol drinkers, alcohol consumption is associated with smoking rather than lung cancer.

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 used to calculate a potential association. This is also called a 2X2 table:

Exposure (ex. Smoking)	Disease (ex. Lung Cancer)		
	Yes	No	Total
Yes	75	10	85
No	25	90	115
Total	100	100	200

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.

CUMULATIVE INCIDENCE. A measure of the risk or probability of acquiring disease. The number of new cases of diseases in a specified time period divided by the number of individuals at the beginning of the observation period. Technically, cumulative incidence is not a rate, but a proportion.

CROSS-SECTIONAL STUDY. A type of descriptive study in which a set of individuals are studied at a single point in time (or over a defined period of time) for the prevalence of disease. Both risk factors and disease status are ascertained at the same time.

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.

DETECTION BIAS. This type of bias can occur when persons with a risk factor are more likely to have disease detected because of more intense follow-up.

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.

Direct contact transmission requires physical contact between an infected person and a susceptible person, and the physical transfer of microorganisms. Direct contact includes touching an infected individual, kissing, sexual contact, contact with oral secretions, or contact with body lesions. This type of transmission requires close contact with an infected individual, and will usually occur between members of the same household or close friends and family.

Diseases spread exclusively by direct contact are unable to survive for significant periods of time away from a host. Sexually transmitted diseases are almost always spread through direct contact, as they are extremely sensitive to drying.

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. 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.

Some diseases can be transferred by infected droplets contacting surfaces of the eye, nose, or mouth. This is referred to as droplet contact transmission. Droplets containing microorganisms can be generated when an infected person coughs, sneezes, or talks. Droplets can also be generated during certain medical procedures, such as bronchoscopy. Droplets are too large to be airborne for long periods of time, and quickly settle out of air.

Droplet transmission can be reduced with the use of personal protective barriers, such as face masks and goggles. Measles and SARS are examples of diseases capable of droplet contact transmission.

E

ECOLOGICAL (OR CORRELATIONAL) STUDY. A type of descriptive study involving the comparison of disease frequency (incidence or prevalence) between populations that are different with respect to one or more risk factors of interest. The risk factor information is not known for individual subjects, but rather as a total population characteristic.

ECOLOGICAL FALLACY. With an ecological study, the risk factor information is not known for individual subjects, but rather as a total population characteristic. For example, an ecological study may find that Japanese people have a high rate of stomach cancer and also a high rate of rice consumption. However, with an ecological study, it is not known if the same individuals who eat rice are the same as those who have stomach cancer.

EFFECT MODIFICATION. This refers to variation in the magnitude of a measure of exposure effect across levels of another variable.

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.) that 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

FECAL-ORAL TRANSMISSION A type of indirect transmission. Fecal-oral transmission is usually associated with organisms that infect the digestive system. Microorganisms enter the body through ingestion of contaminated food and water. Inside the digestive system (usually within the intestines) these microorganisms multiply and are shed from the body in feces. If proper hygienic and sanitation practices are not in place, the microorganisms in the feces may contaminate the water supply through inadequate sewage treatment and water filtration. Fish and shellfish that swim in contaminated water may be used as food sources. If the infected individual is a waiter, cook, or food handler, then inadequate hand washing may result in food being contaminated with microorganisms.

Fecal-oral transmission can be reduced by:

- Proper storage of food at proper temperatures
- Thorough cooking of food
- Frequent and thorough hand washing, especially after washroom use
- Adequate sewage treatment and water filtration/chlorination systems
- Disinfection of frequent touch surfaces to prevent indirect contact transmission
- Increased public awareness of proper hygiene and food handling

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.) that 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, which 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

IMMUNE. Refers to someone who shows no clinical signs of infection after exposure to a pathogen.

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 unapparent 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).

INDEX CASE. The first case to come to the attention of a disease investigator.

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.

