Lesson 6
Investigating an Outbreak

One of the most exciting and challenging tasks facing an epidemiologist working in a public health department is investigating an outbreak. Frequently, the cause and source of the outbreak are unknown. Sometimes large numbers of people are affected. Often, the people in the community are concerned because they fear more people, including themselves, may be stricken unless the cause is found soon. There may be hostilities and defensiveness if an individual, product, or company has been accused of being the cause. Into this pressure-packed situation comes the epidemiologist, sometimes from the local health department, more often from "the outside." In this setting the epidemiologist must remain calm, professional, and scientifically objective. Fortunately, epidemiology provides the scientific basis, the systematic approach, and the population and prevention orientations that are needed.

Objectives

After studying this lesson and answering the questions in the exercises, a student will be able to do the following:

- List the reasons that health agencies investigate reported outbreaks
- List the steps in the investigation of an outbreak
- Define the terms **cluster, outbreak, epidemic**
- Given the initial information of a possible disease outbreak, describe how to determine whether an epidemic exists
- State what a line listing is and what it is used for
- Given information about a community outbreak of disease, execute the initial steps of an investigation and develop biologically plausible hypotheses
- Draw a traditional epidemic curve
- Given data in a two-by-two table, calculate the appropriate measure of association and chi-square test

Introduction to Investigating an Outbreak

Uncovering Outbreaks

One of the uses of surveillance--covered in Lesson 5--is the detection of outbreaks. Outbreaks may be detected when routine, timely analysis of surveillance data reveals an increase in reported cases or an unusual clustering of cases. In a health department, we may detect increases in or unusual patterns of disease from the weekly tabulations of case reports by time and place or from the examination of the exposure information on the case reports themselves. For example, health department staff detected an outbreak of hepatitis B that was transmitted by a dentist because they regularly reviewed and compared the dental exposures reported for hepatitis B cases (19). Similarly, in a hospital, weekly analysis of microbiologic isolates from patients by organism and ward may reveal an increased number of apparent nosocomial (hospital-acquired) infections in one part of the hospital.

Nonetheless, most outbreaks come to the attention of health authorities because an alert clinician is concerned enough to call the health department. The nationwide epidemic of eosinophilia-myalgia syndrome (EMS) was first detected when a physician in New Mexico called a consultant in Minnesota and realized that, together, they had seen three patients with a highly unusual clinical presentation. All three patients said they used L-tryptophan. The local physician promptly called the New Mexico State Health and Environment Department, which set into motion a chain of public health actions leading to the recall of L-tryptophan throughout the country (14,23).

Members of affected groups are another important reporting source for apparent clusters of both infectious and noninfectious disease. For example, someone may call a health department and report that he and several co-workers came down with severe gastroenteritis after attending a banquet several nights earlier. Similarly, a local citizen may call about several cases of cancer diagnosed among his neighbors and express concern that these are more than coincidental. Most health departments have routine procedures for handling calls from the public regarding potential communicable disease outbreaks, and a few states have developed guidelines for how to respond to noninfectious disease cluster reports (2,8,9).

Why Investigate Possible Outbreaks

Health departments investigate suspected outbreaks for a variety of reasons. These include the need to institute control and prevention measures; the opportunity for research and training; program considerations; and public relations, political concerns, and legal obligations.

Control/prevention

The primary public health reason to investigate an outbreak is to control and prevent further disease. Before we can develop control strategies for an outbreak, however, we must identify where the outbreak is in its natural course: Are cases occurring in increasing numbers or is the outbreak just about over? Our goal will be different depending on the answers to these questions.

If cases are continuing to occur in an outbreak, our goal may be to prevent additional cases. Therefore, the objective of our investigation would be to assess the extent of the outbreak and the size and characteristics of the population at risk in order to design and implement appropriate control measures.

On the other hand, if an outbreak appears to be almost over, our goal may be to prevent outbreaks in the future. In that case, the objective of our investigation is more likely to be to identify factors which contributed to the outbreak in order to design and implement measures that would prevent similar outbreaks in the future.

The balance between control measures versus further investigation depends on how much is known about the cause, the source, and the mode of transmission of the agent (11). Table 6.1 illustrates the relative emphasis as influenced by how much we know about these factors.

Table 6.1
Relative priority of investigative and control efforts during an outbreak, based on level of knowledge of the source, mode of transmission, and causative agent

		Source/Mode of Transmission: Known	Source/Mode of Transmission: Unknown
Causative Agent: Known		Investigation + Control +++	Investigation +++ Control +
Causative Agent: Unknown		Investigation +++ Control +++	Investigation +++ Control +

+++ = highest priority
+ = lower priority
Source: 11

If we know little about the source and mode of transmission, as indicated in the right-hand column of the table, we must investigate further before we can design appropriate control measures. In contrast, if we know the source and mode of transmission, as indicated in the left-hand column, control measures can be implemented immediately. However, if we don't know what the agent is, as indicated in the bottom row of the table, we must investigate further to identify the agent.

The public health response to the outbreak of EMS described earlier illustrates this point. Since investigators quickly determined that EMS was associated with the ingestion of L-tryptophan, that product was immediately withdrawn from the market, and persons were warned to avoid taking any they had on hand. However, officials continued the investigation for quite some time until they were certain they had identified the specific contaminant and reason that contamination occurred.

The decisions regarding whether and how extensively to investigate an outbreak are influenced by characteristics of the problem itself: the severity of the illness, the source or mode of transmission, and the availability of prevention and control measures. It is particularly urgent to investigate an outbreak when the disease is severe (serious illness with high risk of hospitalization, complications, or death) and has the potential to affect others unless prompt control measures are taken. For example, in the United States, every case of plague and botulism is investigated immediately to identify and eradicate the source. Cases of syphilis, tuberculosis, and measles are investigated promptly to identify contacts and interrupt further transmission.

Research opportunities

Another important objective of outbreak investigations is, simply, to gain additional knowledge. Each outbreak may be viewed as an experiment of nature waiting to be analyzed and exploited. Each presents a unique opportunity to study the natural history of the disease in question. For a newly recognized disease, field investigation provides an opportunity to define the natural history--including agent, mode of transmission, and incubation period--and the clinical spectrum of disease. Investigators also attempt to characterize the populations at greatest risk and to identify specific risk factors. Acquiring such information was an important motivation for investigators studying such newly recognized diseases as Legionnaires' disease in Philadelphia in 1976, toxic shock syndrome in 1980, acquired immunodeficiency syndrome in the early 1980's, and EMS in 1989.

Even for diseases that are well characterized, an outbreak may provide opportunities to gain additional knowledge by assessing the impact of control measures and the usefulness of new epidemiology and laboratory techniques. For example, an outbreak of measles in a highly immunized community provides a setting for investigators to study vaccine efficacy, the effect of age at vaccination, and the duration of vaccine-induced protection (16). An outbreak of giardiasis was used to study the appropriateness of a new clinical case definition (15), while an outbreak of pertussis was used to study the performance of a new culture medium (7).

Training

Investigating an outbreak requires a combination of diplomacy, logical thinking, problem-solving ability, quantitative skills, epidemiologic know-how, and judgment. These skills improve with practice and experience. Thus many investigative teams pair a seasoned epidemiologist with an epidemiologist-in-training. The latter gains valuable on-the-job training and experience while providing assistance in the investigation and control of the outbreak.

Public, political, or legal concerns

Public, political, or legal concerns sometimes override scientific concerns in the decision to conduct an investigation. Increasingly, the public has taken an interest in disease clusters and potential environmental exposures, and has called upon health departments to investigate. Such investigations almost never identify a causal link between exposure and disease (4,22). Nevertheless, many health departments have learned that it is essential to be "responsibly responsive" to public concerns, even if the concern has little scientific basis (9,2,18). Thus several states, recognizing their need to be responsive and an opportunity to educate the public, have adopted protocols for investigating disease clusters reported by its citizens. Some investigations are conducted because the law requires an agency to do so. For example, CDC's National Institute of Occupational Safety and Health (NIOSH) is required to evaluate the risks to health and safety in a workplace if requested to do so by three or more workers.

Program considerations

Many health departments routinely offer a variety of programs to control and prevent illnesses such as tuberculosis, vaccine-preventable diseases, and sexually transmitted diseases. An outbreak of a disease targeted by a public health program may reveal a weakness in that program and an opportunity to change or strengthen the program's efforts. Investigating the causes of an outbreak may identify populations which have been overlooked, failures in the intervention strategy, changes in the agent, or events beyond the scope of the program. By using an outbreak to evaluate the program's effectiveness, program directors can improve the program's future directions and strategies.

Exercise 6.1

During the previous year, nine residents of a community died from the same type of cancer. List some reasons that might justify an investigation.

Answers on page 398.

Steps of an Outbreak Investigation

In the investigation of an ongoing outbreak, working quickly is essential. Getting the right answer is essential, too. Under such circumstances, epidemiologists find it useful to have a systematic approach to follow, such as the sequence listed in Table 6.2. This approach ensures that the investigation proceeds forward without missing important steps along the way.

