EXCITE

- About Excite
- Epidemiology in the Classroom
- Careers in Public Health
- Science Ambassador
- Science Olympiad
- Resource Library

Epidemiology in the Classroom

An Introduction to Epidemiology

This document is also available in Portable Document Format (891 Kb). You will need Acrobat Reader to view and print these documents.

This background piece provides several working definitions of epidemiology—the basic science of public health; an introduction to the different categories of epidemiology and types of epidemiological studies; and an overview of the disease transmission cycle. First, to set the stage, consider the three incidents that follow, stepping into the shoes of the public health officer who received the initial report and asking yourself the question, "What do I do now?" Some of these examples made national news and may be familiar to you.

First Incident

In March 1985, a nurse epidemiologist in a county health department noted, while reviewing surveillance data, three cases in a single month of hepatitis B of unusual origin. Hepatitis B, or serum hepatitis, is transmitted through sexual contact and by exposure to infected bodily fluids, but these three patients did not seem to have the usual risk factors. All three people did, however, indicate having received injections at the same health care facility.

The nurse's immediate questions were: Is this a coincidence? Did these three cases occur by chance or is there a link? In this instance, the nurse decided to pursue an investigation.

Back to Top

Second Incident

At 8:30 in the morning on August 2, 1976, Dr. Robert B. Craven of CDC's Viral Diseases Division received a call from a nurse at a Veterans' Hospital in Philadelphia, Pennsylvania. The nurse reported two cases of severe respiratory illness, one of which had been fatal. Both people had attended the annual American Legion Convention held July 21-24 . By the evening of August 2, 71 more of the people attending the convention had the same illness, with symptoms of acute onset of fever, chills, headache, malaise, dry cough, and myalgia. Further conversations with local and state public health officials revealed that between July 26 and August 2, 18 conventioneers had died. Deaths were due primarily to pneumonia.

An intense investigation began immediately. The incident became known as the first outbreak of Legionnaires' disease and led to the discovery of the gram-negative pathogen, Legionella pneumophila.

Back to Top

Third Incident

On October 30, 1989, a New Mexico physician notified the state's health department of three patients with marked peripheral eosinophilia and severe myalgia. All three patients had been taking oral preparations of L-tryptophan, a nonprescription drug sold as a dietary supplement in health food stores. Despite extensive clinical evaluation and testing, the illness could not be identified.

An investigation followed and resulted in the characterization of eosinophilia-myalgia syndrome, EMS. The investigation implicated a vehicle for exposure—L-tryptophan dietary supplements—before a suspected agent was identified, and the product was taken off the market. Eventually, the problem was traced to a contaminant that had been introduced by changes in the production process at a single manufacturing facility.

These three examples illustrate some of the key reasons for needing applied, or field, epidemiology:

- They were unexpected.
- They demanded a response.
- The investigators had to go out into the field to solve the problem.

Back to Top

Epidemiology Defined

An apocryphal story is told around CDC that illustrates the confusion sometimes accompanying the term "epidemiology." It seems that one of our scientists, on first arriving at CDC from a clinical practice, found himself somewhat unsure of what epidemiology was all about, so he sought an answer down the street at Emory University. The first person he asked was a medical student, who told him that epidemiology was "the worst taught course in medical school." The second, a clinical faculty member, told him epidemiology was "the science of making the obvious obscure." Finally, knowing that statistics are important to epidemiology, he asked a statistician, who told him that epidemiology is "the science of long division" and provided him with a summary equation. Giving up on finding a real answer, he returned to CDC. On the way, however, he decided to try one more time. He stopped a native Atlantan who told him that epidemiology was "the study of skin diseases."

A less entertaining, but more conventional, definition of epidemiology is "the study of the distribution and determinants of health-related states in specified populations, and the application of this study to control health problems." A look at the key words will help illuminate the meaning:

- **Study**—Epidemiology is the basic science of public health. It's a highly quantitative discipline based on principles of statistics and research methodologies.

