Course Name	**: Principles of Epidemiology**
Course Code	**: APBPH 2102**
Course Level	**: Level 3**
Credit Unit	**: 4 CU**
Contact Hours	**: 60 Hrs**

Course Description

This course of epidemiology explores its definition, its determinants, as well as its application. It also involves the most core epidemiological functions, describing the epidemiologic approach in public health, defining concepts of disease occurrence as well as chain of infection. The course also deals with understanding the concept of Immunology, branches of immunology, the importance of immunology, immune system, studying various disorders of the human immunity, organs of the immune system, cells in the innate immune system, Immunological Biomolecules and pathogenesis.

Course Objectives

· The Course introduces students to epidemiological knowledge that helps them to understand health related issues among the populations.
· It exposes them to measures used to control the chain of infection of diseases in case of its outbreak. Further still
· The course is intended to boast the student's capacity to track the cause of disease by use of methods like evaluation as one of the core epidemiological functions.
· To help students access a more firm understanding on the wider concept of immunology.
· To enable them learn measures to protect and increase their own immunity by eating balanced diets.
· To increase the students' capacity to identify illnesses that affect the human immunity.

Course Content

Introduction to Epidemiology

· Definition of Epidemiology
· Determinants of Epidemiology
· Health-related states or events
· Specified populations
· Application of Epidemiology

Core Epidemiology Functions

· Public Health Surveillance
· Field investigation
· Analytic studies
· Evaluation
· Linkages

· Policy development

The Epidemiologic Approach
· Defining a Case; Components of a case definition for outbreak investigation, criteria in case definition, modifying case definitions, variation definitions
· Using counts and rates
· Descriptive Epidemiology
· Analytic Epidemiology
· Experimental studies
· Observation studies
· Cohort study

Concepts of Disease Occurrence
· Causation
· Agent
· Host
· Environment
· Components causes and causal pies
· Natural History and Spectrum of disease

Chain of Infection
· Portal of exit
· Modes of transmission
· Portal of entry
· Host
· Implications for public health
· Epidemic disease occurrence
· Epidemic patterns
· Glossary

Mode of delivery Face to face lectures
Assessment
Course work 40%
Exams 60%
Total Mark 100%
PRINCIPLES OF EPIDEMIOLOGY

Epidemiology in Public Health

Recently, a news story described an inner-city neighborhood's concern about the rise in the number of children with asthma. Another story reported the revised recommendations for who should receive influenza vaccine this year. A third story discussed the extensive disease-monitoring strategies being implemented in a city recently affected by a massive hurricane. A fourth story described a finding published in a leading medical journal of an association in workers exposed to a particular chemical and an increased risk of cancer. Each

of these news stories included interviews with public health officials or researchers who called themselves epidemiologists. Well, who are these epidemiologists, and what do they do? What is epidemiology? This lesson is intended to answer those questions by describing what epidemiology is, how it has evolved and how it is used today, and what some of the key methods and concepts are. The focus is on epidemiology in public health practice, that is, the kind of epidemiology that is done at health departments.

Objectives

By the end of this Module, you will be able to:

- *Define epidemiology*
- *Summarize the historical evolution of epidemiology*
- *Name some of the key uses of epidemiology*
- *Identify the core epidemiology functions*
- *Describe primary applications of epidemiology in public health practice*
- *Specify the elements of a case definition and state the effect of changing the value of any of the elements*
- *List the key features and uses of descriptive epidemiology*
- *List the key features and uses of analytic epidemiology*
- *List the three components of the epidemiologic triad*
- *Describe the different modes of transmission of communicable disease in a population*

Definition of Epidemiology

The word epidemiology comes from the Greek words *epi*, meaning on or upon, *demos*, meaning people, and *logos*, meaning the study of. In other words, the word epidemiology has its roots in the study of what befalls a population. Many definitions have been proposed, but the following definition captures the underlying principles and public health spirit of epidemiology:

Epidemiology is the ***study*** *of the* ***distribution*** *and* ***determinants*** *of* ***health-related states or events*** *in* ***specifiedpopulations****, and the* ***application*** *of this study to the control of health problems.*

Key terms in this definition reflect some of the important principles of epidemiology.

Study:

Epidemiology is a scientific discipline with sound methods of scientific inquiry at its foundation. Epidemiology is data-driven and relies on a systematic and unbiased approach to the collection, analysis, and interpretation of data. Basic epidemiologic methods tend to rely on careful observation and use of valid comparison groups to assess whether what was observed, such as the number of cases of disease in a particular area during a particular time period or the frequency of an exposure among persons with disease, differs from what might be expected. However, epidemiology also draws on methods from other

scientific fields, including biostatistics and informatics, with biologic, economic, social, and behavioral sciences.

In fact, epidemiology is often described as the basic science of public health, and for good reason. First, epidemiology is a quantitative discipline that relies on a working knowledge of probability, statistics, and sound research methods. Second, epidemiology is a method of causal reasoning based on developing and testing hypotheses grounded in such scientific fields as biology, behavioral sciences, physics, and ergonomics to explain health-related behaviors, states, and events. However, epidemiology is not just a research activity but an integral component of public health, providing the foundation for directing practical and appropriate public health action based on this science and causal reasoning.

Distribution:
Epidemiology is concerned with the **frequency** and **pattern** of health events in a population:
Frequency refers not only to the number of health events such as the number of cases of meningitis or diabetes in a population, but also to the relationship of that number to the size of the population. The resulting rate allows epidemiologists to compare disease occurrence across different populations.
Pattern refers to the occurrence of health-related events by time, place, and person. Time patterns may be annual, seasonal, weekly, daily, hourly, weekday versus weekend, or any other breakdown of time that may influence disease or injury occurrence. Place patterns include geographic variation, urban/rural differences, and location of work sites or schools. Personal characteristics include demographic factors which may be related to risk of illness, injury, or disability such as age, sex, marital status, and socioeconomic status, as well as behaviors and environmental exposures. Characterizing health events by time, place, and person are activities of **descriptive epidemiology**, discussed in more detail later in this lesson.

Determinants:
Epidemiology is also used to search for **determinants**, which are the causes and other factors that influence the occurrence of disease and other health-related events. Epidemiologists assume that illness does not occur randomly in a population, but happens only when the right accumulation of risk factors or determinants exists in an individual. To search for these determinants, epidemiologists use analytic epidemiology or epidemiologic studies to provide the “Why” and “How” of such events. They assess whether groups with different rates of disease differ in their demographic characteristics, genetic or immunologic make-up, behaviors, environmental exposures, or other so-called potential risk factors. Ideally, the findings provide sufficient evidence to direct prompt and effective public health control and prevention measures.

Health-related states or events:

Epidemiology was originally focused exclusively on epidemics of communicable diseases but was subsequently expanded to address endemic communicable diseases and non-communicable infectious diseases. By the middle of the 20th Century, additional epidemiologic methods had been developed and applied to chronic diseases, injuries, birth defects, maternal-child health, occupational health, and environmental health. Then epidemiologists began to look at behaviors related to health and well-being, such as amount of exercise and seat belt use. Now, with the recent explosion in molecular methods, epidemiologists can make important strides in examining genetic markers of disease risk. Indeed, the term healthrelated states or events may be seen as anything that affects the well-being of a population. Nonetheless, many epidemiologists still use the term "disease" as shorthand for the wide range of health-related states and events that are studied.

Specified populations:

Although epidemiologists and direct health-care providers (clinicians) are both concerned with occurrence and control of disease, they differ greatly in how they view "the patient." The clinician is concerned about the health of an individual; the epidemiologist is concerned about the collective health of the people in a community or population. In other words, the clinician's "patient" is the individual; the epidemiologist's "patient" is the community. Therefore, the clinician and the epidemiologist have different responsibilities when faced with a person with illness. For example, when a patient with diarrheal disease presents, both are interested in establishing the correct diagnosis. However, while the clinician usually focuses on treating and caring for the individual, the epidemiologist focuses on identifying the exposure or source that caused the illness; the number of other persons who may have been similarly exposed; the potential for further spread in the community; and interventions to prevent additional cases or recurrences.

Application:

Epidemiology is not just "the study of" health in a population; it also involves applying the knowledge gained by the studies to community-based practice. Like the practice of medicine, the practice of epidemiology is both a science and an art. To make the proper diagnosis and prescribe appropriate treatment for a patient, the clinician combines medical (scientific) knowledge with experience, clinical judgment, and understanding of the patient.Similarly, the epidemiologist uses the scientific methods of descriptive and analytic epidemiology as well as experience, epidemiologic judgment, and understanding of local conditions in "diagnosing" the health of a community and proposing appropriate, practical, and acceptable public health interventions to control and prevent disease in the community.

Summary:

Epidemiology is the study (scientific, systematic, data-driven) of the distribution (frequency, pattern) and determinants (causes, risk factors) of

health-related states and events (not just diseases) in specified populations (patient is community, individuals viewed collectively), and the application of (since epidemiology is a discipline within public health) this study to the control of health problems.

Core Epidemiologic Functions

In the mid-1980s, five major tasks of epidemiology in public health practice were identified: **public health surveillance, fieldinvestigation, analytic studies, evaluation, and linkages.** A sixth task, **policy development**, was recently added. These tasks are described below.

Public health surveillance:

Public health surveillance is the ongoing, systematic collection, analysis, interpretation, and dissemination of health data to help guide public health decision making and action. Surveillance is equivalent to monitoring the pulse of the community. The purpose of public health surveillance, which is sometimes called "information for action,"18 is to portray the ongoing patterns of disease occurrence and disease potential so that investigation, control, and prevention measures can be applied efficiently and effectively. This is accomplished through the systematic collection and evaluation of morbidity and mortality reports and other relevant health information, and the dissemination of these data and their interpretation to those involved in disease control and public health decision making.

Morbidity and mortality reports are common sources of surveillance data for local and state health departments. These reports generally are submitted by health-care providers, infection control practitioners, or laboratories that are required to notify the health department of any patient with a reportable disease such as pertussis, meningococcal meningitis, or AIDS. Other sources of health-related data that are used for surveillance include reports from investigations of individual cases and disease clusters, public health program data such as immunization coverage in a community, disease registries, and health surveys.

Most often, surveillance relies on simple systems to collect a limited amount of information about each case. Although not every case of disease is reported, health officials regularly review the case reports they do receive and look for patterns among them. These practices have proven invaluable in detecting problems, evaluating programs, and guiding public health action.

While public health surveillance traditionally has focused on communicable diseases, surveillance systems now exist that target injuries, chronic diseases, genetic and birth defects, occupational and potentially environmentally-related diseases, and health behaviors. Since September 11, 2001, a variety of systems that rely on electronic reporting have been developed, including those that

report daily emergency department visits, sales of over-the-counter medicines, and worker absenteeism.19,20 Because epidemiologists are likely to be called upon to design and use these and other new surveillance systems, an epidemiologist's core competencies must include design of data collection instruments, data management, descriptive methods and graphing, interpretation of data, andscientific writing and presentation.

Field investigation:

As noted above, surveillance provides information for action. One of the first actions that results from a surveillance case report or report of a cluster is investigation by the public health department.

The investigation may be as limited as a phone call to the healthcare provider to confirm or clarify the circumstances of the reported case, or it may involve a field investigation requiring the coordinated efforts of dozens of people to characterize the extent of an epidemic and to identify its cause.