Table 6.2
Steps of an outbreak investigation

1. Prepare for field work
2. Establish the existence of an outbreak
3. Verify the diagnosis
4. Define and identify cases
   a. establish a case definition
   b. identify and count cases
5. Perform descriptive epidemiology
6. Develop hypotheses
7. Evaluate hypotheses
8. As necessary, reconsider/refine hypotheses and execute additional studies
   a. additional epidemiologic studies
   b. other types of studies – laboratory, environmental
9. Implement control and prevention measures
10. Communicate findings

The steps described in Table 6.2 are in conceptual order. In practice, however, several steps may be done at the same time, or the circumstances of the outbreak may dictate that a different order be followed. For example, control measures should be implemented as soon as the source and mode of transmission are known, which may be early or late in any particular outbreak investigation.

Step 1: Preparing for Field Work

Anyone about to embark on an outbreak investigation should be well prepared before leaving for the field. Preparations can be grouped into three categories: (a) investigation, (b) administration, and (c) consultation. Good preparation in all three categories will facilitate a smooth field experience.

(a) *Investigation*

First, as a field investigator, you must have the appropriate scientific knowledge, supplies, and equipment to carry out the investigation. You should discuss the situation with someone knowledgeable about the disease and about field investigations, and review the applicable literature. You should assemble useful references such as journal articles

and sample questionnaires.

Before leaving for a field investigation, consult laboratory staff to ensure that you take the proper laboratory material and know the proper collection, storage, and transportation techniques. Arrange for a portable computer, dictaphone, camera, and other supplies.

(b) *Administration*
Second, as an investigator, you must pay attention to administrative procedures. In a health agency, you must make travel and other arrangements and get them approved. You may also need to take care of personal matters before you leave, especially if the investigation is likely to be lengthy.

(c) *Consultation*
Third, as an investigator, you must know your expected role in the field. Before departure, all parties should agree on your role, particularly if you are coming from "outside" the local area. For example, are you expected to lead the investigation, provide consultation to the local staff who will conduct the investigation, or simply lend a hand to the local staff? In addition, you should know who your local contacts will be. Before leaving, you should know when and where you are to meet with local officials and contacts when you arrive in the field.

Step 2: Establishing the Existence of an Outbreak

An **outbreak** or an **epidemic** is the occurrence of more cases of disease than expected in a given area or among a specific group of people over a particular period of time. In contrast, a **cluster** is an aggregation of cases in a given area over a particular period without regard to whether the number of cases is more than expected. In an outbreak or epidemic, we usually presume that the cases are related to one another or that they have a common cause.

Many epidemiologists use the terms "outbreak" and "epidemic" interchangeably, but the public is more likely to think that "epidemic" implies a crisis situation. Some epidemiologists restrict the use of the term "epidemic" to situations involving larger numbers of people over a wide geographic area.

Most outbreaks come to the attention of health departments in one of two ways. One way is by regular analysis of surveillance data. As noted in Lesson 5, unusual rises or patterns of disease occurrence can be detected promptly if surveillance data collection and analysis are timely. The second, and probably more common, way is through calls from a health care provider or citizen who knows of "several cases." For example, a member of the public may report three infants born with birth defects within a 1-month period in the same community. This aggregation of cases *seems* to be unusual, but frequently the public does not know the denominator--e.g., the total number of births--or the expected incidence of birth defects.

One of your first tasks as a field investigator is to verify that a purported outbreak is indeed an outbreak. Some will turn out to be true outbreaks with a common cause, some will be sporadic and unrelated cases of the same disease, and others will turn out to be unrelated cases of similar

but unrelated diseases. Often, you must first determine the expected number of cases before deciding whether the observed number exceeds the expected number, i.e., whether a cluster is indeed an outbreak.

Thus, as in other areas of epidemiology, you compare the **observed with the expected**. How then, do you determine what's expected? Usually we compare the current number of cases with the number from the previous few weeks or months, or from a comparable period during the previous few years.

- For a notifiable disease, you can use health department surveillance records.
- For other diseases and conditions, you can usually find existing data locally--hospital discharge records, mortality statistics, cancer or birth defect registries.
- If local data are not available, you can apply rates from neighboring states or national data, or, alternatively, you may conduct a telephone survey of physicians to determine whether they have seen more cases of the disease than usual.
- Finally, you may conduct a survey of the community to establish the background or historical level of disease.

Even if the current number of reported cases exceeds the expected number, the excess may not necessarily indicate an outbreak. Reporting may rise because of changes in local reporting procedures, changes in the case definition, increased interest because of local or national awareness, or improvements in diagnostic procedures. A new physician, infection control nurse, or health care facility may see referred cases and more consistently report cases, when in fact there has been no change in the actual occurrence of the disease. Finally, particularly in areas with sudden changes in population size such as resort areas, college towns, and migrant farming areas, changes in the numerator (number of reported cases) may simply reflect changes in the denominator (size of the population).

Whether you should investigate an apparent problem further is not strictly tied to your verifying that an epidemic exists (observed numbers greater than expected). As noted earlier, the severity of the illness, the potential for spread, political considerations, public relations, available resources, and other factors all influence the decision to launch a field investigation.

Exercise 6.2

For the month of August, 12 new cases of tuberculosis and 12 new cases of aseptic meningitis were reported to a county health department. Would you call either group of cases a cluster? Would you call either group of cases an outbreak? What additional information might be helpful in answering these questions?

Answers on page 398.

Step 3: Verifying the Diagnosis

Closely linked to verifying the existence of an outbreak is establishing what disease is occurring. In fact, as an investigator, you frequently will be able to address these two steps at the same time. Your goals in verifying the diagnosis are (a) to ensure that the problem has been properly diagnosed and (b) to rule out laboratory error as the basis for the increase in diagnosed cases.

In verifying the diagnosis you should review the clinical findings and laboratory results. If you have any question about the laboratory findings, i.e., if the laboratory tests are inconsistent with the clinical and epidemiologic findings, you should have a qualified laboratorian review the laboratory techniques being used. If you plan specialized laboratory work such as confirmation in a reference laboratory, DNA or other chemical or biological fingerprinting, or polymerase chain reaction, you must secure the appropriate specimens, isolates, and other laboratory material as soon as possible, and from a sufficient number of patients.

You should always summarize the clinical findings with frequency distributions (see Lessons 2 and 3 for a discussion of frequency distributions). Such frequency distributions are useful in characterizing the spectrum of illness, verifying the diagnosis, and developing case definitions. Many investigators consider these clinical frequency distributions so important that they routinely present these findings in the first table of their report or manuscript.

Finally, you should visit several patients with the disease. If you do not have the clinical background to verify the diagnosis, a qualified clinician should do so. Nevertheless, regardless of background, you should see and talk to some patients to gain a better understanding of the clinical features, and to develop a mental image of the disease and the patients affected by it. In addition, you may be able to gather critical information from these patients: What were their exposures before becoming ill? What do *they* think caused their illness? Do they know anyone else with the disease? Do they have anything in common with others who have the disease? Conversations with patients are very helpful in generating hypotheses about disease etiology and spread.

Step 4a: Establishing a Case Definition

Your next task as an investigator is to establish a case definition. A case definition is a standard set of criteria for deciding whether an individual should be classified as having the health condition of interest. A case definition includes clinical criteria and--particularly in the setting of an outbreak investigation--restrictions by time, place, and person. You should base the clinical criteria on simple and objective measures such as elevated antibody titers, fever $\geq 101°F$, three or more loose bowel movements per day, or myalgias severe enough to limit the patient's usual activities. You may restrict the case definition by time (for example, to persons with onset of illness within the past 2 months), by place (for example, to residents of the nine-county area or to employees of a particular plant) and by person (for example, to persons with no previous history of musculo-skeletal disease, or to pre-menopausal women). Whatever your criteria, you must apply them consistently and without bias to all persons under investigation.

Be careful that the case definition does not include an exposure or risk factor you want to test. This is a common mistake. For example, do not define a case as "illness X among persons who were in homeless shelter Y" if one of the goals of the investigation is to determine whether the shelter is associated with illness.

Ideally, your case definition will include most if not all of the actual cases, but very few or none of what are called "false-positive" cases (persons who actually do not have the disease in question but nonetheless meet the case definition). Recognizing the uncertainty of some diagnoses, investigators often classify cases as confirmed, probable, or possible.

To be classified as confirmed, a case usually must have laboratory verification. A case classified as probable usually has typical clinical features of the disease without laboratory confirmation. A case classified as possible usually has fewer of the typical clinical features. For example, in an outbreak of bloody diarrhea and hemolytic-uremic syndrome caused by infection with *E. coli* O157:H7, investigators defined cases in the following three classes:

- **Definite** case: *E. coli* O157:H7 isolated from a stool culture or development of hemolytic-uremic syndrome in a school-age child resident of the county with gastrointestinal symptoms beginning between November 3 and November 8, 1990
- **Probable** case: Bloody diarrhea, with the same person, place, and time restrictions
- **Possible** case: Abdominal cramps and diarrhea (at least three stools in a 24-hour period) in a school-age child with onset during the same period (CDC, unpublished data, 1991).

As an investigator, you will find such classifications useful in several situations. First, they will allow you to keep track of a case even if the diagnosis is not confirmed. For example, you might temporarily classify a case as probable or possible while laboratory results are pending. Alternatively, the patient's physician or you may have decided not to order the laboratory test required to confirm the diagnosis because the test is expensive, difficult to obtain, or unnecessary. For example, during a community outbreak of measles, which has a characteristic clinical picture, investigators might follow the usual practice of confirming only a few cases and then relying on clinical features to identify the rest of the cases. Similarly, while investigating an outbreak of diarrhea on a cruise ship, investigators usually try to identify an agent from stool samples from a few afflicted persons. If those few cases are confirmed to be infected with the same agent, the other persons with compatible clinical illness are all presumed to be part of the same outbreak.