- **Distribution**—Epidemiologists study the distribution of frequencies and patterns of health events within groups in a population. To do this, they use descriptive epidemiology, which characterizes health events in terms of time, place, and person.

- **Determinants**—Epidemiologists also attempt to search for causes or factors that are associated with increased risk or probability of disease. This type of epidemiology, where we move from questions of "who," "what," "where," and "when" and start trying to answer "how" and "why," is referred to as analytical epidemiology.

- **Health-related states**—Although infectious diseases were clearly the focus of much of the early epidemiological work, this is no longer true. Epidemiology as it is practiced today is applied to the whole spectrum of health-related events, which includes chronic disease, environmental problems, behavioral problems, and injuries in addition to infectious disease.

- **Populations**—One of the most important distinguishing characteristics of epidemiology is that it deals with groups of people rather than with individual patients.
- **Control**—Finally, although epidemiology can be used simply as an analytical tool for studying diseases and their determinants, it serves a more active role. Epidemiological data steers public health decision making and aids in developing and evaluating interventions to control and prevent health problems. This is the primary function of applied, or field, epidemiology.

Back to Top

A comparison between the practice of public health and the more familiar practice of health care helps in describing epidemiology. First, where health care practitioners collect data on an individual patient by taking a medical history and conducting a physical exam, epidemiologists collect data about an entire population through surveillance systems or descriptive epidemiological studies. The health care practitioner uses his or her data to make a differential diagnosis. The epidemiologist's data is used to generate hypotheses about the relationships between exposure and disease. Both disciplines then test the hypotheses, the health care practitioner by conducting additional diagnostic studies or tests, the epidemiologist by conducting analytical studies such as cohort or case-control studies. The final step is to take action. The health care practitioner prescribes medical treatment, and the epidemiologist, some form of community intervention to end the health problem and prevent its recurrence.

One succinct way to sum up the task of epidemiologists is to say that they "count things." Basically, epidemiologists count cases of disease or injury, define the affected population, and then compute rates of disease or injury in that population. Then they compare these rates with those found in other populations and make inferences regarding the patterns of disease to determine whether a problem exists,. For example, in the hepatitis B example earlier, you might ask: Is the rate of disease among people with no know risk factors greater than we would expect? Is the pattern or distribution of the cases suspicious? Once a problem has been identified, the data are used to determine the cause of the health problem; the modes of transmission; any factors that are related to susceptibility, exposure, or risk; and any potential environmental determinants.

Back to Top

Epidemiological Studies

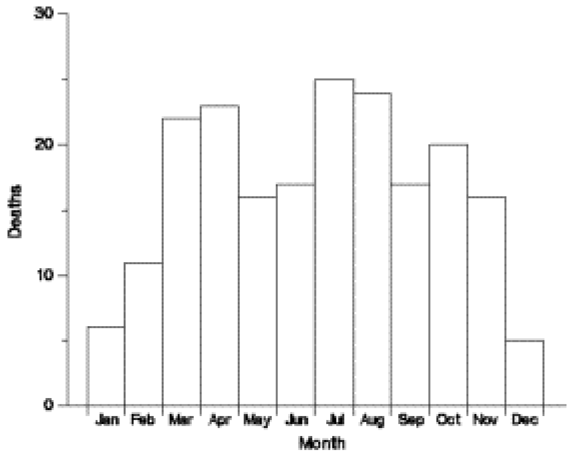
As mentioned earlier, epidemiologists used several different types of studies. Simply speaking, these can be classified as either **experimental**, where the epidemiologists have control over the circumstances from the start, or **observational**, where they do not. Vaccine efficacy trials are a good example of experimental studies because investigators control who gets the vaccine and who doesn't. Observational studies can be further subdivided into **descriptive** and **analytical** studies. In a descriptive study, the epidemiologist collects information to characterize and summarize the health event or problem. In an analytical study, the epidemiologist relies on comparisons between groups to determine the role of various risk factors in causing the problem. Descriptive epidemiology is the most basic of the these categories and is fundamental to the work of an epidemiologists.