Indirect contact transmission refers to situations where a susceptible person is infected from contact with a contaminated surface. Some organisms (such as Norwalk Virus) are capable of surviving on surfaces for an extended period of time. To reduce transmission by indirect contact, frequent touch surfaces should be properly disinfected.

Frequent touch surfaces (fomites) include:

- Door knobs, door handles, handrails
- Tables, beds, chairs
- Washroom surfaces
- Cups, dishes, cutlery, trays
- Medical instruments
- Computer keyboards, mice, electronic devices with buttons
- Pens, pencils, phones, office supplies
- Children's toys

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.

INFORMATION BIAS. The type of bias can occur when interviewers who are aware of the identity of subjects or factors of interest (e.g., case or control) and collect information unevenly between subjects.

INTERACTION. Two or more component causes acting in the same sufficient cause

INTERNATIONAL CLASSIFICATION OF DISEASE (ICD). The ICD provides the ground rules for coding and classifying cause-of-death data. The ICD is developed collaboratively between the World Health Organization (WHO) and ten international centers, one of which is housed at NCHS. The purpose of the ICD is to promote international comparability in the collection, classification, processing, and presentation of health statistics. Since the beginning of the century, the ICD has been modified about once every 10 years, except for the 20-year interval between ICD-9 and ICD-10.

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 unapparent pathologic changes following exposure, ending with the onset of symptoms of chronic disease.

LIFE EXPECTANCY. Life expectancy is the average number of years of life remaining to a person at a particular age and is based on a given set of age-specific death rates, generally the mortality conditions existing in the period mentioned. Life expectancy may be determined by race, sex, or other characteristics using age-specific death rates for the population with that characteristic.

LIVE BIRTH. A live birth is the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of the pregnancy, which, after such separation, breathes or shows any other evidence of life such as heartbeat, umbilical cord pulsation, or definite movement of voluntary muscles, whether the umbilical cord has been cut or the placenta is attached. Each product of such a birth is considered live born.

LIVE BIRTH ORDER. This item from the birth certificate refers to the total number of live births the mother has had, including the present birth as recorded on the birth certificate. Fetal deaths are excluded.

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. For example, five students have the following test scores: 93, 68, 83, 96, and 85. First, the sum of the scores is found $93+68+83+96+85=425$. Then, the sum of the scores is divided by the number of test scores (in this case, 5). $425/5=85$. Thus, the average test score is 85.

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 that divides a set of data into two equal parts. This number may or not be equivalent to the average or arithmetic mean. The median is found two different ways depending on the number of observations. If the number of observations is odd, then there is a unique central point. Suppose five student's test scores are the following: 93, 68, 83, 96, and 85. First, these numbers are placed in order from lowest to highest: 68, 83, 85, 93, and 96. The middle observation is observation #3, 85. This is the median. For samples with an even number of observations, there is no unique central point, and the two middle points would be averaged to obtain the median. Suppose there are six test scores: 68, 75, 83, 85, 93, 96. The two middle scores are 83 and 85. Once averaged, $(83+85/2=84)$, the median is 84.

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.

MISCLASSIFICATION. This type of bias can occur when there is error in classifying subjects by disease or risk factor that tends to distort associations between disease and risk factors. The effect of misclassification with respect to

exposure or disease is dependent on the individual's disease or exposure status. When the misclassification is random or nondifferential, the proportions of subjects erroneously classified in the study groups are approximately equal. With differential misclassification, the proportions of subjects classified incorrectly differ between groups.

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 live 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. Many diseases have certain well-defined stages that, taken together, are referred to as the "natural history of the disease." These stages are as follows: 1. Stage of pathological onset, 2. Pre-symptomatic stage: from the onset to the first appearance of symptoms and/or signs, and 3. Clinical manifestation of disease, which may progress to fatal termination, be subject to remission or relapses, or regress spontaneously, leading to recovery.

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, e.g., the number of cases of disease.

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 that 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.