The objectives of such investigations also vary. Investigations often lead to the identification of additional unreported or unrecognized ill persons who might otherwise continue to spread infection to others. For example, one of the hallmarks of investigations of persons with sexually transmitted disease is the identification of sexual partners or contacts of patients. When interviewed, many of these contacts are found to be infected without knowing it, and are given treatment they did not realize they needed. Identification and treatment of these contacts prevents further spread. For some diseases, investigations may identify a source or vehicle of infection that can be controlled or eliminated.

For example, the investigation of a case of *Escherichia coli* O157:H7 infection usually focuses on trying to identify the vehicle, often ground beef but sometimes something more unusual such as fruit juice. By identifying the vehicle, investigators may be able to determine how many other persons might have already been exposed and how many continue to be at risk. When a commercial product turns out to be the culprit, public announcements and recalling the product may prevent many additional cases.

Occasionally, the objective of an investigation may simply be to learn more about the natural history, clinical spectrum, descriptive epidemiology, and risk factors of the disease before determining what disease intervention methods might be appropriate. Early investigations of the epidemic of SARS in 2003 were needed to establish a case definition based on the clinical presentation, and to characterize the populations at risk by time, place, and person. As more was learned about the epidemiology of the disease and communicability of the virus, appropriate recommendations regarding isolation and quarantine were issued.

Field investigations of the type described above are sometimes referred to as "shoe leather epidemiology," conjuring up images of dedicated, if haggard,

epidemiologists beating the pavement in search of additional cases and clues regarding source and mode of transmission. This approach is commemorated in the symbol of the Epidemic Intelligence Service (EIS), CDC's training program for disease detectives — a shoe with a hole in the sole.

Analytic studies:
Surveillance and field investigations are usually sufficient to identify causes, modes of transmission, and appropriate control and prevention measures. But sometimes analytic studies employing more rigorous methods are needed. Often the methods are used in combination — with surveillance and field investigations providing clues or hypotheses about causes and modes of transmission, and analytic studies evaluating the credibility of those hypotheses.

Clusters or outbreaks of disease frequently are investigated initially with descriptive epidemiology. The descriptive approach involves the study of disease incidence and distribution by time, place, and person. It includes the calculation of rates and identification of parts of the population at higher risk than others. Occasionally, when the association between exposure and disease is quite strong, the investigation may stop when descriptive epidemiology is complete and control measures may be implemented immediately.

John Snow's 1854 investigation of cholera is an example. More frequently, descriptive studies, like case investigations, generate hypotheses that can be tested with analytic studies. While some field investigations are conducted in response to acute health problems such as outbreaks, many others are planned studies.

The hallmark of an analytic epidemiologic study is the use of a valid comparison group. Epidemiologists must be skilled in all aspects of such studies, including design, conduct, analysis, interpretation, and communication of findings.
• **Design** includes determining the appropriate research strategy and study design, writing justifications and protocols, calculating sample sizes, deciding on criteria for subject selection (e.g., developing case definitions), choosing an appropriate comparison group, and designing questionnaires.
• **Conduct** involves securing appropriate clearances and approvals, adhering to appropriate ethical principles, abstracting records, tracking down and interviewing subjects, collecting and handling specimens, and managing the data.
• **Analysis** begins with describing the characteristics of the subjects. It progresses to calculation of rates, creation of comparative tables (e.g., two-by-two tables), and computation of measures of association (e.g., risk ratios or odds ratios), tests of significance (e.g., chi-square test), confidence intervals, and the like. Many epidemiologic studies require more advanced analytic techniques such as stratified analysis, regression, and modeling.

• Finally, **interpretation** involves putting the study findings into perspective, identifying the key take-home messages, and making sound recommendations. Doing so requires that the epidemiologist be knowledgeable about the subject matter and the strengths and weaknesses of the study.

Evaluation

Epidemiologists, who are accustomed to using systematic and quantitative approaches, have come to play an important role in evaluation of public health services and other activities. Evaluation is the process of determining, as systematically and objectively as possible, the relevance, effectiveness, efficiency, and impact of activities with respect to established goals.22

• **Effectiveness** refers to the ability of a program to produce the intended or expected results in the field; effectiveness differs from **efficacy**, which is the ability to produce results under ideal conditions.

• **Efficiency** refers to the ability of the program to produce the intended results with a minimum expenditure of time and resources.

The evaluation itself may focus on plans (formative evaluation), operations (process evaluation), impact (summative evaluation), or outcomes — or any combination of these. Evaluation of an immunization program, for example, might assess the efficiency of the operations, the proportion of the target population immunized, and the apparent impact of the program on the incidence of vaccine-preventable diseases. Similarly, evaluation of a surveillance system might address operations and attributes of the system, its ability to detect cases or outbreaks, and its usefulness.

Linkages:

Epidemiologists working in public health settings rarely act in isolation. In fact, field epidemiology is often said to be a "team sport." During an investigation an epidemiologist usually participates as either a member or the leader of a multidisciplinary team. Other team members may be laboratorians, sanitarians, infection control personnel, nurses or other clinical staff, and, increasingly, computer information specialists.

Many outbreaks cross geographical and jurisdictional lines, so co-investigators may be from local, state, or federal levels of government, academic institutions, clinical facilities, or the private sector. To promote current and future collaboration, the epidemiologists need to maintain relationships with staff of other agencies and institutions. Mechanisms for sustaining such linkages include official memoranda of understanding, sharing of published or on-line information for public health audiences and outside partners, and informal networking that takes place at professional meetings.

Policy development:

The definition of epidemiology ends with the following phrase: "…and the application of this study to the control of health problems." While some academically minded epidemiologists have stated that epidemiologists should

stick to research and not get involved in policy development or even make recommendations, public health epidemiologists do not have this luxury. Indeed, epidemiologists who understand a problem and the population in which it occurs are often in a uniquely qualified position to recommend appropriate interventions. As a result, epidemiologists working in public health regularly provide input, testimony, and recommendations regarding disease control strategies, reportable disease regulations, and health-care policy.

The Epidemiologic Approach

As with all scientific endeavors, the practice of epidemiology relies on a systematic approach. In very simple terms, the epidemiologist:

- **Counts** cases or health events, and describes them in terms of time, place, and person;
- **Divides** the number of cases by an appropriate denominator to calculate rates; and
- **Compares** these rates over time or for different groups of people.

Before counting cases, however, the epidemiologist must decide what a case is. This is done by developing a case definition. Then, using this case definition, the epidemiologist finds and collects information about the case-patients. The epidemiologist then performs descriptive epidemiology by characterizing the cases collectively according to time, place, and person. To calculate the disease rate, the epidemiologist divides the number of cases by the size of the population. Finally, to determine whether this rate is greater than what one would normally expect, and if so to identify factors contributing to this increase, the epidemiologist compares the rate from this population to the rate in an appropriate comparison group, using analytic epidemiology techniques. These epidemiologic actions are described in more detail below. Subsequent tasks, such as reporting the results and recommending how they can be used for public health action, are just as important, but are beyond the scope of this lesson.

Defining a case:

Before counting cases, the epidemiologist must decide what to count, that is, what to call a case. For that, the epidemiologist uses a **case definition**. A case definition is a set of standard criteria for classifying whether a person has a particular disease, syndrome, or other health condition. Some case definitions, particularly those used for national surveillance, have been developed and adopted as national standards that ensure comparability. Use of an agreedupon standard case definition ensures that every case is equivalent, regardless of when or where it occurred, or who identified it.

Furthermore, the number of cases or rate of disease identified in one time or place can be compared with the number or rate from another time or place. For example, with a standard case definition, health officials could compare the number of cases of listeriosis that occurred in Forsyth County, North Carolina

in 2000 with the number that occurred there in 1999. Or they could compare the rate of listeriosis in Forsyth County in 2000 with the national rate in that same year. When everyone uses the same standard case definition and a difference is observed, the difference is likely to be real rather than the result of variation in how cases are classified.

Components of a case definition for outbreak investigations:

A case definition consists of clinical criteria and, sometimes, limitations on time, place, and person. The clinical criteria usually include confirmatory laboratory tests, if available, or combinations of symptoms (subjective complaints), signs (objective physical findings), and other findings. Case definitions used during outbreak investigations are more likely to specify limits on time, place, and/or person than those used for surveillance. Contrast the case definition used for surveillance of listeriosis with the case definition used during an investigation of a listeriosis outbreak in North Carolina in 2000.

Both the national surveillance case definition and the outbreak case definition require a clinically compatible illness and laboratory confirmation of *Listeria monocytogenes*from a normally sterile site, but the outbreak case definition adds restrictions on time and place, reflecting the scope of the outbreak.

Many case definitions, such as that shown for listeriosis, require laboratory confirmation. This is not always necessary, however; in fact, some diseases have no distinctive laboratory findings.

Kawasaki syndrome, for example, is a childhood illness with fever and rash that has no known cause and no specifically distinctive laboratory findings. Notice that its case definition (see box below) is based on the presence of fever, at least four of five specified clinical findings, and the lack of a more reasonable explanation.

Criteria in case definitions:

A case definition may have several sets of criteria, depending on how certain the diagnosis is. For example, during an investigation of a possible case or outbreak of measles, a person with a fever and rash might be classified as having a suspected, probable, or confirmed case of measles, depending on what evidence of measles is present. A case might be classified as suspected or probable while waiting for the laboratory results to become available. Once the laboratory provides the report, the case can be reclassified as either confirmed or "not a case," depending on the laboratory results. In the midst of a large outbreak of a disease caused by a known agent, some cases may be permanently classified as suspected or probable because officials may feel that running laboratory tests on every patient with a consistent clinical picture and a history of exposure (e.g., chickenpox) is unnecessary and even wasteful. Case definitions should not rely on laboratory culture results alone, since organisms are sometimes present without causing disease.

Modifying case definitions:
Case definitions can also change over time as more information is obtained. The first case definition for SARS, based on clinical symptoms and either contact with a case or travel to an area with SARS transmission, was published in CDC's Morbidity and Mortality Weekly Report (MMWR) on March 21, 2003 (see box below). Two weeks later it was modified slightly. On March 29, after a novel coronavirus was determined to be the causative agent, an interim surveillance case definition was published that included laboratory criteria for evidence of infection with the SARSassociated coronavirus. By June, the case definition had changed several more times. In anticipation of a new wave of cases in 2004, a revised and much more complex case definition was published in December 2003.

Variation in case definitions:
Case definitions may also vary according to the purpose for classifying the occurrences of a disease. For example, health officials need to know as soon as possible if anyone has symptoms of plague or anthrax so that they can begin planning what actions to take. For such rare but potentially severe communicable diseases, for which it is important to identify every possible case, health officials use a sensitive case definition. A sensitive case definition is one that is broad or "loose," in the hope of capturing most or all of the true cases. For example, the case definition for a suspected case of rubella (German measles) is "any generalized rash illness of acute onset." This definition is quite broad, and would include not only all cases of rubella, but also measles, chickenpox, and rashes due to other causes such as drug allergies. So while the advantage of a sensitive case definition is that it includes most or all of the true cases, the disadvantage is that it sometimes includes other illnesses as well.

On the other hand, an investigator studying the causes of a disease outbreak usually wants to be certain that any person included in a study really had the disease. That investigator will prefer a specific or "strict" case definition. For instance, in an outbreak of *Salmonella* Agona infection, the investigators would be more likely to identify the source of the infection if they included only persons who were confirmed to have been infected with that organism, rather than including anyone with acute diarrhea, because some persons may have had diarrhea from a different cause. In this setting, the only disadvantages of a strict case definition are the requirement that everyone with symptoms be tested and an underestimation of the total number of cases if some people with salmonellosis are not tested.