Another way of comparing descriptive and analytical epidemiology is to say that in the descriptive process, we are concerned with "person" (Who was affected?), "place" (Where were they affected?), and time (When were they affected?). Once we know the answers to these questions, we can enter the realm of analytical epidemiology and ask how and why these people were affected.

A descriptive study of fatalities associated with the use of farm tractors illustrates the usefulness of person, place, and time for drawing inferences about health problems. The study was conducted in the early 1980s, using data from death certificate records, which are a readily available surveillance system. Take a moment to study the graphics below and consider what the data might mean, then click the "Inferences" bar for an explanation.

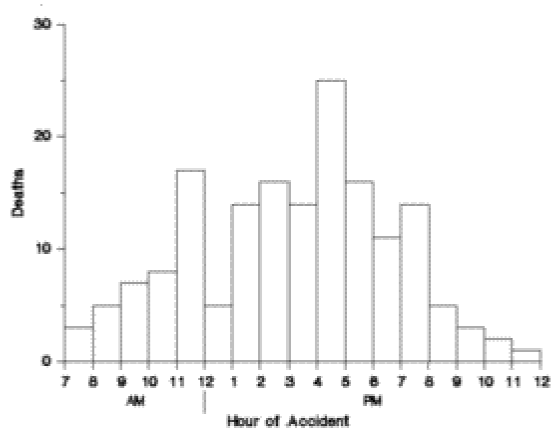
Back to Top

Deaths associated with tractor injuries, by month of death, Georgia 1971-1981



Deaths peaked in the spring and fall months, so they may be related to planting and harvest.

Deaths associated with tractor injuries, by time of day, Georgia 1971-1981



Because deaths peaked just before lunch and during the late afternoon and decrease from 12:00-1:00, they might be related to fatigue. The dip at noon could be because most people are taking a break for lunch. The fact that children are home from school by 4:00 could contribute to the peak in the afternoon.

Back to Top

Deaths associated with tractor injuries, by place, Georgia 1971-1981

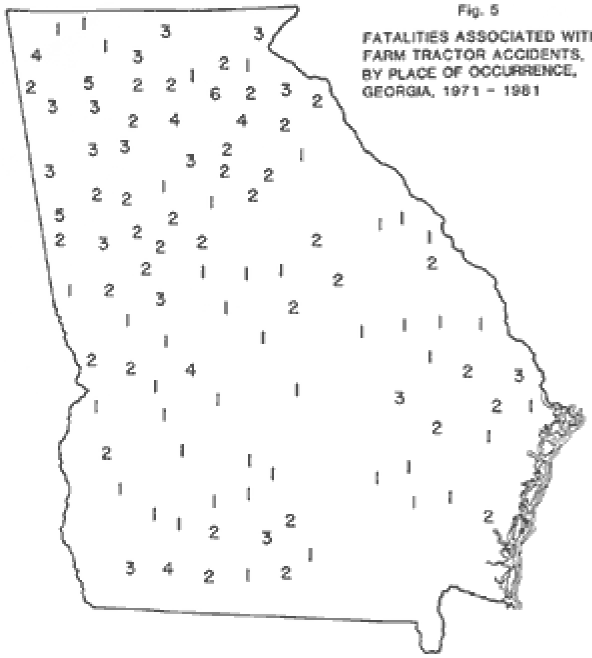


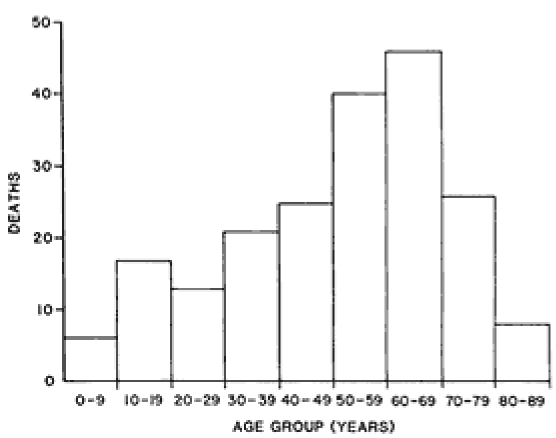
Fig. 5
FATALITIES ASSOCIATED WITH FARM TRACTOR ACCIDENTS, BY PLACE OF OCCURRENCE, GEORGIA, 1971 - 1981