OVERSAMPLE. A sampling procedure designed to give a demographic or geographic population a larger proportion of representation in the sample than the population's proportion of representation in the overall population.

P

PANDEMIC. An epidemic occurring over a very wide area (several countries or continents) and usually affecting a large proportion of the population. An example is the 1918 Spanish Influenza pandemic.

PARITY. Parity is defined as the total number of live births ever had by the woman. This number is distinguished from gravidity, which is the total number of times she has been pregnant. Nulliparous women are those who have had no live births, and parous women are those who have given birth to at least one baby.

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.

POVERTY LEVEL. Poverty statistics are based on definitions originally developed by the Social Security Administration. These include a set of money income thresholds that vary by family size and composition. Families or individuals with income below their appropriate thresholds are classified as below the poverty level. These thresholds are updated annually by the U.S. Bureau of the Census to reflect changes in the Consumer Price Index for all urban consumers (CPI-U). For example, the average poverty threshold for a family of four was \$17,029 in 1999 and \$13,359 in 1990.

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.

PREFERRED PROVIDER ORGANIZATION. This is a health plan generally consisting of hospital and physician providers. The PPO provides health care services to plan members usually at discounted rates in return for expedited claims payment. Plan members can use PPO or non-PPO health care providers, however, financial incentives are built into the benefit structure to encourage utilization of PPO providers.

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. Prevalence rates do not differentiate between old and new cases of disease. It is a function of the incidence rate and the mean duration of illness.

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.

PROSPECTIVE STUDY. At the time a study begins, either the exposure or the outcome of interest has not yet 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.

P-VALUE. A measure of how likely the observed results would occur when there is no real association between a food and an illness (i.e. how likely the observed results would occur by chance alone).

Q

QUESTIONNAIRE. A set of questions used to collect information. Typically questions can be grouped into the following categories: identifying, demographic, clinical, risk, and reported information.

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.

RECALL BIAS. This type of bias can occur in retrospective studies when persons with disease tend to report past exposures and events differently from persons without disease.

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

RELIABILITY. Refers to the consistency of a measurement when repeated on the same subjects.

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.

RETROSPECTIVE STUDY. At the time a study begins, both the exposure and the outcome of interest has occurred.

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.

SELECTION BIAS. This type of bias can occur when there are differential selection rates of subjects by disease and/or risk factor characteristics.

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.

SIGNIFICANCE TEST. A test of statistical significance shows how likely one is to get a measure of association as strong as the observed one if there is no difference between the groups.

SKEWED. A distribution that is asymmetrical.

SOURCE. The object, animal, or person from which infection is acquired.

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. There are two types of transmission: direct and indirect.

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

TYPE I ERRORS. Type I errors occur when the null hypothesis is rejected when it is true.

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.

USUAL SOURCE OF CARE. A place a patient usually goes when sick or when needing advice about health. Persons who report the emergency department as their usual source of care are defined as having no usual source of care.

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.

VECTOR-BORNE TRANSMISSION

Vectors are animals that are capable of transmitting diseases. Examples of vectors are flies, mites, fleas, ticks, rats, and dogs. The most common vector for disease is the mosquito. Mosquitoes transfer disease through the saliva, which comes in contact with their hosts when they are withdrawing blood. Mosquitoes are vectors for malaria, West Nile virus, dengue fever, and yellow fever.

Vectors add an extra dimension to disease transmission. Since vectors are mobile, they increase the transmission range of a disease. Changes in vector behavior will affect the transmission pattern of a disease. It is important to study the behavior of the vector as well as the disease-causing microorganism in order to establish a proper method of disease prevention. In the case of malaria, insecticides were sprayed and breeding grounds for mosquitoes were eliminated in an attempt to control the spread of malaria.

Biting is not the only way vectors can transmit diseases. Diseases may be spread through the feces of a vector. Microorganisms could also be located on the outside surface of a vector (such as a fly) and spread through physical contact with food, a common touch surface, or a susceptible individual.

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.