Early in an investigation, investigators often use a sensitive or "loose" case definition which includes confirmed, probable, and even possible cases. Later on, when hypotheses have come into sharper focus, the investigator may "tighten" the case definition by dropping the possible category. You will find this a useful strategy in investigations that require you to travel to different hospitals, homes, or other sites to gather information, because it is better to collect extra

data while you're there than to have to go back. This illustrates an important axiom of field epidemiology: "Get it while you can."

A "loose" case definition is used early in the investigation to identify the extent of the problem and the populations affected. Important hypotheses may arise from this process. However, in analytic epidemiology, inclusion of false-positive cases can produce misleading results. Therefore, to test these hypotheses using analytic epidemiology (see page 375), specific or "tight" case definitions must be used.

Step 4b: Identifying and Counting Cases

As noted earlier, many outbreaks are brought to the attention of health authorities by concerned health care providers or citizens. However, the cases which prompted the concern are often only a small and nonrepresentative fraction of the total number of cases. Public health workers must therefore "cast the net wide" to determine the geographic extent of the problem and the populations affected by it.

When you need to identify cases, use as many sources as you can. You may have to be creative, aggressive, and diligent in identifying these sources. Your methods for identifying cases must be appropriate for the setting and disease in question.

First, direct your case finding at health care facilities where the diagnosis is likely to be made: physicians' offices, clinics, hospitals, and laboratories. If you send out a letter describing the situation and asking for reports, that is called "stimulated or enhanced passive surveillance." Alternatively, if you telephone or visit the facilities to collect information on cases, that is called "active surveillance."

In some outbreaks, public health officials may decide to alert the public directly, usually through the local media. For example, in outbreaks caused by a contaminated food product such as salmonellosis caused by contaminated milk (21) or L-tryptophan-induced EMS (14), announcements in the media alerted the public to avoid the implicated product and to see a physician if they had symptoms compatible with the disease in question.

If an outbreak affects a restricted population, such as on a cruise ship, in a school, or at a worksite, and if a high proportion of cases are unlikely to be diagnosed (if, for example, many cases are mild or asymptomatic), you may want to conduct a survey of the entire population. You could administer a questionnaire to determine the true occurrence of clinical symptoms, or you could collect laboratory specimens to determine the number of asymptomatic cases.

Finally, you can ask case-patients if they know anyone else with the same condition. Frequently, one person with an illness knows or hears of others with the same illness.

Regardless of the particular disease you are investigating, you should collect the following types of information about every case:

- identifying information
- demographic information
- clinical information
- risk factor information
- reporter information

Identifying information–name, address, and telephone number–allows you and other investigators to contact patients for additional questions, and to notify them of laboratory results and the outcome of the investigation. Names will help you in checking for duplicate records, while the addresses allow you to map the geographic extent of the problem.

Demographic information–age, sex, race, and occupation–provides the "person" characteristics of descriptive epidemiology you need to characterize the populations at risk.

Clinical information allows you to verify that the case definition has been met. Date of onset allows you to chart the time course of the outbreak. Supplementary clinical information, including whether hospitalization or death occurred, will help you describe the spectrum of illness.

You must tailor risk factor information to the specific disease in question. For example, in an investigation of hepatitis A, you would ascertain exposure to food and water sources.

Finally, by identifying the person who provided the case report, you will be able to seek additional clinical information or report back the results of your investigation.

Traditionally, we collect the information described above on a standard case report form, questionnaire, or data abstraction form. We then abstract selected critical items on a form called a line listing. An example of a line listing is shown in Figure 6.1.

In a line listing, each column represents an important variable, such as name or identification number, age, sex, case classification, etc., while each row represents a different case. New cases are added to a line listing as they are identified. Thus, a line listing contains key information on every case, and can be scanned and updated as necessary. Even in the era of microcomputers, many epidemiologists still maintain a hand-written line listing of key data items, and turn to their computers for more complex manipulations, cross-tabulations, and the like.

Figure 6.1
Example of line listing for an outbreak of hepatits A

Line Listing of reported suspect cases, page 1

Case #	Initials	Date of Report	Date of Onset	Diagnostic								Lab		Age	Sex
				MD Dx	Signs and Symptoms							HA IgM	Other		
					N	V	A	F	DU	J					
1	JG	10/12	10/6	Hep A	+	+	+	+	+	+		+	SGOT ↑	37	M
2	BC	10/12	10/5	Hep A	+	–	+	+	+	+		+	ALT ↑	62	F
3	HP	10/13	10/4	Hep A	±	–	+	+	+	S*		+	SGOT ↑	30	F
4	MC	10/15	10/4	Hep A	–	–	+	+	?	–		+	HBs Ag -	17	F
5	NG	10/15	10/9	NA	–	–	+	–	+	+		NA	NA	32	F
6	RD	10/15	10/8	Hep A	+	+	+	+	+	+		+		38	M
7	KR	10/16	10/13	Hep A	±	–	+	+	+	+		+	SGOT = 240	43	M
8	DM	10/16	10/12	Hep A	–	–	+	+	+	–		+		57	M
9	PA	10/18	10/7	Hep A	±	–	+	±	+	+		+		52	F
10	SS	10/11	10/11	R/o Hep A Hep	+	+	+		+	–		pending	HBsAg pending	31	M

S* = scleral
N = nausea
V = vomiting
A = anorexia
F = fever
DU = dark urine
J = jaundice
HA IgM = hepatitis A IgM antibody test

Exercise 6.3

Review the six case report forms in Appendix G. Create a line listing based on this information.

Answers on page 399.

Step 5: Performing Descriptive Epidemiology

Once you have collected some data, you can begin to characterize an outbreak by time, place, and person. In fact, you may wind up performing this step several times during the course of an outbreak. Characterizing an outbreak by these variables is called **descriptive epidemiology**, because you describe what has occurred in the population under study. This step is critical for several reasons. First, by looking at the data carefully, you become familiar with them. You can learn what information is reliable and informative (such as if many cases report the same unusual exposure) and learn what may not be as reliable (for example, many missing or "don't know" responses to a particular question). Second, you provide a comprehensive description of an outbreak by portraying its trend over time, its geographic extent (place), and the populations (persons) affected by the disease. You can assess your description of the outbreak in light of what is known about the disease (usual source, mode of transmission, risk factors and populations affected, etc.) to develop causal hypotheses. You can, in turn, test these hypotheses using the techniques of analytic epidemiology, described under Step 7.

Note that you should begin descriptive epidemiology early, and should update it as you collect additional data. To keep an investigation moving quickly and in the right direction, you must discover both errors and clues in the data as early as possible.

Time

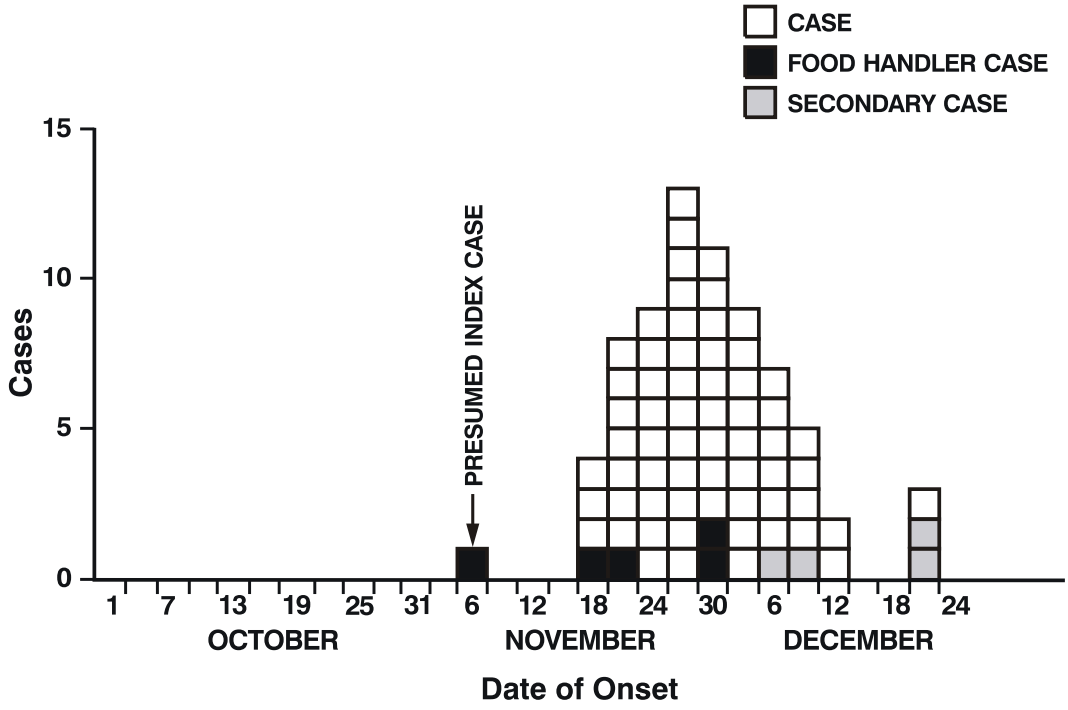
Traditionally, we depict the time course of an epidemic by drawing a histogram of the number of cases by their date of onset. This graph, called an **epidemic curve**, or **epi curve** for short, gives us a simple visual display of the outbreak's magnitude and time trend. Figure 6.2 shows a typical epidemic curve. This visual display can be understood by both epidemiologists and non-epidemiologists alike.