Using counts and rates:
As noted, one of the basic tasks in public health is identifying and counting cases. These counts, usually derived from case reports submitted by health-care workers and laboratories to the health department, allow public health officials to determine the extent and patterns of disease occurrence by time,

place, and person. They may also indicate clusters or outbreaks of disease in the community. Counts are also valuable for health planning. For example, a health official might use counts (i.e., numbers) to plan how many infection control isolation units or doses of vaccine may be needed. However, simple counts do not provide all the information a health department needs. For some purposes, the counts must be put into context, based on the population in which they arose. Rates are measures that relate the numbers of cases during a certain period of time (usually per year) to the size of the population in which they occurred.

Descriptive Epidemiology

As noted earlier, every novice newspaper reporter is taught that a story is incomplete if it does not describe the what, who, where, when, and why/how of a situation, whether it be a space shuttle launch or a house fire. Epidemiologists strive for similar comprehensiveness in characterizing an epidemiologic event, whether it be a pandemic of influenza or a local increase in allterrain vehicle crashes. However, epidemiologists tend to use synonyms for the five W's listed above: case definition, person, place, time, and causes/risk factors/modes of transmission.

Descriptive epidemiology covers **time**, **place**, and **person**. Compiling and analyzing data by time, place, and person is desirable for several reasons.

• First, by looking at the data carefully, the epidemiologist becomes very familiar with the data. He or she can see what the data can or cannot reveal based on the variables available, its limitations (for example, the number of records with missing information for each important variable), and its eccentricities (for example, all cases range in age from 2 months to 6 years, plus one 17-yearold.).

• Second, the epidemiologist learns the extent and pattern of the public health problem being investigated — which months, which neighborhoods, and which groups of people have the most and least cases.

• Third, the epidemiologist creates a detailed description of the health of a population that can be easily communicated with tables, graphs, and maps.

• Fourth, the epidemiologist can identify areas or groups within the population that have high rates of disease. This information in turn provides important clues to the causes of the disease, and these clues can be turned into testable hypotheses.

Time:

The occurrence of disease changes over time. Some of these changes occur regularly, while others are unpredictable. Two diseases that occur during the same season each year include influenza (winter) and West Nile virus infection (August– September). In contrast, diseases such as hepatitis B and salmonellosis can occur at any time. For diseases that occur seasonally, health officials can anticipate their occurrence and implement control and prevention measures, such as an influenza vaccination campaign or mosquito spraying.

For diseases that occur sporadically, investigators can conduct studies to identify the causes and modes of spread, and then develop appropriately targeted actions to control or prevent further occurrence of the disease. In either situation, displaying the patterns of disease occurrence by time is critical for monitoring disease occurrence in the community and for assessing whether the public health interventions made a difference.

Day of week and time of day. For some conditions, displaying data by day of the week or time of day may be informative. Analysis at these shorter time periods is particularly appropriate for conditions related to occupational or environmental exposures that tend to occur at regularly scheduled intervals. Farm tractor fatalities are displayed by days of the week.32 Note that the number of farm tractor fatalities on Sundays was about half the number on the other days. The pattern of farm tractor injuries by hour, as displayed in Figure 1.8 peaked at 11:00 a.m., dipped at noon, and peaked again at 4:00 p.m. These patterns may suggest hypotheses and possible explanations that could be evaluated with further study. The hourly number of survivors and rescuers presenting to local hospitals in New York following the attack on the World Trade Center on September 11, 2001.

Epidemic period. To show the time course of a disease outbreak or epidemic, epidemiologists use a graph called an epidemic curve. As with the other graphs presented so far, an epidemic curve's y-axis shows the number of cases, while the x-axis shows time as either date of symptom onset or date of diagnosis. Depending on the incubation period (the length of time between exposure and onset of symptoms) and routes of transmission, the scale on the x-axis can be as broad as weeks (for a very prolonged epidemic) or as narrow as minutes (e.g., for food poisoning by chemicals that cause symptoms within minutes). Conventionally, the data are displayed as a histogram (which is similar to a bar chart but has no gaps between adjacent columns). Sometimes each case is displayed as a square. The shape and other features of an epidemic curve can suggest hypotheses about the time and source of exposure, the mode of transmission, and the causative agent.

Place:

Describing the occurrence of disease by place provides insight into the geographic extent of the problem and its geographic variation. Characterization by place refers not only to place of residence but to any geographic location relevant to disease occurrence. Such locations include place of diagnosis or report, birthplace, site of employment, school district, hospital unit, or recent travel destinations. The unit may be as large as a continent or country or as small as a street address, hospital wing, or operating room.

Sometimes place refers not to a specific location at all but to a place category such as urban or rural, domestic or foreign, and institutional or noninstitutional.

Analyzing data by place can identify communities at increased risk of disease. Even if the data cannot reveal why these people have an increased risk, it can help generate hypotheses to test with additional studies. For example, is a community at increased risk because of characteristics of the people in the community such as genetic susceptibility, lack of immunity, risky behaviors, or exposure to local toxins or contaminated food? Can the increased risk, particularly of a communicable disease, be attributed to characteristics of the causative agent such as a particularly virulent strain, hospitable breeding sites, or availability of the vector that transmits the organism to humans? Or can the increased risk be attributed to the environment that brings the agent and the host together, such as crowding in urban areas that increases the risk of disease transmission from person to person, or more homes being built in wooded areas close to deer that carry ticks infected with the organism that causes Lyme disease

Person:

Because personal characteristics may affect illness, organization and analysis of data by "person" may use inherent characteristics of people (for example, age, sex, race), biologic characteristics (immune status), acquired characteristics (marital status), activities (occupation, leisure activities, use of medications/tobacco/drugs), or the conditions under which they live (socioeconomic status, access to medical care). Age and sex are included in almost all data sets and are the two most commonly analyzed "person" characteristics. However, depending on the disease and the data available, analyses of other person variables are usually necessary. Usually epidemiologists begin the analysis of person data by looking at each variable separately. Sometimes, two variables such as age and sex can be examined simultaneously. Person data are usually displayed in tables or graphs.

Age. Age is probably the single most important "person" attribute, because almost every health-related event varies with age. A number of factors that also vary with age include: susceptibility, opportunity for exposure, latency or incubation period of the disease, and physiologic response (which affects, among other things, disease development). When analyzing data by age, epidemiologists try to use age groups that are narrow enough to detect any age-related patterns that may be present in the data. For some diseases, particularly chronic diseases, 10-year age groups may be adequate. For other diseases, 10-year and even 5-year age groups conceal important variations in disease occurrence by age. Consider the graph of pertussis occurrence by standard 5-year age groups.

Ethnic and racial groups. Sometimes epidemiologists are interested in analyzing person data by biologic, cultural or social groupings such as race, nationality, religion, or social groups such as tribes and other geographically or socially isolated groups. Differences in racial, ethnic, or other group variables may reflect differences in susceptibility or exposure, or differences in other

factors that influence the risk of disease, such as socioeconomic status and access to health care. Infant mortality rates for 2002 are shown by race and Hispanic origin of the mother.

Socioeconomic status. Socioeconomic status is difficult to quantify. It is made up of many variables such as occupation, family income, educational achievement or census track, living conditions, and social standing. The variables that are easiest to measure may not accurately reflect the overall concept.
Nevertheless, epidemiologists commonly use occupation, family income, and educational achievement, while recognizing that these variables do not measure socioeconomic status precisely. The frequency of many adverse health conditions increases with decreasing socioeconomic status. For example, tuberculosis is more common among persons in lower socioeconomic strata. Infant mortality and time lost from work due to disability are both associated with lower income.

These patterns may reflect more harmful exposures, lower resistance, and less access to health care. Or they may in part reflect an interdependent relationship that is impossible to untangle: Does low socioeconomic status contribute to disability, or does disability contribute to lower socioeconomic status, or both? What accounts for the disproportionate prevalence of diabetes and asthma in lower socioeconomic areas? A few adverse health conditions occur more frequently among persons of higher socioeconomic status. Gout was known as the "disease of kings" because of its association with consumption of rich foods. Other conditions associated with higher socioeconomic status include breast cancer, Kawasaki syndrome, chronic fatigue syndrome, and tennis elbow. Differences in exposure account for at least some if not most of the differences in the frequency of these conditions.

Analytic Epidemiology
As noted earlier, descriptive epidemiology can identify patterns among cases and in populations by time, place and person. From these observations, epidemiologists develop hypotheses about the causes of these patterns and about the factors that increase risk of disease. In other words, epidemiologists can use descriptive epidemiology to generate hypotheses, but only rarely to test those hypotheses. For that, epidemiologists must turn to analytic epidemiology.

The key feature of analytic epidemiology is a comparison group. Consider a large outbreak of hepatitis A that occurred in Pennsylvania in 2003.38 Investigators found almost all of the casepatients had eaten at a particular restaurant during the 2–6 weeks (i.e., the typical incubation period for hepatitis A) before onset of illness. While the investigators were able to narrow down their hypotheses to the restaurant and were able to exclude the food preparers and servers as the source, they did not know which particular food may have

been contaminated. The investigators asked the case-patients which restaurant foods they had eaten, but that only indicated which foods were popular.

The investigators, therefore, also enrolled and interviewed a comparison or control group — a group of persons who had eaten at the restaurant during the same period but who did not get sick. Of 133 items on the restaurant's menu, the most striking difference between the case and control groups was in the proportion that ate salsa (94% of case-patients ate, compared with 39% of controls). Further investigation of the ingredients in the salsa implicated green onions as the source of infection. Shortly thereafter, the Food and Drug Administration issued an advisory to the public about green onions and risk of hepatitis A. This action was in direct response to the convincing results of the analytic epidemiology, which compared the exposure history of case-patients with that of an appropriate comparison group. When investigators find that persons with a particular characteristic are more likely than those without the characteristic to contract a disease, the characteristic is said to be associated with the disease. The characteristic may be a:

- Demographic factor such as age, race, or sex;
- Constitutional factor such as blood group or immune status;
- Behavior or act such as smoking or having eaten salsa; or
- Circumstance such as living near a toxic waste site.

Identifying factors associated with disease help health officials appropriately target public health prevention and control activities. It also guides additional research into the causes of disease.

Thus, analytic epidemiology is concerned with the search for causes and effects, or the why and the how. Epidemiologists use analytic epidemiology to quantify the association between exposures and outcomes and to test hypotheses about causal relationships. It has been said that epidemiology by itself can never prove that a particular exposure caused a particular outcome. Often, however, epidemiology provides sufficient evidence to take appropriate control and prevention measures.

Epidemiologic studies fall into two categories: **experimental** and **observational**.

Experimental studies:

In an experimental study, the investigator determines through a controlled process the exposure for each individual (clinical trial) or community (community trial), and then tracks the individuals or communities over time to detect the effects of the exposure. For example, in a clinical trial of a new vaccine, the investigator may randomly assign some of the participants to receive the new vaccine, while others receive a placebo shot. The investigator then tracks all participants, observes who gets the disease that the new vaccine is intended to prevent, and compares the two groups (new vaccine vs. placebo) to see whether the vaccine group has a lower rate of disease. Similarly, in a trial to prevent onset of diabete among high-risk individuals, investigators

randomly assigned enrollees to one of three groups — placebo, an anti-diabetes drug, or lifestyle intervention. At the end of the follow-up period, investigators found the lowest incidence of diabetes in the lifestyle intervention group, the next lowest in the anti-diabetic drug group, and the highest in the placebo group.