Most of the deaths occurred in northern Georgia, which is hilly and mountainous; south-central Georgia, where fewer deaths occurred, is much flatter. This distribution might implicate the rugged terrain, but because the map shows numbers of deaths, not rates, we don't know whether the numbers could instead reflect differences in population size or even perhaps the number of tractors being used. As for the prevalence of tractors, south-central Georgia is more agrarian than northern Georgia, so the number of tractors probably isn't a factor. Another possible association is differences in experience and skill in using tractors.

[Back to Top](#)

Deaths associated with tractor injuries, by age group, Georgia 1971-1981

Fig. 1 FATALITIES ASSOCIATED WITH FARM TRACTOR ACCIDENTS, BY AGE GROUP, GEORGIA , 1971 - 1981



The number of deaths is clearly higher in the older age group. This could mean that tractor users are predominantly older, but it could also indicate that older users are less likely to survive an accident.

Any inferences you make are likely bases for hypotheses, which would then have to be tested using one of three analytical study designs: cross-sectional, cohort, and case-control. In all three types, the epidemiologist is attempting to discover the relationship between an exposure or risk factor and a health outcome. For example: Did the chicken salad at the company picnic cause the salmonella outbreak? Does cigarette smoke cause lung cancer? Are alcohol use and motor vehicle crashes related? Does the supplement L-tryptophan cause EMS?

The first type of design, a **cross-sectional study**, is basically the same as a survey. In this type of study, the epidemiologist defines the population to be studied and then collects information from members of the group about their disease and exposure status. Since the data represent a point in time, it's like taking a "snapshot" of the population. Cross-sectional studies are good for examining the relationship between a variable and a disease, but not for determining cause and effect, which requires data over time. Cohort studies and case-control studies are much better suited to examining cause-and-effect relationships.

In a **cohort study**, you select the study population according to their exposure, regardless of whether they have the disease or health outcome you're studying. You then determine the outcomes and compare them on the basis of the individuals' exposures. Cohort studies are often referred to as **prospective** studies because they follow the study population forward in time, from suspected cause to effect. An example would be dividing a group of people on the basis of their smoking status and following them for 20 years to see if they develop lung cancer. Cohort studies are also used to investigate outbreaks in small, well-defined populations. For example, you would use a cohort study to answer the question posed earlier regarding the cause of a salmonella outbreak at a company picnic. In this situation, you would ask each attendee about their exposure (e.g., what foods and beverages they consumed at the picnic) and whether they became ill afterward. The relationship between exposure and outcome in a cohort study is quantified by calculating the **relative risk** for the exposure.

Back to Top

Cohort studies have several advantages:

- You can examine multiple outcomes for a single exposure. For instance, if you select a group based on their smoking exposure, you can look not only at the incidence of lung cancer, but also at the incidence of cardiovascular disease, emphysema, burns, other smoking-related outcomes.
- Cohort studies are very useful in examining rare exposures, such as asbestos exposure and lung cancer.
- You can directly calculate the incidence of disease for each of the exposure groups.
- The temporal sequence is logical: you are starting with exposure and following forward in time to the development of disease.

Disadvantages of cohort studies are that they are costly in time and resources and, if the disease is rare, you need a large number of subjects. Also, because you are following forward in time, logistical problems may develop and subjects can be lost to follow-up. (Cohort studies are discussed in more detail in How to Investigate an Outbreak.)