An epidemic curve will provide you with a great deal of information about an epidemic. First, you will usually be able to tell where you are in the time course of an epidemic, and what the future course might be. Second, if you have identified the disease and know its usual incubation period, you usually can deduce a probable time period of exposure and can develop a questionnaire focusing on that time period. Finally, you may be able to draw inferences about the epidemic pattern--whether it is common source or propagated, or both. For a review of epidemic patterns see Lesson 1.

How To Draw an Epidemic Curve. To draw an epidemic curve, you first must know the time of onset of illness for each case. For most diseases, date of onset is sufficient; for a disease with a very short incubation period, hours of onset may be more suitable.

Next, select the unit of time on the *x*-axis. We usually base these units on the incubation period of the disease (if known) and the length of time over which cases are distributed. As a rule of thumb, select a unit that is one-eighth to one-third, i.e., roughly one-quarter as long as the incubation period. Thus, for an outbreak of *Clostridium perfringens* food poisoning (usual incubation period 10-12 hours), with cases confined to a few days, you could use an *x*-axis unit of 2 or 3 hours. Unfortunately, we often need to draw an epidemic curve when we don't know the

Figure 6.2
Typical epidemic curve: Hepatitis A cases by date of onset, Fayetteville, Arkansas, November-December 1978



Source: CDC, unpublished data, 1978

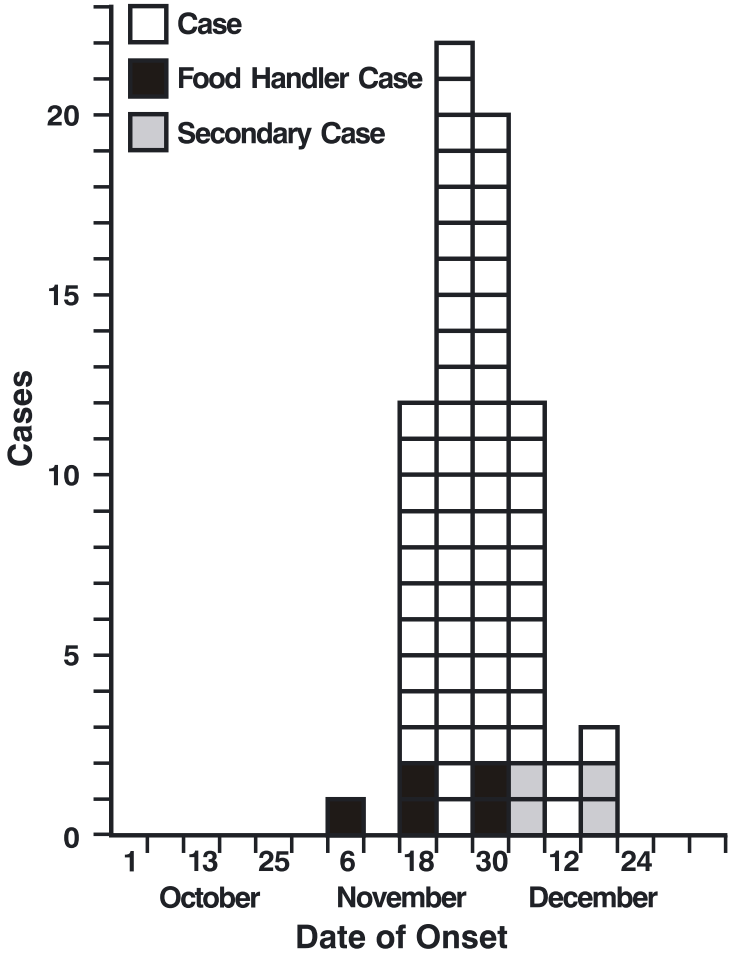
disease and/or its incubation time. In that circumstance, it is useful to draw several epidemic curves with different units on the *x*-axis to find one that seems to portray the data best. For example, Figure 6.3 shows an epidemic curve of the same data as in Figure 6.2; in Figure 6.2 the *x*-axis unit is 3 days and in Figure 6.3 the *x*-axis unit is 6 days. Which unit seems to provide the most useful information about the course of the epidemic?

The units used for the *x*-axis in Figures 6.2 and 6.3 are both useful. They both demonstrate a point-source epidemic. The unit selected for Figure 6.2 is preferred because (1) it distributes the cases more clearly, and (2) it separates out the presumed index case more clearly.

Finally, show the pre-epidemic period on your graph to illustrate the background or "expected" number of cases. (Remember, an epidemic is defined as more cases than expected.) For a disease with a human host, such as hepatitis A, one of the early cases may be a foodhandler who is the source of the epidemic! Notice that both Figure 6.2 and 6.3 show a relatively long pre-epidemic period.

Interpreting an Epidemic Curve. The first step in interpreting an epidemic curve is to consider its overall shape. The shape of the epidemic curve is determined by the epidemic pattern (common source versus propagated), the period of time over which susceptible persons are exposed, and the minimum, average, and maximum incubation periods for the disease.

Figure 6.3
Epidemic curve with different units on *x*-axis:
Hepatitis A cases by date of onset, Fayetteville, Arkansas,
November-December 1978



Source: CDC, unpublished data, 1978

An epidemic curve which has a steep upslope and a more gradual downslope (a log-normal curve) indicates a **point source** epidemic in which persons are exposed to the same source over a relative brief period. In fact, any sudden rise in the number of cases suggests sudden exposure to a common source.

In a point source epidemic, all the cases occur within one incubation period. If the duration of exposure was prolonged, the epidemic is called a **continuous common source** epidemic, and the epidemic curve will have a plateau instead of a peak. Intermittent common source epidemics produce irregularly jagged epidemic curves which reflect the intermittency and duration of

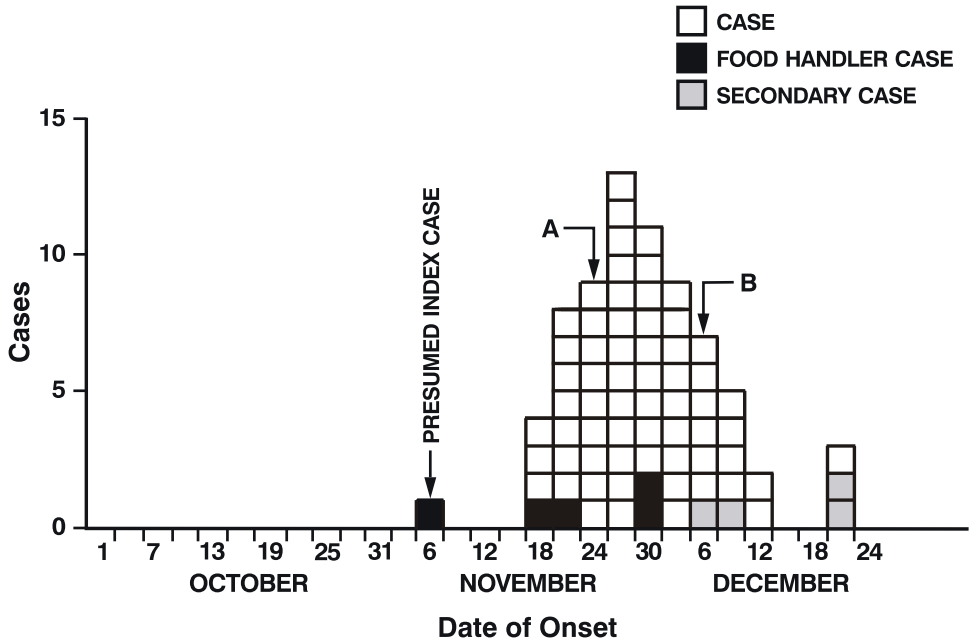
exposure, and the number of persons exposed. Person-to-person spread – a **propagated** epidemic – should have a series of progressively taller peaks one incubation period apart, but in reality few produce this classic pattern.

When you examine an epidemic curve, you should determine where you are in the epidemic. For example, suppose you plotted an epidemic curve of the data in Figure 6.4 when you had only data through November 26 – that is, only through point A. At that point, it should seem clear to you that the outbreak is still on the upswing, and you could safely predict that new cases would continue to occur. On the other hand, if you plotted an epidemic curve using the data through point B, you should realize that the outbreak has peaked and may soon be over, although, depending on the disease, a few late or secondary cases might still occur.

The cases that stand apart may be just as informative as the overall pattern. An early case may represent a background or unrelated case, a source of the epidemic, or a person who was exposed earlier than most of the cases (the cook who tasted her dish hours before bringing it to the big picnic!). Similarly, late cases may represent unrelated cases, long-incubation-period cases, secondary cases, or persons exposed later than most of the cases. On the other hand, these outliers sometimes represent miscoded or erroneous data. All outliers are worth examining carefully because if they are part of the outbreak, their unusual exposures may point directly to the source.

In a point-source epidemic of a known disease with a known incubation period, you can use the epidemic curve to identify a likely period of exposure. This is critical to asking the right questions to identify the source of the epidemic.