Observational studies:
In an observational study, the epidemiologist simply observes the exposure and disease status of each study participant. John Snow's studies of cholera in London were observational studies. The two most common types of observational studies are cohort studies and case-control studies; a third type is cross-sectional studies.

Cohort study. A cohort study is similar in concept to the experimental study. In a cohort study the epidemiologist records whether each study participant is exposed or not, and then tracks the participants to see if they develop the disease of interest. Note that this differs from an experimental study because, in a cohort study, the investigator observes rather than determines the participants' exposure status. After a period of time, the investigator compares the disease rate in the exposed group with
the disease rate in the unexposed group. The unexposed group serves as the comparison group, providing an estimate of the baseline or expected amount of disease occurrence in the community. If the disease rate is substantively different in the exposed group compared to the unexposed group, the exposure is said to be associated with illness.

The length of follow-up varies considerably. In an attempt to respond quickly to a public health concern such as an outbreak, public health departments tend to conduct relatively brief studies.
On the other hand, research and academic organizations are more likely to conduct studies of cancer, cardiovascular disease, and other chronic diseases which may last for years and even decades.
The Framingham study is a well-known cohort study that has followed over 5,000 residents of Framingham, Massachusetts, since the early 1950s to establish the rates and risk factors for heart disease. The Nurses Health Study and the Nurses Health Study II are cohort studies established in 1976 and 1989, respectively, that have followed over 100,000 nurses each and have provided useful information on oral contraceptives, diet, and lifestyle risk factors. These studies are sometimes called **follow-up** or **prospective** cohort studies, because participants are enrolled as the study beginsand are then followed prospectively over time to identifyoccurrence of the outcomes of interest.

An alternative type of cohort study is a **retrospective** cohort study. In this type of study both the exposure and the outcomes have already occurred. Just as in a prospective cohort study, the investigator calculates and compares rates of

disease in the exposed and unexposed groups. Retrospective cohort studies are commonly used in investigations of disease in groups of easily identified people such as workers at a particular factory or attendees at a wedding. For example, a retrospective cohort study was used to determine the source of infection of cyclosporiasis, a parasitic disease that caused an outbreak among members of a residential facility in Pennsylvania in 2004. The investigation indicated that consumption of snow peas was implicated as the vehicle of the cyclosporiasis outbreak.

Case-control study. In a case-control study, investigators start by enrolling a group of people with disease (at CDC such persons are called case-patients rather than cases, because case refers to occurrence of disease, not a person). As a comparison group, the investigator then enrolls a group of people without disease (controls). Investigators then compare previous exposures between the two groups. The control group provides an estimate of the baseline or expected amount of exposure in that population. If the amount of exposure among the case group is substantially higher than the amount you would expect based on the control group, then illness is said to be associated with that exposure.

The study of hepatitis A traced to green onions, described above, is an example of a case-control study. The key in a case-control study is to identify an appropriate control group, comparable to the case group in most respects, in order to provide a reasonable estimate of the baseline or expected exposure.

Cross-sectional study. In this third type of observational study, a sample of persons from a population is enrolled and their exposures and health outcomes are measured simultaneously. The cross-sectional study tends to assess the presence (prevalence) of the health outcome at that point of time without regard to duration. For example, in a cross-sectional study of diabetes, some of the enrollees with diabetes may have lived with their diabetes for many years, while others may have been recently diagnosed. From an analytic viewpoint the cross-sectional study is weaker than either a cohort or a case-control study because a crosssectional study usually cannot disentangle risk factors for occurrence of disease (incidence) from risk factors for survival with the disease. On the other hand, a cross-sectional study is a perfectly fine tool for descriptive epidemiology purposes. Crosssectional studies are used routinely to document the prevalence in a community of health behaviors (prevalence of smoking), health states (prevalence of vaccination against measles), and health outcomes, particularly chronic conditions (hypertension, diabetes).

In summary, the purpose of an analytic study in epidemiology is to identify and quantify the relationship between an exposure and a health outcome. The hallmark of such a study is the presence of at least two groups, one of which serves as a comparison group. In an experimental study, the investigator determines the exposure for the study subjects; in an observational study, the subjects are exposed under more natural conditions. In an observational cohort study, subjects are enrolled or grouped on the basis of their exposure,

then are followed to document occurrence of disease. Differences in disease rates between the exposed and unexposed groups lead investigators to conclude that exposure is associated with disease. In an observational case-control study, subjects are enrolled according to whether they have the disease or not, then are questioned or tested to determine their prior exposure. Differences in exposure prevalence between the case and control groups allow investigators to conclude that the exposure is associated with the disease. Cross-sectional studies measure exposure and disease status at the same time, and are better suited to descriptive epidemiology than causation.

Concepts of Disease Occurrence

A critical premise of epidemiology is that disease and other healthevents do not occur randomly in a population, but are more likelyto occur in some members of the population than others because ofrisk factors that may not be distributed randomly in the population.As noted earlier, one important use of epidemiology is to identifythe factors that place some members at greater risk than others.

Causation:

A number of models of disease causation have been proposed. Among the simplest of these is the epidemiologic triad or triangle, the traditional model for infectious disease. The triad consists of an external **agent**, a susceptible **host**, and an **environment** that brings the host and agent together. In this model, disease results from the interaction between the agent and the susceptible host in an environment that supports transmission of the agent from a source to that host. Two ways of depicting this model are shown in Agent, host, and environmental factors interrelate in a variety of complex ways to produce disease. Different diseases require different balances and interactions of these three components. Development of appropriate, practical, and effective public health measures to control or prevent disease usually requires assessment of all three components and their interactions.

Agent originally referred to an infectious microorganism or pathogen: a virus, bacterium, parasite, or other microbe. Generally, the agent must be present for disease to occur; however, presence of that agent alone is not always sufficient to cause disease. A variety of factors influence whether exposure to an organism will result in disease, including the organism's pathogenicity (ability to cause disease) and dose. Over time, the concept of agent has been broadened to include chemical and physical causes of disease or injury. These include chemical contaminants (such as the L-tryptophan contaminant responsible for eosinophilia-myalgia syndrome), as well as physical forces (such as repetitive mechanical forces associated with carpal tunnel syndrome). While the epidemiologic triad serves as a useful model for many diseases, it has proven inadequate for cardiovascular disease, cancer, and other diseases that appear to have multiple contributing causes without a single necessary one.

Host refers to the human who can get the disease. A variety of factors intrinsic to the host, sometimes called risk factors, can influence an individual's exposure, susceptibility, or response to a causative agent. Opportunities for exposure are often influenced by behaviors such as sexual practices, hygiene, and other personal choices as well as by age and sex. Susceptibility and response to an agent are influenced by factors such as genetic composition, nutritional and immunologic status, anatomic structure, presence of disease or medications, and psychological makeup.

Environment refers to extrinsic factors that affect the agent and the opportunity for exposure. Environmental factors include physical factors such as geology and climate, biologic factors such as insects that transmit the agent, and socioeconomic factors such as crowding, sanitation, and the availability of health services.

Component causes and causal pies:
Because the agent-host-environment model did not work well for many non-infectious diseases, several other models that attempt to account for the multifactorial nature of causation have been proposed. One such model was proposed by Rothman in 1976, and has come to be known as the Causal Pies. An individual factor that contributes to cause disease is shown as a piece of a pie. After all the pieces of a pie fall into place, the pie is complete — and disease occurs. The individual factors are called **component causes**. The complete pie, which might be considered a causal pathway, is called a **sufficientcause**. A disease may have more than one sufficient cause, with each sufficient cause being composed of several component causes that may or may not overlap. A component that appears in every pie or pathway is called a **necessary cause**, because without it, disease does not occur.

The component causes may include intrinsic host factors as well as the agent and the environmental factors of the agent-hostenvironment triad. A single component cause is rarely a sufficient cause by itself. For example, even exposure to a highly infectious agent such as measles virus does not invariably result in measles disease. Host susceptibility and other host factors also may play a role. At the other extreme, an agent that is usually harmless in healthy persons may cause devastating disease under different conditions.

*Pneumocystis carinii*is an organism that harmlessly colonizes the respiratory tract of some healthy persons, but can cause potentially lethal pneumonia in persons whose immune systems have been weakened by human immunodeficiency virus (HIV). Presence of *Pneumocystis carinii*organisms is therefore a necessary but not sufficient cause of pneumocystis pneumonia. As the model indicates, a particular disease may result from a variety of different sufficient causes or pathways. For example, lung cancer may result from a sufficient cause that includes smoking as a component cause. Smoking is not a

sufficient cause by itself, however, because not all smokers develop lung cancer.

Neither is smoking a necessary cause, because a small fraction of lung cancer victims have never smoked. Suppose Component Cause B is smoking and Component Cause C is asbestos. Sufficient Cause I includes both smoking (B) and asbestos (C). Sufficient Cause II includes asbestos without smoking, and Sufficient Cause C includes smoking without asbestos. But because lung cancer can develop in persons who have never been exposed to either smoking or asbestos, a proper model for lung cancer would have to show at least one more Sufficient Cause Pie that does not include either component B or component C.

Note that public health action does not depend on the identification of every component cause. Disease prevention can be accomplished by blocking any single component of a sufficient cause, at least through that pathway. For example, elimination of smoking (component B) would prevent lung cancer from sufficient causes I and II, although some lung cancer would still occur through sufficient cause III.

Natural History and Spectrum of Disease

Natural history of disease refers to the progression of a disease process in an individual over time, in the absence of treatment. For example, untreated infection with HIV causes a spectrum of clinical problems beginning at the time of seroconversion (primary HIV) and terminating with AIDS and usually death. It is now recognized that it may take 10 years or more for AIDS to develop after seroconversion.43 Many, if not most, diseases have a characteristic natural history, although the time frame and specific manifestations of disease may vary from individual to individual and are influenced by preventive and therapeutic measures.

The process begins with the appropriate exposure to or accumulation of factors sufficient for the disease process to begin in a susceptible host. For an infectious disease, the exposure is a microorganism. For cancer, the exposure may be a factor that initiates the process, such as asbestos fibers or components in tobacco smoke (for lung cancer), or one that promotes the process, such as estrogen (for endometrial cancer). After the disease process has been triggered, pathological changes then occur without the individual being aware of them. This stage of subclinical disease, extending from the time of exposure to onset of disease symptoms, is usually called the **incubation period** for infectious diseases, and the **latency period** for chronic diseases. During this stage, disease is said to be asymptomatic (no symptoms) or inapparent. This period may be as brief as seconds for hypersensitivity and toxic reactions to as long as decades for certain chronic diseases. Even for a single disease, the characteristic incubation period has a range. For example, the typical incubation period for hepatitis A is as long as 7 weeks. The latency

period for leukemia to become evident among survivors of the atomic bomb blast in Hiroshima ranged from 2 to 12 years, peaking at 6-7 years.44 Incubation periods of selected exposures and diseases varying from minutes to decades are displayed in

Although disease is not apparent during the incubation period, some pathologic changes may be detectable with laboratory, radiographic, or other screening methods. Most screening programs attempt to identify the disease process during this phase of its natural history, since intervention at this early stage is likely to be more effective than treatment given after the disease has progressed and become symptomatic. The onset of symptoms marks the transition from subclinical to clinical disease. Most diagnoses are made during the stage of clinical disease. In some people, however, the disease process may never progress to clinically apparent illness. In others, the disease process may result in illness that ranges from mild to severe or fatal. This range is called the **spectrum of disease**. Ultimately, the disease process ends either in recovery, disability or death.