In a **case-control study**, the epidemiologist is working backward, from the effect to the suspected cause. For this reason, case-control studies are often referred to as **retrospective** studies. Participants are selected on the basis of the presence or absence of the disease or outcome in question, so that you have one group of people (case-subjects) with the health problem and one without (controls). These groups are then compared to determine the presence of specific exposures or risk factors. For example, you could pick a group of people with lung cancer and a group without and then compare them for their history of exposure to smoking. The relationship between exposure and outcome in a case-control study is quantified by calculating the **odds ratio**.

Back to Top

Case-control studies have three primary advantages:

- You can examine multiple exposures for a single outcome.
- They are well suited for studying rare diseases and those with long latency periods.
- They require fewer case-subjects and are generally quicker and less expensive to conduct than cohort studies, which makes them well suited for the conditions of an outbreak investigation.

The disadvantages of case-control studies are that they aren't suitable for studying rare exposures;

they are subject to bias because of the method used to select controls; and they don't allow you to directly measure the incidence of disease. Also, because they look backward, case-control studies may create uncertainty about the temporal relationship between exposure and disease. (Case-control studies are discussed in more detail in How to Investigate an Outbreak.)

One important point must be made about analytical epidemiologic studies: In these studies, we are attempting to answer the how and why questions, and we are able to quantify the relationship between an exposure and an outcome. However, a mathematical, or quantitative, relationship between the two is not enough to establish causation. We can demonstrate increased relative risk of lung cancer for smokers in a cohort study, or we can demonstrate in a case-control study that people with lung cancer are much more likely to have smoked in the past; but that alone doesn't establish cause and effect.

In general, five criteria must be met to establish a cause-and-effect relationship:

- **Strength of association**—the relationship must be clear.
- **Consistency**—observation of the association must be repeatable in different populations at different times.
- **Temporality**—the cause must precede the effect.
- **Plausibility**—the explanation must make sense biologically.
- **Biological gradient**—there must be a dose-response relationship.

Back to Top

Disease Transmission

In talking about epidemiology, it is important to review how outbreaks occur. First, we'll look at three commonly used, and often misunderstood, terms: "epidemic," "outbreak," and "cluster." An **epidemic** is the occurrence of more cases of disease than would normally be expected in a specific place or group of people over a given period of time. To an epidemiologist, "**outbreak**" means basically the same thing. In the public's mind, however, "epidemic" has a far more serious connotation than "outbreak." For this reason, "outbreak" is often used to avoid sensationalism. The third term, "**cluster**," is occasionally used, incorrectly, in place of "epidemic" or "outbreak." A cluster is a group of cases in a specific time and place that may or may not be greater than the expected rate. Often the aim of investigating clusters is to determine the baseline rate of disease for that time and place. Two other terms you will come across are "**endemic**," meaning a high background rate of disease, and "**pandemic**," meaning very widespread, often global, disease.

For an outbreak, or epidemic, to occur, the basic elements of disease causation and an adequate chain of transmission must be present. Disease occurs when an outside **agent** capable of causing the disease meets a **host** that is vulnerable to the agent in an **environment** that allows the agent and host to interact. Then, given a chain of transmission from one host to another and a suitable mode of spread, an outbreak can develop. These basic concepts help guide the selection of public health strategies to prevent health problems. Depending on which approach might be most effective, we might direct efforts at the specific agent (e.g., guinea worm), host (e.g., immunization to prevent measles), or environment (e.g., sanitation improvements to prevent salmonella). We can also target a specific point in the chain of transmission. This was the response in the E.coli outbreak in Washington State in the early 1990s, when health officials called for the thorough cooking of hamburgers to interrupt transmission of the bacterium.