Figure 6.4
Typical epidemic curve with point A on upslope and point B on downslope



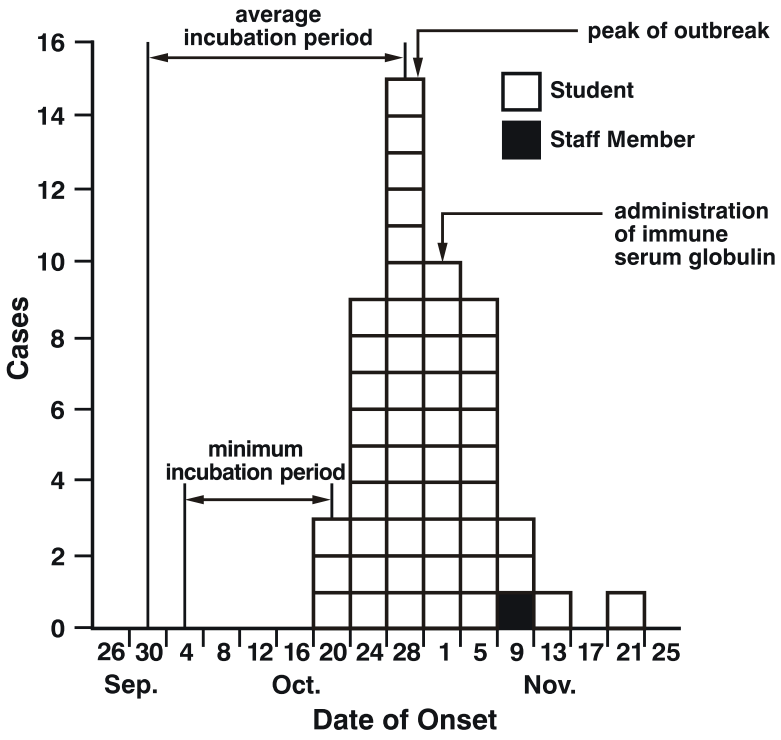
To identify the likely period of exposure from an epidemic curve,

1. Look up the average and minimum incubation periods of the disease. This information can be found in *Control of Communicable Diseases in Man* (3).
2. Identify the peak of the outbreak or the median case and count back on the *x*-axis one average incubation period. Note the date.
3. Start at the earliest case of the epidemic and count back the minimum incubation period, and note this date as well.

Ideally, the two dates are similar, and represent the probable period of exposure. This technique is not precise, however, and you usually should widen the period of exposure by 10-20% on either side of these dates. You should then ask about exposures during the wider period in an attempt to identify the source.

For example, consider the outbreak of hepatitis A illustrated by the epidemic curve in Figure 6.5. The incubation period for hepatitis A ranges from 15 to 50 days, with an average incubation period of 28-30 days (roughly one month). First, is this epidemic curve consistent with a point source? That is, do all 48 cases fall within one incubation period?

Figure 6.5
Hepatitis A cases in Colbert County, Alabama, October-November 1972



The epidemic is consistent with a point source because the last case is within 35 days (50 - 15) of the first case. Therefore, we can use the epidemic curve to identify the likely period of exposure by making the following determinations:

1. What is the peak of the outbreak or the median date of onset?

   The peak of the outbreak occurred during the 4-day interval beginning on October 28. The median date of onset of the 48 cases lies between the 24th and 25th case. Both of these occurred during the same 4-day period.

2. What would be the beginning of one average incubation period prior to the peak (median date) of the outbreak?

   Since the interval containing both the peak and the median of the outbreak includes the last four days of October, one month earlier would fall during the last few days of September.

3. What would be the beginning of one minimum incubation period before the first case?

   The earliest case occurred on October 20. Subtracting 15 days from October 20 points us to October 5.

Thus we would look for exposures around the end of September and the beginning of October. This turned out to be the exact period during which there had been a temporary lapse in chlorination of the school's water supply (4)!

Exercise 6.4

Using the data from a hepatitis A outbreak, draw an epidemic curve. From your epidemic curve and your knowledge of the average and minimum incubation periods for hepatitis A, identify the likely exposure period. Work space provided on page 368.

Case #	Age	Sex	Date of Onset	Case #	Age	Sex	Date of Onset
2	16	F	4-3	41	37	F	5-9
3	34	M	4-6	43	16	M	5-10
6	15	M	4-28	45	29	F	5-10
7	46	M	4-30	46	5	M	5-10
8	21	F	5-1	47	8	F	5-11
9	14	M	5-1	48	15	F	5-11
11	13	M	5-2	49	14	M	5-11
12	43	M	5-2	50	16	M	5-11
13	14	M	5-3	52	16	M	5-12
15	37	M	5-3	53	19	M	5-12
16	5	F	5-3	54	15	M	5-12
17	11	F	5-3	55	10	F	5-12
18	19	M	5-4	56	6	M	5-12
19	14	F	5-4	57	20	M	5-12
20	35	F	5-4	58	43	M	5-12
21	11	F	5-4	59	15	F	5-12
22	14	M	5-4	60	12	F	5-12
23	14	M	5-4	61	14	M	5-13
25	15	M	5-5	62	34	M	5-13
26	12	M	5-5	63	15	F	5-13
27	50	M	5-5	64	30	M	5-13
29	50	M	5-6	65	16	M	5-13
31	11	M	5-7	66	15	M	5-14
32	15	M	5-7	67	15	M	5-14
33	18	F	5-7	68	16	M	5-14
34	14	M	5-7	69	16	M	5-14
35	15	M	5-8	70	18	F	5-15
36	30	M	5-8	72	12	M	5-18
37	20	F	5-9	74	22	F	5-20
38	14	F	5-9	75	15	F	5-24
39	17	M	5-9	76	14	M	5-26
40	15	M	5-9				

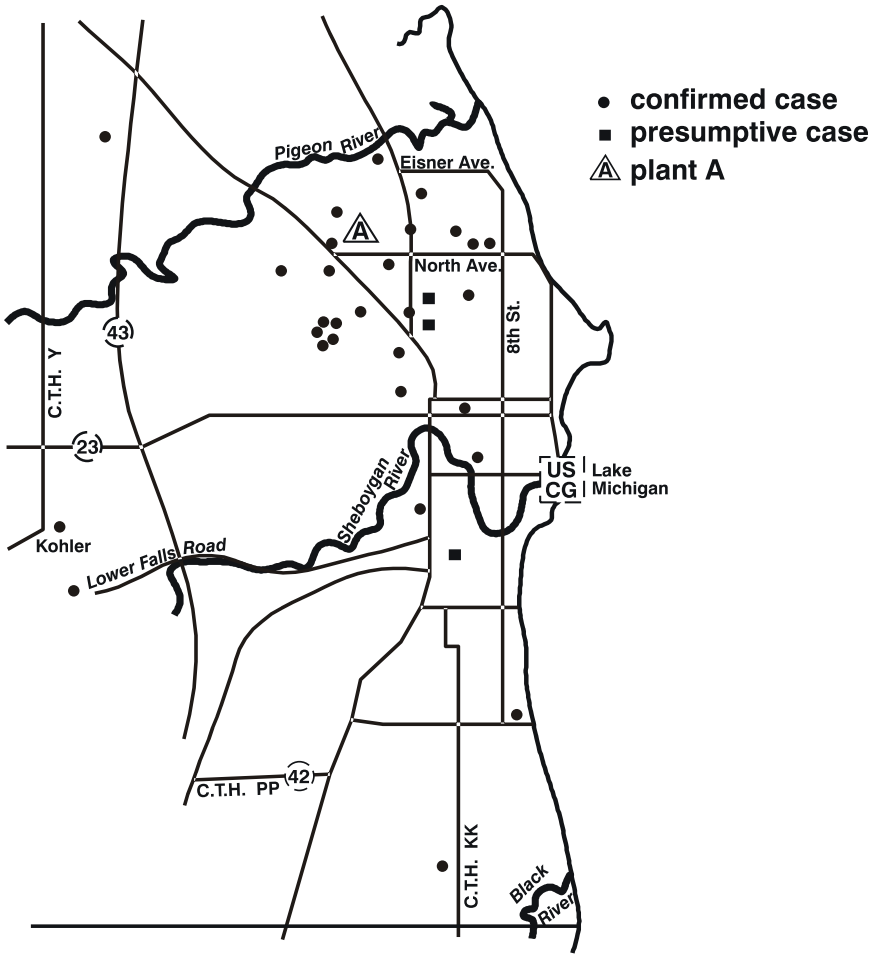
Answers on page 400.

Place

Assessment of an outbreak by place not only provides information on the geographic extent of a problem, but may also demonstrate clusters or patterns that provide important etiologic clues. A spot map is a simple and useful technique for illustrating where cases live, work, or may have been exposed.

On a spot map of a community, clusters or patterns may reflect water supplies, wind currents, or proximity to a restaurant or grocery. In Figure 6.6, for example, the homes of patients with Legionnaires' disease is shown in relation to the cooling tower at plant A (1).