For an infectious agent, **infectivity** refers to the proportion of exposed persons who become infected. **Pathogenicity** refers to the proportion of infected individuals who develop clinically apparent disease. **Virulence** refers to the proportion of clinically apparent cases that are severe or fatal. Because the spectrum of disease can include asymptomatic and mild cases, the cases of illness diagnosed by clinicians in the community often represent only the tip of the iceberg. Many additional cases may be too early to diagnose or may never progress to the clinical stage. Unfortunately, persons with inapparent or undiagnosed infections may nonetheless be able to transmit infection to others. Such persons who are infectious but have subclinical disease are called **carriers**. Frequently, carriers are persons with incubating disease or inapparent infection. Persons with measles, hepatitis A, and several other diseases become infectious a few days before the onset of symptoms. However carriers may also be persons who appear to have recovered from their clinical illness but remain infectious, such as chronic carriers of hepatitis B virus, or persons who never exhibited symptoms. The challenge to public health workers is that these carriers, unaware that they are infected and infectious to others, are sometimes more likely to unwittingly spread infection than are people with obvious illness.

Chain of Infection

As described above, the traditional epidemiologic triad model holds that infectious diseases result from the interaction of agent, host, and environment. More specifically, transmission occurs when the agent leaves its **reservoir** or host through a **portal ofexit**, is conveyed by some **mode of transmission**, and enters through an appropriate **portal of entry** to infect a **susceptiblehost**. This sequence is sometimes called the chain of infection.

Portal of exit:
Portal of exit is the path by which a pathogen leaves its host. The portal of exit usually corresponds to the site where the pathogen is localized. For example, influenza viruses and *Mycobacteriumtuberculosis* exit the respiratory tract, schistosomes through urine, cholera vibrios in feces, *Sarcoptesscabiei* in scabies skin lesions, and enterovirus 70, a cause of hemorrhagic conjunctivitis, in conjunctival secretions. Some bloodborne agents can exit by crossing the placenta from mother to fetus (rubella, syphilis, toxoplasmosis), while others exit through cuts or needles in the skin (hepatitis B) or blood-sucking arthropods (malaria).

Modes of transmission:
An infectious agent may be transmitted from its natural reservoir to a susceptible host in different ways. There are different classifications for modes of transmission. Here is one classification:
• Direct (Direct contact and Droplet spread)
• Indirect (Airborne and VehicleborneVectorborne (mechanical or biologic))

In **direct transmission**, an infectious agent is transferred from a reservoir to a susceptible host by direct contact or droplet spread.
Direct contact occurs through skin-to-skin contact, kissing, and sexual intercourse. Direct contact also refers to contact with soil or vegetation harboring infectious organisms. Thus, infectious mononucleosis ("kissing disease") and gonorrhea are spread from person to person by direct contact. Hookworm is spread by direct contact with contaminated soil.
Droplet spread *r*efers to spray with relatively large, short-range aerosols produced by sneezing, coughing, or even talking. Droplet spread is classified as direct because transmission is by direct spray over a few feet, before the droplets fall to the ground. Pertussis and meningococcal infection are examples of diseases transmitted from an infectious patient to a susceptible host by droplet spread.
Indirect transmission refers to the transfer of an infectious agent from a reservoir to a host by suspended air particles, inanimate objects (vehicles), or animate intermediaries (vectors).
Airborne transmission occurs when infectious agents are carried by dust or droplet nuclei suspended in air. Airborne dust includes material that has settled on surfaces and become resuspended by air currents as well as infectious particles blown from the soil by the wind. Droplet nuclei are dried residue of less than 5 microns in size. In contrast to droplets that fall to the ground within a few feet, droplet nuclei may remain suspended in the air for long periods of time and may be blown over great distances. Measles, for example, has occurred in children who came into a physician's office after a child with measles had left, because the measles virus remained suspended in the air.

Vehicles that may indirectly transmit an infectious agent include food, water, biologic products (blood), and fomites (inanimate objects such as handkerchiefs, bedding, or surgical scalpels). A vehicle may passively carry a pathogen — as food or water may carry hepatitis A virus. Alternatively, the vehicle may provide an environment in which the agent grows, multiplies, or produces toxin — as improperly canned foods provide an environment that supports production of botulinum toxin by *Clostridiumbotulinum*.
Vectors such as mosquitoes, fleas, and ticks may carry an infectious agent through purely mechanical means or may support growth or changes in the agent. Examples of mechanical transmission are flies carrying *Shigella*on their appendages and fleas carrying *Yersinia pestis*, the causative agent of plague, in their gut. In contrast, in biologic transmission, the causative agent of malaria or guinea worm disease undergoes maturation in an intermediate host before it can be transmitted to humans.

Portal of entry:
The portal of entry refers to the manner in which a pathogen enters a susceptible host. The portal of entry must provide access to tissues in which the pathogen can multiply or a toxin can act. Often, infectious agents use the same portal to enter a new host that they used to exit the source host. For example, influenza virus exits the respiratory tract of the source host and enters the respiratory tract of the new host. In contrast, many pathogens that cause gastroenteritis follow a so-called “fecal-oral” route because they exit the source host in feces, are carried on inadequately washed hands to a vehicle such as food, water, or utensil, and enter a new host through the mouth. Other portals of entry include the skin (hookworm), mucous membranes (syphilis), and blood (hepatitis B, human immunodeficiency virus).

Host:
The final link in the chain of infection is a susceptible host. Susceptibility of a host depends on genetic or constitutional factors, specific immunity, and nonspecific factors that affect an individual’s ability to resist infection or to limit pathogenicity. An individual’s genetic makeup may either increase or decrease susceptibility. For example, persons with sickle cell trait seem to be at least partially protected from a particular type of malaria. Specific immunity refers to protective antibodies that are directed against a specific agent. Such antibodies may develop in responsemto infection, vaccine, or toxoid (toxin that has been deactivated but retains its capacity to stimulate production of toxin antibodies) or may be acquired by transplacental transfer from mother to fetus or by injection of antitoxin or immune globulin. Nonspecific factors that defend against infection include the skin, mucous membranes, gastric acidity, cilia in the respiratory tract, the cough reflex, and nonspecific immune response. Factors that may increase susceptibility to infection by disrupting host defenses include malnutrition, alcoholism, and disease or therapy that impairs the nonspecific immune response.

Implications for public health:
Knowledge of the portals of exit and entry and modes of transmission provides a basis for determining appropriate control measures. In general, control measures are usually directed against the segment in the infection chain that is most susceptible to intervention, unless practical issues dictate otherwise.
For some diseases, the most appropriate intervention may be directed at controlling or eliminating the agent at its source. A patient sick with a communicable disease may be treated with antibiotics to eliminate the infection. An asymptomatic but infected person may be treated both to clear the infection and to reduce the risk of transmission to others. In the community, soil may be decontaminated or covered to prevent escape of the agent.

Some interventions are directed at the mode of transmission:
Interruption of direct transmission may be accomplished by isolation of someone with infection, or counseling persons to avoid the specific type of contact associated with transmission. Vehicleborne transmission may be interrupted by elimination or decontamination of the vehicle. To prevent fecal-oral transmission, efforts often focus on rearranging the environment to reduce the risk of contamination in the future and on changing behaviors, such as promoting handwashing. For airborne diseases, strategies may be directed at modifying ventilation or air pressure, and filtering or treating the air. To interrupt vectorborne transmission, measures may be directed toward controlling the vector population, such as spraying to reduce the mosquito population.

Some strategies that protect portals of entry are simple and effective. For example, bed nets are used to protect sleeping persons from being bitten by mosquitoes that may transmit malaria. A dentist's mask and gloves are intended to protect the dentist from a patient's blood, secretions, and droplets, as well to protect the patient from the dentist. Wearing of long pants and sleeves and use of insect repellent are recommended to reduce the risk of Lyme disease and West Nile virus infection, which are transmitted by the bite of ticks and mosquitoes, respectively.

Some interventions aim to increase a host's defenses. Vaccinations promote development of specific antibodies that protect against infection. On the other hand, prophylactic use of antimalarial drugs, recommended for visitors to malaria-endemic areas, does not prevent exposure through mosquito bites, but does prevent infection from taking root.

Finally, some interventions attempt to prevent a pathogen from encountering a susceptible host. The concept of **herd immunity** suggests that if a high enough proportion of individuals in a population are resistant to an agent, then those few who are susceptible will be protected by the resistant majority, since the pathogen will be unlikely to "find" those few susceptible individuals. The degree

of herd immunity necessary to prevent or interrupt an outbreak varies by disease. In theory, herd immunity means that not everyone in a community needs to be resistant (immune) to prevent disease spread and occurrence of an outbreak. In practice, herd immunity has not prevented outbreaks of measles and rubella in populations with immunization levels as high as 85% to 90%. One problem is that, in highly immunized populations, the relatively few susceptible persons are often clustered in subgroups defined by socioeconomic or cultural factors. If the pathogen is introduced into one of these subgroups, an outbreak may occur.

Epidemic Disease Occurrence

Level of disease

The amount of a particular disease that is usually present in a community is referred to as the baseline or **endemic** level of the disease. This level is not necessarily the desired level, which may in fact be zero, but rather is the observed level. In the absence of intervention and assuming that the level is not high enough to deplete the pool of susceptible persons, the disease may continue to occur at this level indefinitely. Thus, the baseline level is often regarded as the expected level of the disease. While some diseases are so rare in a given population that a single case warrants an epidemiologic investigation (e.g., rabies, plague, polio), other diseases occur more commonly so that only deviations from the norm warrant investigation.

Sporadic refers to a disease that occurs infrequently and irregularly.

Endemic refers to the constant presence and/or usual prevalence of a disease or infectious agent in a population within a geographic area.

Hyperendemicrefers to persistent, high levels of disease occurrence. Occasionally, the amount of disease in a community rises above the expected level.

Epidemic refers to an increase, often sudden, in the number of cases of a disease above what is normally expected in that population in that area.

Outbreak carries the same definition of epidemic, but is often used for a more limited geographic area. **Cluster** refers to an aggregation of cases grouped in place and time that are suspected to be greater than the number expected, even though the expected number may not be known.

Pandemic refers to an epidemic that has spread over several countries or continents, usually affecting a large number of people.

Epidemics occur when an agent and susceptible hosts are present in adequate numbers, and the agent can be effectively conveyed from a source to the susceptible hosts. More specifically, an epidemic may result from:

• A recent increase in amount or virulence of the agent,

• The recent introduction of the agent into a setting where it has not been before,

• An enhanced mode of transmission so that more susceptible persons are exposed,

• A change in the susceptibility of the host response to the agent, and/or

• Factors that increase host exposure or involve introduction through new portals of entry.
The previous description of epidemics presumes only infectious agents, but non-infectious diseases such as diabetes and obesity exist in epidemic proportion.

Epidemic Patterns:
Epidemics can be classified according to their manner of spread through a population:
• Common-source
• Point
• Continuous
• Intermittent
• Propagated
• Mixed
• Other
A **common-source outbreak** is one in which a group of persons are all exposed to an infectious agent or a toxin from the same source. If the group is exposed over a relatively brief period, so that everyone who becomes ill does so within one incubation period, then the common-source outbreak is further classified as a **pointsourceoutbreak**. The epidemic of leukemia cases in Hiroshima following the atomic bomb blast and the epidemic of hepatitis A among patrons of the Pennsylvania restaurant who ate green onions each had a point source of exposure.38, 44 If the number of cases during an epidemic were plotted over time, the resulting graph, called an epidemic curve, would typically have a steep upslope and a more gradual downslope (a so-called "log-normal distribution").