Back to Top

To elaborate further, the **agent** is the entity necessary to cause disease in a susceptible host. Classically, we think of an agent as being biological: a bacterium, parasite, or virus. But an agent can also be a physical force, as in the example of motor vehicle crashes; a chemical, as in an environmental problem; or a nutritional imbalance, as with rickets. When we talk about agents, several characteristics are important to consider:

- **Infectivity**—the capacity to cause infection in a susceptible host.
- **Pathogenicity**—the capacity to cause disease in a host.
- **Virulence**—the severity of disease that the agent causes in the host.

The **host** is the person, or in a more generic definition, the organism, that is susceptible to the effect of the agent. The status of the host is quite important and is generally classifiable as susceptible, immune, or infected. Finally, and also quite important, is that the host's response to exposure can vary widely, from showing no effect to manifesting subclinical disease, atypical symptoms, straightforward illness, or severe illness.

The **environment** is the conditions or influences that are not part of either the agent or the host, but that influence their interaction. A wide variety of factors, including physical, climatologic, biologic, social, and economic conditions, can come into play. For instance, in a study of motor vehicle injuries, the agent (mechanical energy) and the host (driver) could be affected by the topography, the weather, and the actions of other drivers. In many infectious disease outbreaks, social and economic conditions cause overcrowding and lead to high levels of exposure.

Back to Top

Agent, host, and environment alone are not sufficient to cause an epidemic; an adequate chain of transmission must be present. This process requires a source for the agent, a portal of exit, a mode of transmission, and a portal of entry. The first element, the **source for the agent**, is often the place where the agent originates, where it lives, grows, and multiplies, but this is not always the case. The agent that causes botulism (*Clostridium botulinum*), for example, originates in soil, but the source of most botulism infections is improperly canned food containing *C. botulinum* spores.

The second element, a **portal of exit**, is a pathway by which the agent can leave the source. This pathway is usually related to the place where the agent is localized. For instance, the agents causing tuberculosis and the flu are released through the respiratory tract, whereas agents for many stomach ailments are released through the digestive tract. Agents found in the blood, such as hepatitis B and HIV, can be released through cuts or needles.

Once the agent leaves the source, a **mode of transmission**, or means of carrying it to the host, is needed. This can happen in a number of ways, some of which are direct and some indirect. **Direct transmission** includes contact with soil or plants as well as contact between people. In **indirect transmission**, the agent can be **airborne**, **vector borne**, or **vehicle borne**. In airborne transmission, the agent is carried from the source to the host suspended in air particles. Vector-borne diseases are transmitted indirectly by a live carrier, usually an arthropod, such as mosquitos, fleas, or ticks. Vehicle-borne diseases are carried by inanimate objects, such as food or water, blood, or items like handkerchiefs, bedding, and surgical instruments.

Finally, there must be a pathway into the host, a **portal of entry**, that gives the agent access to tissue where it can multiply or act. Often the agent enters the host in the same way that it left the source. This is the case with the flu virus, which leaves the source through the respiratory tract and enters a new host through the respiratory tract.

Back to Top

A Few More Words about Field Epidemiology

Field epidemiology is, in the most general terms, the practice or application of epidemiology to control and prevent health problems. Epidemiologists are mobilized under a variety of circumstances, prime ones being when a problem is acute and unexpected and when quick action is required. The Legionnaires' disease outbreak in Philadelphia, mentioned at the beginning of this discussion, is an excellent example. These criteria are also met when a commercial product presents an imminent threat to public health and safety, as was the case with L-tryptophan and EMS. High levels of community concern often mandate a quick response. Involvement of the press is occasionally the driving force behind an investigation, and political pressure is also often part of the equation.

Field investigations are action oriented, with the main goal being to solve a pressing public health problem. Uppermost in investigators minds is the need to institute the controls necessary to safeguard health as soon as possible, and often, as in the example of L-tryptophan and EMS, this step is taken before the entire investigation is complete. Limited control over the situation, little time for planning a study, and limited data sources and laboratory samples challenge investigators. However, the obligation remains to do the best science possible under the circumstances.

Learn more about How to Investigate an Outbreak

Back to Top

This page last reviewed November 17, 2004