Figure 6.6
Residence of patients with Legionnaires' disease, Sheboygan, Wisconsin, 1986

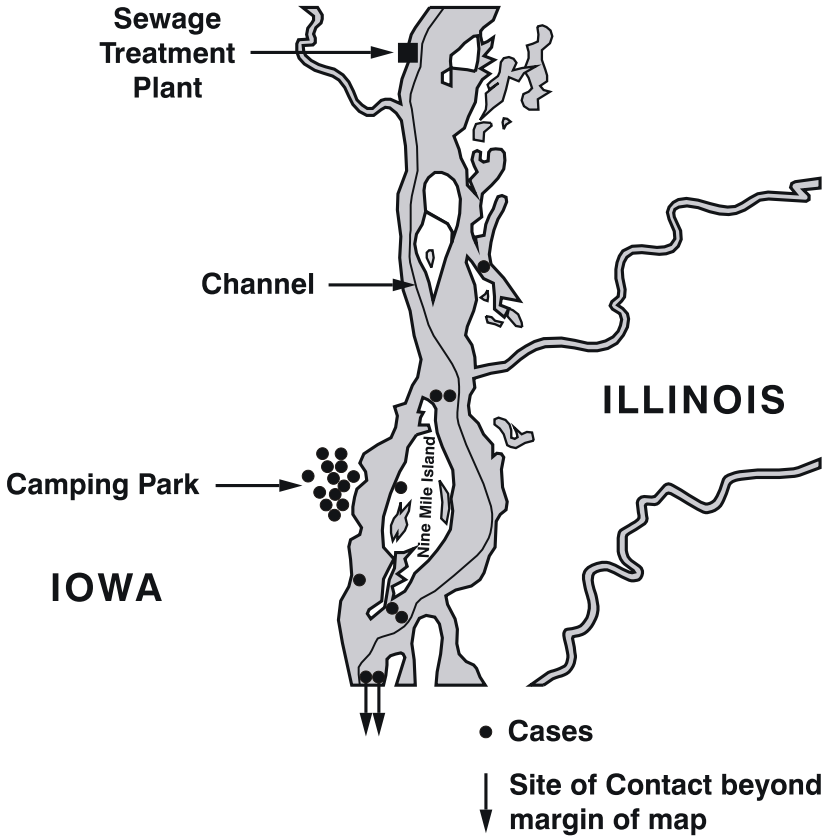


Source: 1

On a spot map of a hospital, nursing home, or other such facility, clustering is consistent with either a focal source or person-to-person spread, while scattering of cases throughout the facility is more consistent with a widely disseminated vehicle or a source common to the residents that is not associated with room assignment, such as a common dining hall.

Although we often use spot maps to plot location of residence, place of work is sometimes more revealing. Certainly, place of work is important in assessing "sick building syndrome" and other disorders related to air-flow patterns in buildings. In studying an outbreak of surgical wound infections in a hospital, we might plot cases by operating room, recovery room, and ward room to look for clustering. We can even use maps to plot recreational opportunities. For example, Figure 6.7 shows persons with shigellosis plotted by where they swam in the Mississippi River (20).

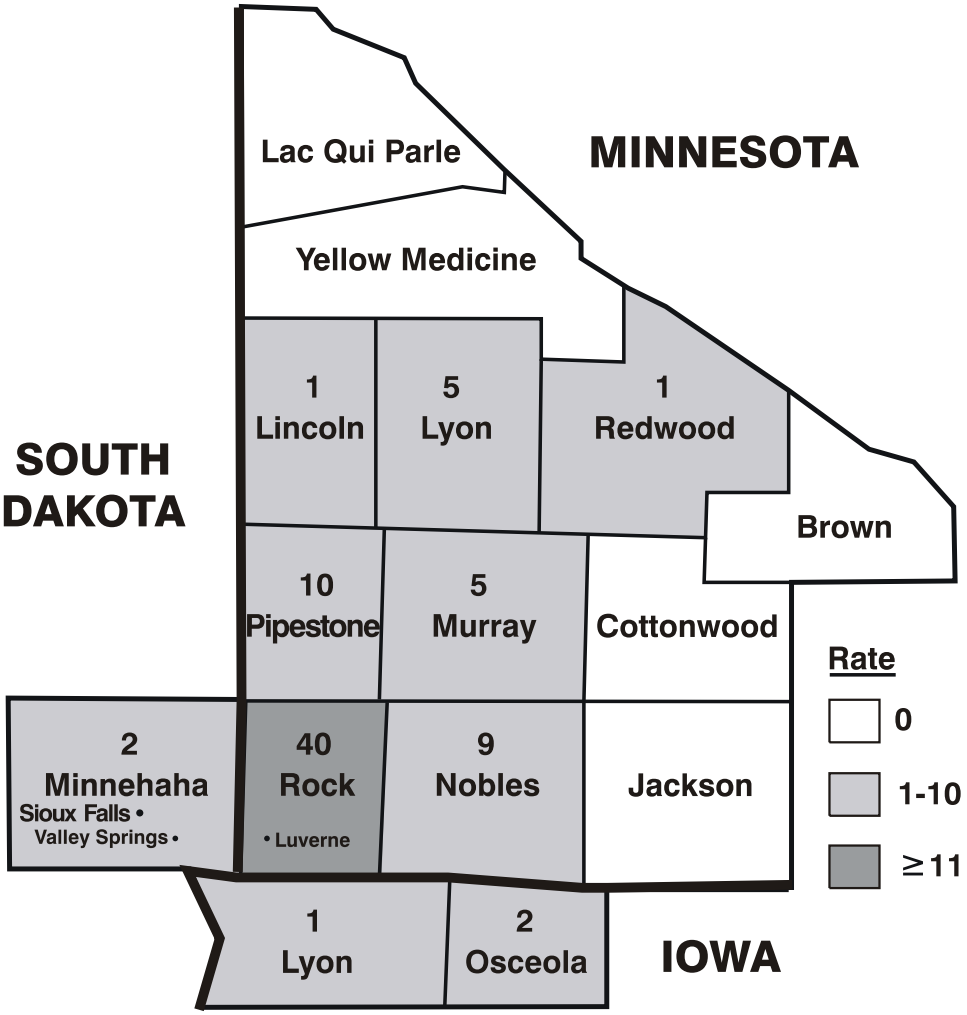
Figure 6.7
Mississippi River sites where 22 culture-positive cases swam within three days of onset of illness



Source: 20

If the size of the population varies between the areas you are comparing, a spot map--which shows numbers of cases–can be misleading. This is a weakness of spot maps. In such an instance, you should show area-specific attack rates with an area map. For example, Figure 6.8 is an area map that shows county-specific attack rates of thyrotoxicosis in 15 counties near the junction of Minnesota, South Dakota, and Iowa (13). If we had used a spot map to plot cases rather than rates, we might have misinterpreted the risk among Minnehaha residents. Seventeen residents of that county were affected, exceeded only by Rock County (43) and Nobles County (20). But because the population of Minnehaha is much larger than the population of the other counties, the risk was actually fairly low. Since this outbreak crosses state lines, it alerts us to maintain broad perspective and not restrict our thinking to artificial geopolitical boundaries.

Figure 6.8
Rate per 10,000 persons of thyrotoxicosis by county, Minnesota, South Dakota, and Iowa, February 1984-August 1985



Source: 13

Person

Characterizing an outbreak by person is how we determine what populations are at risk for the disease. We usually define such populations by host characteristics (age, race, sex, or medical status) or by exposures (occupation, leisure activities, use of medications, tobacco, drugs). Both of these influence susceptibility to disease and opportunities for exposure. As described in Lesson 2, we use rates to identify high-risk groups. In order to calculate rates, we must first have both numerators (numbers of cases) and denominators (number of people at risk).

Usually, age and sex are the two host factors we assess first, because they are often the person characteristics most strongly related to exposure and to the risk of disease. The categories used for age and sex in a frequency distribution should be appropriate for the particular disease and should match the available denominator data.

In many outbreaks, occupation is another important person characteristic. Although we like to calculate rates, it may be difficult to get denominator data for occupation. Nonetheless, the distribution of the cases themselves may suggest hypotheses worth pursuing.

Other person characteristics to analyze will be more specific to the disease under investigation and the setting of the outbreak. For example, if you were investigating an outbreak of hepatitis B, you should consider the usual high-risk exposures for that infection, such as intravenous drug use, sexual contacts, and health care employment. You might characterize an outbreak centered in a school by grade or classroom, and by student versus teacher or other staff.

Summarizing by Time, Place, and Person

After characterizing an outbreak by time, place, and person, it is useful to summarize what you know. For example, during an investigation of a different outbreak of Legionnaires' disease, this time in Louisiana, members of the investigative team discussed what they knew based on the descriptive epidemiology (6). Specifically, the epidemic curve indicated that the outbreak was basically over; no new case had been reported in the last two weeks. The affected population had a greater proportion of persons who were black, female, young, and less likely to smoke than persons in the usual Legionnaires' outbreak. There appeared to be no clustering by either residence or worksite, and no connection with exposure to the town's cooling towers. Thus the investigators were forced to develop new hypotheses about a source of Legionnaires' disease to explain this outbreak.

Step 6: Developing Hypotheses

The next conceptual step in an investigation is formulating hypotheses. However, in reality we usually begin to generate hypotheses with the first phone call. But at this point in an investigation, after talking with some case-patients and with local public health officials, and having characterized the outbreak by time, place, and person, our hypotheses will be sharpened and more accurately focused. The hypotheses should address the source of the agent, the mode (and vehicle or vector) of transmission, and the exposures that caused the disease. Also, the

hypotheses should be testable, since evaluating hypotheses is one of the goals of the next step in an investigation.

You can generate hypotheses in a variety of ways. First, consider what you know about the disease itself: What is the agent's usual reservoir? How is it usually transmitted? What vehicles are commonly implicated? What are the known risk factors? In other words, simply by becoming familiar with the disease, you can, at the very least, "round up the usual suspects."