GLOSSARY
Active immunity see immunity, active.
active surveillance see surveillance, active.
age-adjusted mortality rate see mortality rate, age-adjusted.
agenta factor (e.g., a microorganism or chemical substance) or form of energy whose presence, excessive presence, or in the case of deficiency diseases, relative absence is essential for the occurrence of a disease or other adverse health outcome.
age-specific mortality rate see mortality rate, age-specific.
alternative hypothesis see hypothesis, alternative.
analytic epidemiology see epidemiology, analytic.
analytic study see study, analytic.
antibodyany of a variety of proteins in the blood that are produced in response to an antigen as an immune response.
antigenany substance (e.g., a toxin or the surface of a microorganism or transplanted organ) recognized as foreign by the human body and that stimulates the production of antibodies.

applied epidemiology see epidemiology, applied.
arbovirusany of a group of viruses that are transmitted between hosts by mosquitoes, ticks, and other arthropods.
arithmetic mean see mean, arithmetic.
arithmetic-scale line graph see line graph, arithmetic-scale.
arthropodan organism that has jointed appendages and segmented external skeleton (e.g., flies, mosquitoes, ticks, or mites).
associationthe statistical relation between two or more events, characteristics, or other variables.
asymmetricala type of distribution where the shape to the right and left of the central location is not the same. Often referred to as a skewed distribution; the mean, median, and mode of an asymmetrical distribution are not the same.
asymptomaticwithout symptoms.
attack rate a form of incidence that measures the proportion of persons in a population who experience an acute health event during a limited period (e.g., during an outbreak), calculated as the number of new cases of a health problem during an outbreak divided by the size of the population at the beginning of the period, usually expressed as a percentage or per 1,000 or 100,000 population (see also **incidence proportion**).
attack rate, secondary a measure of the frequency of new cases of a disease among the contacts of known patients.
attributable proportion see**proportion, attributable**.
attributable risk percent see **proportion, attributable**.
attributea risk factor that is an intrinsic characteristic of the individual person, animal, plant, or other type of organism under study (e.g., genetic susceptibility, age, sex, breed, weight).
axisone of the dimensions of a graph in a rectangular graph, the x-axis is the horizontal axis, and the y-axis is the vertical axis.

Bar chart a visual display in which each category of a variable is represented by a bar or column bar charts are used to illustrate variations in size among categories.
bar chart, 100% component a stacked bar chart in which all bars or columns are the same length, and the measured axis represents 0%–100%.
bar chart, deviation a bar chart displaying either positive or negative differences from a baseline.
bar chart, grouped a bar chart displaying quantities of two variables, represented by adjoining bars or columns (i.e., a group) of categories of one variable, separated by space between groups.
bar chart, stacked a bar chart displaying quantities of two variables, represented by subdivided bars or columns (the subdivisions representing the categories of one variable) separated by space between bars or columns.
biasa systematic deviation of results or inferences from the truth or processes leading to such systematic deviation; any systematic tendency in the collection, analysis, interpretation,

publication, or review of data that can lead to conclusions that are systematically different from the truth. In epidemiology, does not imply intentional deviation.

bias, information systematic difference in the collection of data regarding the participants in a study (e.g., about exposures in a case-control study, or about health outcomes in a cohort study) that leads to an incorrect result (e.g., risk ratio or odds ratio) or inference.

bias, selection systematic difference in the enrollment of participants in a study that leads to an incorrect result (e.g., risk ratio or odds ratio) or inference.

bimodalhaving two data peaks.

biologic transmission see **transmission, biologic**.

birth cohort see **cohort, birth**.

birth rate, crude the number of live births during a specified period divided by the mid-period population, usually expressed per 1,000 population.

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

Carrier a person or animal that harbors the infectious agent for a disease and can transmit it toothers, but does not demonstrate signs of the disease. A carrier can be asymptomatic (neverindicate signs of the disease) or can display signs of the disease only during the incubationperiod, convalescence, or postconvalescence. The period of being a carrier can be short (atransient carrier) or long (a chronic carrier).

casean instance of a particular disease, injury, or other health conditions that meets selected criteria (see also **case definition**). Using the term to describe the person rather than the health condition is discouraged (see also **case-patient**).

case-control study see **study, case-control**.

case definition a set of uniformly applied criteria for determining whether a person should be identified as having a particular disease, injury, or other health condition. In epidemiology, particularly for an outbreak investigation, a case definition specifies clinical criteria and details of time, place, and person.

case-fatality rate (also called **case-fatality ratio)** the proportion of persons with a particular condition (e.g., patients) who die from that condition. The denominator is the number of persons with the condition; the numerator is the number of cause-specific deaths among those persons.

case, index the first case or instance of a patient coming to the attention of health authorities.

case-patientin a case-control study, a person who has the disease, injury, or other health condition that meets the case definition (see also **case**).

case, source the case or instance of a patient responsible for transmitting infection to others; the instance of a patient who gives rise to an outbreak or epidemic.

cause, component a factor that contributes to a sufficient cause (see **cause, sufficient**).
cause of disease a factor (e.g., characteristic, behavior, or event) that directly influences the occurrence of a disease. Reducing such a factor among a population should reduce occurrence of the disease.
cause, necessary a factor that must be present for a disease or other health problem to occur.
cause-specific mortality rate see **mortality rate, cause-specific**.
cause, sufficient a factor or collection of factors whose presence is always followed by the occurrence of a particular health problem.
census the enumeration of an entire population, usually including details on residence, age, sex, occupation, racial/ethnic group, marital status, birth history, and relationship to the head of household.
central location (also called **central tendency)** a statistical measurement to quantify the middle or the center of a distribution. Of the multiple ways to define central tendency, the most common are the mean, median, and mode.
chain of infection the progression of an infectious agent that leaves its reservoir or host through a portal of exit, is conveyed by a mode of transmission, and then enters through an appropriate portal of entry to infect a susceptible host.
"**chartjunk**" unnecessary or confusing visual elements in charts, illustrations, or graphs. The term was first used by Edward Tufte in his book, *The Visual Display of Quantitative Information* (1983).
class interval the span of values of a continuous variable that are grouped into a single category (see **class**), usually to create a frequency distribution for that variable.
class limits the values at the upper and lower ends of a class interval.
clinical criteria the medical features (e.g., symptoms, medical examination findings, and laboratory results) that are used in a case definition.
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, injury, or other health condition (particularly cancer and birth defects) in a circumscribed area during a particular period without regard to whether the number of cases is more than expected (often the expected number is not known).
cohorta well-defined group of persons 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 events.
cohort, birth a group of persons born during a particular period or year.
cohort study see **study, cohort**.
common-source outbreak see **outbreak, common-source**.
community immunity see **immunity, herd**.
community trial see **trial, community**.
comparison group a group in an analytic study (e.g., a cohort or case-control study) with whom the primary group of interest (exposed group in a cohort

study or case-patients in a case-control study) is compared. The comparison group provides an estimate of the background or expected incidence of disease (in a cohort study) or exposure (in a case-control study).
confidence interval a range of values for a measure (e.g., rate or odds ratio) constructed so that the range has a specified probability (often, but not necessarily, 95%) of including the true value of the measure.
confidence limits the end points (i.e., the minimum and maximum values) of a confidence.
confoundingthe distortion of the association between an exposure and a health outcome by a third variable that is related to both.
contactexposure to a source of an infection; a person who has been exposed.
contact, direct exposure or transmission of an agent from a source to a susceptible host through touching (e.g., from a human host by kissing, sexual intercourse, or skin-to-skin contact) or from touching an infected animal or contaminated soil or vegetation.
contagiouscapable of being transmitted from one person to another by contact or close
proximity.
contingency table a two-variable table of cross-tabulated data.
continuous variable see **variable, continuous**.
controlin a case-control study, a member of the group of persons without the health problem under study (see also **comparison group** and **study, case-control**).
crudewhen referring to a rate, an overall or summary rate for a population, without adjustment.
crude birth rate see **birth rate, crude**.
crude death rate see **mortality rate, crude**.
crude mortality rate see **mortality rate, crude**.
cumulative frequency in a frequency distribution, the number or proportion of observations with a particular value and any smaller value.
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.

Death-to-case ratio the number of deaths attributed to a particular disease, injury, or other healthcondition during a specified period, divided by the number of new cases of that disease, injury,or condition identified during the same period.
decision analysis application of quantitative methods to decision-making.
decision tree a branching chart that represents the logical sequence or pathway of a clinical or public health decision.
demographic information personal characteristics of a person or group (e.g., age, sex,
race/ethnicity, residence, and occupation) demographic information is used in descriptive epidemiology to characterize patients or populations.
dendrogramsee **phylogenetic tree**.

denominatorthe lower portion of a fraction; used in calculating a ratio, proportion, or rate. For a rate, the denominator is usually the midinterval population.
dependent variable see **variable, dependent**.
descriptive epidemiology see **epidemiology, descriptive**.
determinantany factor that brings about change in a health condition or in other defined
characteristics (see also **cause** and **risk factor**).
direct transmission see **transmission, direct**.
discrete variable (or **data**) see **variable (**or **data), discrete**.
distributionin epidemiology, the frequency and pattern of health-related characteristics and events in a population. In statistics, the frequency and pattern of the values or categories of a variable.
dose-responseassociation between an exposure and health outcome that varies in a consistently increasing or decreasing fashion as the amount of exposure (dose) increases.
dot plot a visual display of the specific data points of a variable.
droplet nuclei the residue of dried droplets of infectious agents that is easily inhaled and exhaled and can remain suspended in air for relatively long periods or 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 that travel only a short distance before falling to the ground.

Effect the result of a cause.
effectivenessthe ability of an intervention or program to produce the intended or expected results in the field.
efficacythe ability of an intervention or program to produce the intended or expected results under ideal conditions.
efficiencythe ability of an intervention or program to produce the intended or expected results with a minimum expenditure of time and resources.
EIS Epidemic Intelligence Service; CDC's 2-year training program in applied epidemiology for public health professionals (http://www.cdc.gov/eis).
endemicthe constant presence of an agent or health condition within a given geographic area or population; can also refer to the usual prevalence of an agent or condition.
environmental factor an extrinsic factor (e.g., geology, climate, insects, sanitation, or health services) that affects an agent and the opportunity for exposure.
epidemicthe occurrence of more cases of disease, injury, or other health condition than expected in a given area or among a specific group of persons during a particular period. Usually, the cases are presumed to have a common cause or to be related to one another in some way (see also **outbreak**).
epidemic curve a histogram that displays the course of an outbreak or epidemic by plotting the number of cases according to time of onset.
epidemic period the time span of an outbreak or epidemic.

epidemiologic triad the traditional model of infectious disease causation having three components: an external agent, a susceptible host, and an environment that brings the host and agent together so that disease occurs.
epidemiologythe study of the distribution and determinants of health conditions or events among populations and the application of that 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 or expected values so that associations between exposures and outcomes can be quantified and hypotheses about the cause of the problem can be tested (see also **study, analytic)**.
epidemiology, applied the application or practice of epidemiology to control and prevent health problems.
epidemiology, descriptive the aspect of epidemiology concerned with organizing and summarizing data regarding the persons affected (e.g., the characteristics of those who became ill), time (e.g., when they become ill), and place (e.g., where they might have been exposed to the cause of illness).
epidemiology, field applied epidemiology (i.e., the application or practice of epidemiology to control and prevent health problems), particularly when the epidemiologist(s) must travel to and work in the community in which the health problem is occurring or has occurred.
evaluationsystematic and objective examination of activities to determine their relevance,
effectiveness, and impact.
excess risk risk difference, calculated as the risk among the exposed group minus the risk among the unexposed group.
experimental study see **study, experimental**.
exposed group a group whose members have had contact with a suspected cause of, or possess a characteristic that is a suspected determinant of, a particular health problem.
exposurehaving come into contact with a cause of, or possessing a characteristic that is a
determinant of, a particular health problem.