Another useful way you can generate hypotheses is to talk to a few of the case-patients, as discussed under "Step 3: Verifying the Diagnosis." Your conversations about possible exposures should be open-ended and wide-ranging, not necessarily confined to the known sources and vehicles. In some difficult investigations which yielded few clues, investigators have convened a meeting of several case-patients to search for common exposures. In addition,investigators have sometimes found it useful to visit the homes of case-patientsand look through their refrigerators and shelves for clues.

Just as case-patients may have important insights into causes, so too may the local health department staff. The local staff know the people in the community and their practices, and often have hypotheses based on their knowledge.

The descriptive epidemiology often provides some hypotheses. If the epidemic curve points to a narrow period of exposure, what events occurred around that time? Why do the people living in a particular area have the highest attack rates? Why are some groups with particular age, sex, or other person characteristics, at greater risk than other groups with different person characteristics? Such questions about the data should lead to hypotheses which can be tested by appropriate analytic techniques.

As noted earlier, outliers also can provide important clues. In the outbreak of thyrotoxicosis presented in Figure 6.8, most cases came from Luverne, Minnesota, and the surrounding areas. Only one case was identified in Sioux Falls, South Dakota, 60 miles away. Did this person ever go to Luverne? *Yes.* Was she a friend or acquaintance of any of the Luverne cases? *Not really.* What does she do when she goes to Luverne? *Visit my father and buy the locally-produced ground beef that he sells in his store.* Aha! The hypothesis that the locally-produced ground beef was the vehicle could easily be tested by asking cases and noncases whether they ate ground beef from the same source. Cases did, noncases didn't (13).

Step 7: Evaluating Hypotheses

The step after developing hypotheses to explain an outbreak is evaluating the credibility of those hypotheses. In a field investigation, you can evaluate hypotheses in one of two ways: either by comparing the hypotheses with the established facts, or by using analytic epidemiology to quantify relationshipsand explore the role of chance.

You would use the first method when the clinical, laboratory, environmental, and/or epidemiologic evidence so obviously supports the hypotheses that formal hypothesis testing is unnecessarily. For example, in an outbreak of hypervitaminosis D that occurred in Massachusetts

in 1991 it was found that all of the case-patients drank milk delivered to their homes by a local dairy. Therefore, investigators hypothesized that the dairy was the source and the milk was the vehicle. When they visited the dairy, they quickly recognized that the dairy was inadvertently adding far more than the recommended dose of vitamin D to the milk. No analytic epidemiology was really necessary to evaluate the basic hypotheses in this setting (CDC, unpublished data, 1991).

In many other settings, however, the circumstances are not as straightforward. In those instances, you should use analytic epidemiology to test your hypotheses. The key feature of **analytic epidemiology** is a comparison group. With a comparison group, you are able to quantify relationships between exposures and disease, and to test hypotheses about causal relationships. Careful analysis of the series of cases is insufficient for these purposes; a comparison group is essential. You can use comparison groups in two types of studies: cohort and case-control.

Cohort studies

A cohort study is the best technique for an outbreak in a small, well-defined population. For example, you would use a cohort study if an outbreak of gastroenteritis occurred among persons who attended a wedding and a complete list of wedding guests was available.

In this situation, you would contact each attendee and ask a series of questions. You would determine not only whether the attendee had become ill (and met whatever case definition you had developed), but also what foods and drinks he/she had consumed. You might even try to quantify how much of each item he/she had consumed.

After collecting similar information from each attendee, you would be able to calculate an attack rate for those who ate a particular item and an attack rate for those who did not eat that item. Generally, you should look for three characteristics:

1. The attack rate is high among those exposed to the item
2. The attack rate is low among those not exposed, so the difference or ratio between attack rates is high
3. Most of the cases were exposed, so that the exposure could “explain” most, if not all, of the cases

You could, in addition, compute the ratio of these attack rates. Such a ratio is called a **relative risk**, and is a measure of the association between exposure (the food item) and disease. You could also compute a chi-square or other test of statistical significance to determine the likelihood of finding an association as large or larger on the basis of chance alone.

Table 6.3, which is based on a famous outbreak of gastroenteritis following a church supper in Oswego, New York in 1940, illustrates the use of a cohort study in an outbreak investigation (12). Of 80 persons who attended the supper, 75 were interviewed. Forty-six persons met the case definition. Attack rates for those who did and did not eat each of 14 items are presented in Table 6.3.

Table 6.3
Attack rates by items served at the church supper, Oswego, New York, April 1940

	Number of persons who ate specified item				Number of persons who did not eat specified item			
	Ill	Well	Total	Attack Rate (%)	Ill	Well	Total	Attack Rate (%)
Baked ham	29	17	46	63	17	12	29	59
Spinach	26	17	43	60	20	12	32	62
Mashed Potato*	23	14	37	62	23	14	37	62
Cabbage salad	18	10	28	64	28	19	47	60
Jello	16	7	23	70	30	22	52	58
Rolls	21	16	37	57	25	13	38	66
Brown bread	18	9	27	67	28	20	48	58
Milk	2	2	4	50	44	27	71	62
Coffee	19	12	31	61	27	17	44	61
Water	13	11	24	54	33	18	51	65
Cakes	27	13	40	67	19	16	35	54
Ice cream (van.)	43	11	54	80	3	18	21	14
Ice cream (choc.)*	25	22	47	53	20	7	27	74
Fruit salad	4	2	6	67	42	27	69	61

*Excludes 1 person with indefinite history of consumption of that food.
Source: 12

Scan the column of attack rates among those who ate the specified items. Which item shows the highest attack rate? Were most of the 46 cases exposed to that food item? Is the attack rate low among persons not exposed to that item?

You should have identified vanilla ice cream as the implicated vehicle. The data for an individual item are often presented in a two-by-two table. The following two-by-two table shows the data on vanilla ice cream.

Table 6.4
Attack rate by consumption of vanilla ice cream, Oswego, New York, April 1940

		Ill	Well	Total	Attack Rate (%)
Ate vanilla ice cream?	Yes	43	11	54	79.6
	No	3	18	21	14.3
	Total	46	29	75	61.3

The relative risk is calculated as 79.6 / 14.3, or 5.6. The relative risk indicates that persons who ate the vanilla ice cream were 5.6 times more likely to become ill than those who did not eat the vanilla ice cream. Sometimes, attack rate tables such as Table 6.3 include an additional column on the far right for relative risks.

Statistical significance testing. We use tests of statistical significance to determine how likely it is that our results could have occurred by chance alone, if exposure was not actually related to disease. We are not able to cover this broad topic in detail in this course. Instead, we will present only the key features and formulas. For more information, we suggest that you consult one of the many statistics texts that cover the subject well.

The first step in testing for statistical significance is to assume that the exposure is not related to disease. This assumption is known as the **null hypothesis**. (The **alternative hypothesis**, which may be adopted if the null hypothesis proves to be implausible, is that exposure is associated with disease.) Next, you should compute a measure of association, such as a relative risk or odds ratio. Then, you calculate a chi-square or other statistical test. This test tells you the probability of finding an association as strong as, or stronger than, the one you have observed if the null hypothesis is really true. This probability is called the **p-value**. A very small p-value means that you are very unlikely to observe such an association if the null hypothesis is true. If you find a p-value smaller than some cutoff that you have decided on in advance, such as 5%, you may discard or reject the null hypothesis in favor of the alternative hypothesis.

Recall the notation of the two-by-two table described in Lesson 4:

Table 6.5
Standard notation of a two-by-two table

	Ill	Well	Total
Exposed	a	b	H1
Unexposed	c	d	H2
Total	V1	V2	T

The most common statistical test in the outbreak setting is the chi-square test. For a two-by-two table, the chi-square formula is:

$$\text{Chi-square} = \frac{T[|ad - bc| - (T/2)]^2}{V1 \times V2 \times H1 \times H2}$$

Once you have a value for chi-square, you look up its corresponding p-value in a table of chi-squares, such as Table 6.6. Since a two-by-two table has 1 degree of freedom, a chi-square larger than 3.84 corresponds to a p-value smaller than 0.05. This means that if you have planned to reject the null hypothesis when the p-value is less than 0.05, you can do so if your value for chi-square is greater than 3.84.

Table 6.6
Table of Chi Squares

Degree of Freedom	Probability						
	.50	.20	.10	.05	.02	.01	.001
1	.455	1.642	2.706	3.841	5.412	6.635	10.827
2	1.386	3.219	4.605	5.991	7.824	9.210	13.815
3	2.366	4.642	6.251	7.815	9.837	11.345	16.268
4	3.357	5.989	7.779	9.488	11.668	13.277	18.465
5	4.351	7.289	9.236	11.070	13.388	15.086	20.517
10	9.342	13.442	15.987	18.307	21.161	23.209	29.588
15	14.339	19.311	22.307	24.996	28.259	30.578	37.697
20	19.337	25.038	28.412	31.410	35.020	37.566	43.315
25	24.337	30.675	34.382	37.652	41.566	44.314	52.620
30	29.336	36.250	40.256	43.773	47.962	50.892	59.703

The chi-square test works well if the number of people in the study is greater than about 30. For smaller studies, a test called the **Fisher Exact Test** may be more appropriate. Again, we refer you to any statistics book for further discussion of this topic.

Case-control studies

In many outbreak settings, the population is not well defined. Therefore, cohort studies are not feasible. However, since cases have been identified in an earlier step of the investigation, the case-control study is ideal. Indeed, case-control studies are more common than cohort studies in the investigation of an outbreak.