False-negative a negative test result for a person who actually has the condition similarly, aperson who has the disease (perhaps mild or variant) but who does not fit the case definition, or apatient or outbreak not detected by a surveillance system.
false-positivea positive test result for a person who actually does not have the condition.
Similarly, a person who does not have the disease but who nonetheless fits the case definition, or a patient or outbreak erroneously identified by a surveillance system.
field epidemiology see **epidemiology, field**.
follow-up study see **study, cohort**.

fomitean inanimate object that can be the vehicle for transmission of an infectious agent (e.g., bedding, towels, or surgical instruments).
forest plot a graph that displays the point estimates and confidence intervals of individual studies included in a meta-analysis or systematic review as a series of parallel lines.
frequencythe amount or number of occurrences of an attribute or health outcome among a population.
frequency distribution a complete summary of the frequencies of the values or categories of a variable, 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 straight lines.

Geometric mean see **mean, geometric**.
grapha visual display of quantitative data arranged on a system of coordinates.

Health a state of complete physical, mental, and social well-being and not merely the absence ofdisease or other infirmity.
health indicator any of a variety of measures (e.g., mortality rate) that indicate the state of health of a given population.
health information system a combination of health statistics from different sources. Data from these systems are used to learn about health status, health care, provision and use of services, and the impact of services and programs on health.
healthy worker effect the observation that employed persons generally have lower mortality rates than the general population, because persons with severe, disabling disease (who have higher mortality rates) tend to be excluded from the workforce.
herd immunity see **immunity, herd**.
high-risk group a group of persons whose risk for a particular disease, injury, or other health condition is greater than that of the rest of their community or population.
HIPAA the Health Insurance Portability and Accountability Act, enacted in 1996, which addresses the privacy of a person's medical information as well as postemployment insurance and other health-related concerns.
histograma 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 grouped on the vertical axis. Columns are drawn so that their bases equal the class intervals (i.e., so that columns of adjacent intervals touch), and their heights correspond to the class frequencies.
hosta person or other living organism that is susceptible to or harbors an infectious agent under natural conditions.

host factor an intrinsic factor (e.g., age, race/ethnicity, sex, or behaviors) that influences a person's exposure, susceptibility, or response to an agent.
hyperendemicthe constant presence at high incidence and prevalence of an agent or health condition within a given geographic area or population.
hypothesisa 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 (see **hypothesis, null**) proves implausible.
hypothesis, null the supposition that two (or more) groups do not differ in the measure of interest (e.g., incidence or proportion exposed); the supposition that an exposure is not associated with the health condition under study, so that the risk ratio or odds ratio equals 1. The null hypothesis is used in conjunction with statistical testing.

Immunity, active resistance developed in response to an antigen (i.e., an infecting agent orvaccine), usually characterized by the presence of antibody produced by the host.
immunity, herd the resistance to an infectious agent of an entire group or community (and, in particular, protection of susceptible persons) as a result of a substantial proportion of the population being immune to the agent. Herd immunity is based on having a substantial number of immune persons, thereby reducing the likelihood that an infected person will come in contact with a susceptible one among human populations, also called **community immunity**.
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 (e.g., antiserum or immune globulin).
incidencea measure of the frequency with which new cases of illness, injury, or other health condition occurs among a population during a specified period.
incidence proportion the fraction of persons with new cases of illness, injury, or other health condition during a specified period, calculated as the number of new cases divided by the size of the population at the start of the study period (see also **attack rate**).
incidence rate a measure of the frequency with which new cases of illness, injury, or other health condition occur, expressed explicitly per a time frame. Incidence rate is calculated as the number of new cases over a specified period divided either by the average population (usually mid-period) or by the cumulative person-time the population was at risk.
incubation period the time interval from exposure to an infectious agent to the onset of symptoms of an infectious disease.
independent variable see **variable, independent**.
index case see **case, index**.
indirect transmission see **transmission, indirect**.

individual data values or observations from each record (also called raw data).
infant mortality rate see **mortality rate, infant**.
infectioninvasion of the body tissues of a host by an infectious agent, whether or not it causes disease.
infectivitythe ability of an infectious agent to cause infection, measured as the proportion of persons exposed to an infectious agent who become infected.
information bias see bias, information.
interquartile range a measure of spread representing the middle 50% of the observations, calculated as the difference between the third quartile (75th percentile) and the first quartile (25th percentile).
isolationthe separation of infected persons to prevent transmission to susceptible ones. Isolation refers to separation of ill persons; **quarantine** refers to separation of potentially exposed but well persons.

Latency period the time from exposure to a causal agent to onset of symptoms of a (usuallynoninfectious) disease (see also **incubation period**).
life expectancy a statistical projection of the average number of years a person of a given age is expected to live, if current mortality rates continue to apply.
line graph, arithmetic-scale a graph that displays patterns or trends by plotting the frequency (e.g., number, proportion, or rate) of a characteristic or event during some variable, usually time.
The y-axis, measuring frequency, uses an arithmetic scale.
line graph, semilogarithmic-scale a graph that displays patterns or trends by plotting the frequency (e.g., number, proportion, or rate) of a characteristic or event during some variable, usually time. The y-axis, measuring frequency, uses a logarithmic scale.
line listing a type of epidemiologic database, organized similar to a spreadsheet with rows and columns in which information from cases or patients are listed each column represents a variable, and each row represents an individual case or patient.
logarithmic transformation conversion of nominal or ordinal data to logarithmic data. The purpose is to examine rate of change instead of amount of change only.

Map, area (shaded, choropleth) a visual display of the geographic pattern of a health problem,in which a marker is placed on a map to indicate where each affected person lives, works, ormight have been exposed.
mean (or average) commonly called the average; it is the most common measure of central tendency.
mean, arithmetic the measure of central location, commonly called the average, 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 (e.g., risk ratio, 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 see **measure of spread**.
measure of spread a measure of the distribution of observations out from its central value. Measures of spread used in epidemiology include the interquartile range, variance, and the standard deviation.
measurement scale the complete range of possible values for a measurement.
mechanical transmission see **transmission, mechanical**.
medianthe measure of central location that divides a set of data into two equal parts, above and below which lie an equal number of values (see also **measure of central location**).
medical surveillance see **surveillance, medical**.
midrangethe halfway point, or midpoint, in a set of observations. For the majority 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.
modethe most frequently occurring value in a set of observations (see also **measure of central location**).
mode of transmission the manner in which an agent is transmitted from its reservoir to a susceptible host (see also **transmission**).
morbiditydisease; any departure, subjective or objective, from a state of physiological or psychological health and well-being.
mortalitydeath.
mortality rate a measure of the frequency of occurrence of death among a defined population during a specified time interval.
mortality rate, age-adjusted a mortality rate that has been statistically modified to eliminate the effect of different age distributions among different populations.
mortality rate, age-specific a mortality rate limited to a particular age group, calculated as the number of deaths among the age group divided by the number of persons in that age group, usually expressed per 100,000.
mortality rate, cause-specific the mortality rate from a specified cause, calculated as the number of deaths attributed to a specific cause during a specified time interval among a population divided by the size of the midinterval population.
mortality rate, crude a mortality rate from all causes of death for an entire population, without adjustment.
mortality rate, infant the mortality rate for children aged <1 year, calculated as the number of deaths reported among this age group during a given period divided by the number of live births reported during the same period, and expressed per 1,000 live births. Infant mortality rate is a universally accepted

indicator of the health of a nation's population and the adequacy of its health-care system.
mortality rate, neonatal the mortality rate for children from age birth up to, but not including, 28 days. In calculating neonatal mortality rates, the numerator is the number of deaths among this age group during a given period, and the denominator is the number of live births reported during the same period. The neonatal mortality rate is usually expressed per 1,000 live births.
mortality rate, postneonatalthe mortality rate for children from age 28 days up to, but not including, 1 year. In calculating postneonatal mortality rates, the numerator is the number of deaths among this age group during a given 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/ethnic-specific a mortality rate limited to a specified racial or ethnic group both numerator and denominator are limited to that group.
mortality rate, sex-specific a mortality rate among either males or females.

Natural history of disease the progression of a disease process in a person from the time itbegins to the time it resolves, in the absence of treatment.
NCHS The National Center for Health Statistics, the US governmental organization responsible for national vital statistics and multiple national health surveys. Organizationally, NCHS is a component of the Centers for Disease Control and Prevention, one of the agencies of the US Department of Health and Human Services.
NHANES The National Health and Nutrition Examination Survey, a representative survey of the civilian, noninstitutionalized US population conducted by the National Center for Health
Statistics, designed to (1) estimate the proportion of the US population and designated groups with selected disease and risk factors; (2) monitor trends in selected behaviors, exposures, and diseases; and (3) study the associations among diet, nutrition, and health.
necessary cause see **cause, necessary**.
neonatal mortality rate see **mortality rate, neonatal**.
nominal scale see **scale, nominal**.
normal curve the bell-shaped curve that results when a normal distribution is graphed.
normal distribution a distribution represented as a bell shape, symmetrical on both sides of the peak, which is simultaneously the mean, median, and mode, and with both tails extending to infinity.
notifiable disease a disease that, by law, must be reported to public health authorities upon diagnosis.
null hypothesis see **hypothesis, null**.
numeratorthe upper portion of a fraction (see also **denominator**).

Observational study see **study, observational**.
odds ratio a measure of association used in comparative studies, particularly case-control studies, that quantifies the association between an exposure and a health outcome; also called the cross-product ratio.
ordinal scale see **scale, ordinal**.
outbreakthe occurrence of more cases of disease, injury, or other health condition than expected in a given area or among a specific group of persons during a specific period. Usually, the cases are presumed to have a common cause or to be related to one another in some way. Sometimes distinguished from an epidemic as more localized, or the term less likely to evoke public panic (see also **epidemic**).
outbreak, common-source an outbreak that results from persons being exposed to the same harmful influence (e.g., an infectious agent or toxin). The exposure period can be brief or can extend over 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 spreads from person to person rather than from a common source.
outcome(s)any or all of the possible results that can 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.
outliera value substantively or statistically different from all (or approximately all) of the other values in a distribution.

Pandemic an epidemic occurring over a widespread area (multiple countries or continents) and usually affecting a substantial proportion of the population.
***P* value** the probability of observing an association between two variables or a difference
between two or more groups as large or larger than that observed, if the null hypothesis were true. Used in statistical testing to evaluate the plausibility of the null hypothesis (i.e., whether the observed association or difference plausibly might have occurred by chance).
passive immunity see **immunity, passive**.
passive surveillance see**surveillance, passive**.
pathogenicitythe ability of an agent to cause disease after infection, measured as the proportion of persons infected by an agent who then experience clinical disease.
percentilea 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 or 1% of the observations. Forexample, 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 rate the incidence rate calculated as the number of new cases among a population divided by the cumulative person-time of that population, usually expressed as the number of events per persons per unit of time.

person-timethe amount of time each participant in a cohort study is observed and disease-free, often summed to provide the denominator for a person-time rate.
phylogenetic tree a branching chart that indicates the evolutionary lineage or genetic relatedness of organisms.
pie chart a circular graph of a frequency distribution in which each segment of the pie is proportional in size to the frequency of corresponding category.
pointprevalencesee**prevalence, point**.
point-source outbreak see **outbreak, point-source**.
populationthe total number of inhabitants of a geographic area or the total number of persons in a particular group (e.g., the number of persons engaged in a certain occupation).
population pyramid a graphical display of the age-sex distribution of a population, constructed with a horizontal histogram of the age distribution of males pointing to the left, and the corresponding horizontal histogram of age distribution of females pointing to the right.
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 host.
postneonatal mortality rate see **mortality rate, postneonatal**.
predictive value positive the proportion of cases identified by a test, reported by a surveillance system, or classified by a case definition that are true cases, calculated as the number of true positives divided by the number of true-positives plus false-positives.
prevalencethe number or proportion of cases or events or attributes among a given population.
prevalence rate the proportion of a population that has a particular disease, injury, other health condition, or attribute at a specified point in time (point prevalence) or during a specified period (period prevalence).
prevalence, period the amount of a particular disease, chronic condition, or type of injury present among a population at any time during a particular period.
prevalence, point the amount of a particular disease, chronic condition, or type of injury present among a population at a single point in time.
privacy rule a set of regulations based on the Health Insurance Portability and Accountability Act to protect the privacy of individually identifiable health information.
propagated outbreak see **outbreak, propagated**.
proportiona 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 health state or event among exposed persons that can be attributed to the exposure also called attributable risk percent.
proportionate mortality the proportion of deaths among a population attributable to a particular cause during a selected period. Each cause of death

is expressed as a percentage of all deaths, and the sum of the proportionate mortality for all causes must equal 100%. These proportions are not mortality rates because, in proportionate mortality, the denominator is all deaths instead of the population among whom the deaths occurred.
prospective study see **study, prospective**.

Quarantine the separation of well persons who have been exposed or are suspected to have beenexposed to a communicable disease, to monitor for illness and to prevent potential transmissionof infection to susceptible persons during the incubation period. Quarantine refers to separationof potentially exposed but well persons; **isolation** refers to separation of ill persons.

Race/ethnic-specific mortality rate see **mortality rate, race/ethnic-specific**.
random sample see **sample, random**.
rangein statistics, the difference between the largest and smallest values in a distribution; in common use, the span of values from smallest to largest.
ratean expression of the relative frequency with which an event occurs among a defined population per unit of time, calculated as the number of new cases or deaths during a specified period divided by either person-time or the average (midinterval) population. In epidemiology, it is often used more casually to refer to proportions that are not truly rates (e.g., attack rate or case-fatality rate).
rate ratio a measure of association that quantifies the relation between an exposure and a health outcome from an epidemiologic study, calculated as the ratio of incidence rates or mortality rates of two groups.
ratiothe relative size of two quantities, calculated by dividing one quantity by the other.
recordin a line listing, each row is a record or observation. A record represents data related to a single case.
relative risk a general term for measures of association calculated from the data in a two-by-two table, including risk ratio, rate ratio, and odds ratio (see **risk ratio**).
representative sample see **sample, representative**.
reservoirthe habitat in which an infectious agent normally lives, grows, and multiplies, which can include humans, animals, or the environment.
retrospective study see **study, retrospective**.
riskthe probability that an event will occur (e.g., that a person will be affected by, or die from, an illness, injury, or other health condition within a specified 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, injury, or other health condition.
risk ratio a measure of association that quantifies the association between an exposure and a health outcome from an epidemiologic study, calculated as the ratio of incidence proportions of two groups.

Sample a selected subset of a population a sample can be random or nonrandom andrepresentative or nonrepresentative.
sample, random a sample of persons 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.
scale, interval a measurement scale consisting of quantitative categories whose values are measured on a scale of equally spaced units, but without a true zero point (e.g., date of birth).
scale, nominal a measurement scale consisting of qualitative categories whose values have no inherent statistical order or rank (e.g., categories of race/ethnicity, religion, or country of birth).
scale, ordinal a measurement scale consisting of qualitative categories whose values have a distinct order but no numerical distance between their possible values (e.g., stage of cancer, I, II, III, or IV).
scale, ratio a measurement scale consisting of quantitative categories whose values are intervals with a true zero point (e.g., height in centimeters or duration of illness).
scatter diagram (or **scattergram**) a graphical display of the association between two variables in which 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.
seasonalitychange in physiologic 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**.
selection bias see **bias, selection**.
semilogarithmic-scale line graph see **line graph, semilogarithmic-scale**
sensitivitythe ability of a test, case definition, or surveillance system to identify true cases; the proportion of people with a health condition (or the proportion of outbreaks) that are identified by a screening test or case definition (or surveillance system).
sentinel surveillance see**surveillance, sentinel**.
sex-specific mortality rate see **mortality rate, sex-specific**.
skeweda distribution that is not symmetrical.
source (of infection) the person, animal, object, or substance from which an infectious agent is transmitted to a host.
source case see **case, source**.
specificity the ability or a test, case definition, or surveillance system to exclude persons without the health condition of interest; the proportion of persons without a health condition that are correctly identified as such by a screening test, case definition, or surveillance system.
spectrum of illness the range of manifestations a disease process can take (e.g., from asymptomatic to mild clinical illness to severe illness and death).
sporadican event that occurs infrequently and irregularly.

spot map a visual display of the geographic pattern of a health problem, in which a marker is placed on a map to indicate where each affected person lives, works, or might have been exposed.
standard deviation a statistical summary of how dispersed the values of a variable are around its mean, calculated as the 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 (see also ***P* value**).
study, analytic a study, usually observational, 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 observational analytic study that enrolls one group of persons with a certain disease, chronic condition, or type of injury (case-patients) and a group of persons without the health problem (control subjects) and compares differences in exposures, behaviors, and other characteristics to identify and quantify associations, test hypotheses, and identify causes.
study, cohort 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 documented and compared. Cohort studies can be either prospective or retrospective.
study, cross-sectional a study in which a sample of persons from a population are enrolled and their exposures and health outcomes are measured simultaneously; a survey.
study, experimental a study in which the investigator specifies the type of exposure for each person (clinical trial) or community (community trial) then follows the persons' or communities' health status to determine the effects of the exposure.
study, observational a study in which the investigator observes rather than influences exposure and disease among participants. Case-control and cohort studies are observational studies (see also **study, experimental**).
study, prospective an analytic study in which participants are enrolled before the health outcome of interest has occurred.
study, retrospective an analytic study in which participants are enrolled after the health outcome of interest has occurred. Case-control studies are inherently retrospective.
subclinicalwithout apparent symptoms.

surveillance, active public health surveillance in which the health agency solicits reports.
surveillance, medical monitoring of a person who might have been exposed to an infectious, chemical, radiologic, or other potentially causal agent, for the purpose of detecting early symptoms.
surveillance, passive public health surveillance in which data are sent to the health agency without prompting.
surveillance, sentinel a surveillance system that uses a prearranged sample of sources (e.g., physicians, hospitals, or clinics) who have agreed to report all cases of one or more notifiable diseases.
surveillance, syndromic(1) the monitoring of the frequency of illnesses with a specified set of clinical features among a given population without regard to the specific diagnoses, if any, that are assigned to them by clinicians. (2) A system for early detection of outbreaks whereby health department staff, assisted by automated acquisition of data routinely collected for other purposes and computer generation of statistical signals, monitor disease indicators, particularly those associated with possible terrorism-related biologic and chemical agents, continually or at least daily to detect outbreaks earlier than would otherwise be possible with traditional public health methods.
surveya systematic canvassing of persons to collect information, often from a representative sample of the population.
survival curve a line graph that begins with 100% of the study population and displays the percentage of the population still surviving at successive points in time. A survival curve can also be used to depict freedom from a health problem, complication, or another endpoint.
symmetricala type of distribution where the shapes to the right and left of the central location are the same. Normal, bell-shaped distributions are symmetrical; the mean, median, and mode are the same.
symptomany indication of disease noticed or felt by a patient.
syndromea combination of symptoms characteristic of a disease or health condition; sometimes refers to a health condition without a clear cause (e.g., chronic fatigue syndrome).
syndromic surveillance see**surveillance, syndromic**.

Table an arrangement of data in rows and columns. In epidemiology, the data are usually
summaries of the frequency of occurrence of an event or characteristic occurring among different groups.
trendmovement or change in frequency over time, usually upwards or downwards**.**
trend, secular changes occurring over a substantial period, generally years or decades.
trial, clinical an experimental study that uses data from individual persons. The investigator specifies the type of exposure for each study participant and then follows each person's health status to determine the effects of the exposure.

trial, community an experimental study that uses data from communities. The investigator specifies the type of exposure for each community and then follows the communities' health status to determine the effects of the exposure.
trial, randomized clinical a clinical trial in which persons are randomly assigned to exposure or treatment groups.
two-by-two table see **table, two-by-two.**

Validity the degree to which a measurement, questionnaire, test, or study or any other datacollectiontool measures what it is intended to measure.
variableany characteristic or attribute that can be measured and can have different values.
variable(or **data**), **discrete** a variable that is limited to a finite number of values; data for such a variable.
variable, continuous a variable that has the potential for having an infinite number of values along a continuum (e.g., height and weight).
variable, dependent in a statistical analysis, a variable whose values are a function of one or more other variables.
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 spread in a set of observations, calculated as the sum of the squares of deviations from the mean, divided by the number of observations minus 1 (see also **standarddeviation**).
vectora living intermediary that carries an agent from a reservoir to a susceptible host (see also **transmission, biologic** and **transmission, mechanical**) (e.g., mosquitoes, fleas, or ticks).
vehiclean inanimate object that can carry an agent from a reservoir to a susceptible host (e.g., food, water, blood products, and bedding) (see also **transmission, indirect**).
virulencethe ability of an infectious agent to cause severe disease, measured as the proportion of persons with the disease who become severely ill or die.
vital statistics systematically tabulated data about recorded births, marriages, divorces, and deaths.

X-axis the horizontal axis of a rectangular graph, usually displaying the independent variable(e.g., time).
Y-axis the vertical axis of a rectangular graph, usually displaying the dependent variable (e.g.,frequency — number, proportion, or rate).
years of potential life lost (YPLL) a measure of the impact of premature death on a population,
calculated as the sum of the differences between a predetermined minimally acceptable age (e.g., 65 years or current life expectancy) and the age at death for everyone who died earlier than that age.

Zoonosis an infectious disease that is transmissible from animals to humans.

References

1. WHO Definition of Health Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, 1946
2. http://www.globalhandwashing.org/Publications/Lit_review.htm
3. *The Solid Facts: Social Determinants of Health* edited by Richard Wilkinson and Michael Marmot, WHO, 2003
4. SYED M. AHMED, BARBRA BECK, CHERYL A. MAURANA & GAIL NEWTON Overcoming Barriers to Effective Community-Based Participatory Research in US Medical Schools. Education for Health, Vol. 17, No. 2, July 2004, 141 – 151.
5. Asia-Pacific Academic Consortium for Public Health. "Asia-Pacific Journal of Public Health". http://aph.sagepub.com/. Retrieved on 12 January 2009.
6. *Annual Reviews.* "Annual Reviews of Public Health". http://www.annualreviews.org/default.aspx. Retrieved on 29 December 2008.
7. Public Health Association of Australia. "Australian and New Zealand Journal of Public Health". http://blackwellpublishing.com/journal.asp?ref=1326-0200&site=1. Retrieved on 12 January 2009.