As we discussed in Lesson 1, in a case-control study you ask both case-patients and a comparison group of persons without disease ("controls") about their exposures. You then compute a measure of association–an **odds ratio**–to quantify the relationship between exposure and disease. Finally, as in a cohort study, you can compute a chi-square or other test of statistical significance to determine your likelihood of finding this relationship by chance alone.

This method, while not *proving* that a particular exposure caused disease, certainly has served epidemiologists well over time in implicating sources and vehicles associated with disease, and leading them to appropriate control and prevention measures.

Choosing controls. When you design a case-control study, your first, and perhaps most important, decision is who the controls should be. Conceptually, the controls must not have the disease in question, but should represent the population that the cases come from. In other words, they should be similar to the cases except that they don't have the disease. If the null hypothesis were true, the controls would provide us with the level of exposure that you should expect to find among the cases. If exposure is much higher among the cases than the controls, you might choose to reject the null hypothesis in favor of a hypothesis that says exposure is associated with disease.

In practice, it is sometimes difficult to know who the controls should be. Precisely what is the population that the cases came from? In addition, we must consider practical matters, such as how to contact potential controls, gain their cooperation, ensure that they are free of disease, and get appropriate exposure data from them. In a community outbreak, a random sample of the healthy population may, in theory, be the best control group. In practice, however, persons in a random sample may be difficult to contact and enroll. Nonetheless, many investigators attempt to enroll such "population-based" controls through dialing of random telephone numbers in the community or through a household survey.

Other common control groups consist of:

- neighbors of cases
- patients from the same physician practice or hospital who do not have the disease in question
- friends of cases

While controls from these groups may be more likely to participate in the study than randomly identified population-based controls, they may not be as representative of the population. These **biases** in the control group can distort the data in either direction, masking an association between the exposure and disease, or producing a spurious association between an innocent exposure and disease.

In designing a case-control study, you must consider a variety of other issues about controls, including how many to use. Sample size formulas are widely available to help you make this decision. In general, the more subjects (cases and controls) you use in a study, the easier it will be to find an association.

Often, the number of cases you can use will be limited by the size of the outbreak. For example, in a hospital, 4 or 5 cases may constitute an outbreak. Fortunately, the number of potential controls will usually be more than you need. In an outbreak of 50 or more cases, 1 control per case will usually suffice. In smaller outbreaks, you might use 2, 3, or 4 controls per case. More than 4 controls per case will rarely be worth your effort.

As an example, consider again the outbreak of Legionnaires' disease which occurred in Louisiana. Twenty-seven cases were enrolled in a case-control study. The investigators enrolled 2 controls per case, or a total of 54 controls. Using descriptive epidemiology, the investigators did not see any connection with the town's various cooling towers. Using analytic epidemiology, the investigators determined quantitatively that cases and controls were about equally exposed to cooling towers. However, cases were far more likely to shop at Grocery Store A, as shown in the following two-by-two table (6).

Table 6.7
Exposure to Grocery Store A among cases and controls, Legionellosis outbreak, Louisiana, 1990

		Cases	Controls	Total
Shopped at Grocery Store A?	Yes	25	28	53
	No	2	26	28
	Total	27	54	81

In a case-control study, we are unable to calculate attack rates, since we do not know the total number of people in the community who did and did not shop at Grocery Store A. Since we cannot calculate attack rates, we cannot calculate a relative risk. The measure of association of choice in a case-control study is the **odds ratio**. Fortunately, for a rare disease such as legionellosis or most other diseases which cause occasional outbreaks, the odds ratio approximately equals the relative risk we would have found if we had been able to conduct a cohort study.

The odds ratio is calculated as ad / bc. The odds ratio for Grocery Store A is thus 25 x 26 / 28 x 2, or 11.6. These data indicate that persons exposed to Grocery Store A were 11.6 times more likely to develop Legionnaires' disease than persons not exposed to that store!

To test the statistical significance of this finding, we can compute a chi-square test using the following formula:

$$\text{Chi-square} = \frac{T\left[|ad - bc| - (T/2)\right]^2}{V1 \times V2 \times H1 \times H2}$$

For Grocery Store A, the chi-square becomes:

$$= \frac{81 \times \left[|25 \times 26 - 28 \times 2| - 81/2\right]^2}{27 \times 54 \times 53 \times 28}$$

$$= 24{,}815{,}342.25 / 2{,}163{,}672$$

$$= 11.47$$

Referring to Table 6.6, a chi-square of 11.47 corresponds to a p-value less than 0.001. A p-value this small indicates that the null hypothesis is highly improbable, and the investigators rejected the null hypothesis.

Exercise 6.5

You are called to help investigate a cluster of 17 men who developed leukemia in a community. Some of them worked as electrical repair men, and others were ham radio operators. Which study design would you choose to investigate a possible association between exposure to electromagnetic fields and leukemia?

Answers on page 401.

Exercise 6.6

To study rash illness among grocery store workers, investigators conducted a cohort study. The following table shows the data for exposure to celery. What is the appropriate measure of association? Calculate this measure and a chi-square test of statistical significance.

		Rash	No rash	Total	Attack Rate (%)
Exposed to celery?	Yes	25	31	56	44.64
	No	5	65	70	7.14
	Total	30	96	126	23.81

How would you interpret your results?

Answer on page 401.

Step 8: Refining Hypotheses and Executing Additional Studies

Epidemiologic studies

Unfortunately, analytic studies sometimes are unrevealing. This is particularly true if the hypotheses were not well founded at the outset. It is an axiom of field epidemiology that if you cannot generate good hypotheses (by talking to some cases or local staff and examining the descriptive epidemiology and outliers), then proceeding to analytic epidemiology, such as a case-control study, is likely to be a waste of time.

When analytic epidemiology is unrevealing, you need to reconsider your hypotheses. This is the time to convene a meeting of the case-patients to look for common links and to visit their homes to look at the products on their shelves. Consider new vehicles or modes of transmission.

An investigation of an outbreak of *Salmonella muenchen* in Ohio illustrates how a reexamination of hypotheses can be productive. In that investigation, a case-control study failed to implicate any plausible food source as a common vehicle. Interestingly, *all* case-households, but only 41% of control households, included persons 15 to 35 years. The investigators thus began to consider vehicles of transmission to which young adults were commonly exposed. By asking about drug use in a second case-control study, the investigators implicated marijuana as the likely vehicle. Laboratory analysts subsequently isolated the outbreak strain of *S. muenchen* from several samples of marijuana provided by case-patients (24).

Even when your analytic study identifies an association between an exposure and disease, you often will need to refine your hypotheses. Sometimes you will need to obtain more specific exposure histories. For example, in the investigation of Legionnaires' disease (page 380), what about Grocery Store A linked it to disease? The investigators asked cases and controls how much time they spent in the store, and where they went in the store. Using the epidemiologic data, the investigators were able to implicate the ultrasonic mist machine that sprayed the fruits and vegetables. This association was confirmed in the laboratory, where the outbreak subtype of the Legionnaires' disease bacillus was isolated from the water in the mist machine's reservoir (6).

Sometimes you will need a more specific control group to test a more specific hypothesis. For example, in many hospital outbreaks, investigators use an initial study to narrow their focus. They then conduct a second study, with more closely matched controls, to identify a more specific exposure or vehicle. In a large community outbreak of botulism in Illinois, investigators used three sequential case-control studies to identify the vehicle. In the first study, investigators compared exposures of cases and controls from the general public to implicate a restaurant. In a second study they compared restaurant exposures of cases and healthy restaurant patrons to identify a specific menu item, a meat and cheese sandwich. In a third study, investigators used radio broadcast appeals to identify healthy restaurant patrons who had eaten the implicated sandwich. Compared to cases who had also eaten the sandwich, controls were more likely to have avoided the onions that came with the sandwich. Type A Clostridium botulinum was then identified from a pan of leftover sauteed onions used only to make that particular sandwich (17).

Finally, recall that one reason to investigate outbreaks is research, that is, to expand our knowledge. An outbreak may provide an "experiment of nature," which would be unethical for us to set up deliberately, but which we can learn from when it occurs naturally. For example, in the previously described outbreak of hypervitaminosis D in Massachusetts, investigators quickly traced the source to a dairy that was adding too much vitamin D to its milk. After they had instituted the appropriate control measures, the investigators used the "experiment of nature" to characterize the spectrum of health effects caused by overexposure to vitamin D (CDC, unpublished data, 1991). Thus the investigation led to increased knowledge about an unusual problem as well as to prompt action to remove the source.

When an outbreak occurs, whether it is routine or unusual, consider what questions remain unanswered about that particular disease and what kind of study you might do in this setting to answer some of those questions. The circumstances may allow you to learn more about the disease, its modes of transmission, the characteristics of the agent, host factors, and the like. For example, an outbreak of mumps in a highly immunized population may be an opportunity to study vaccine efficacy and duration of protection.

Laboratory and environmental studies

While epidemiology can implicate vehicles and guide appropriate public health action, laboratory evidence can clinch the findings. The laboratory was essential in both the outbreak of salmonellosis linked to marijuana and in the Legionellosis outbreak traced to the grocery store mist machine. You may recall that the investigation of Legionnaires' disease in Philadelphia in 1976 was not considered complete until the new organism was isolated in the laboratory some 6 months later (10).

Environmental studies are equally important in some settings. They are often helpful in explaining **why** an outbreak occurred. For example, in the investigation of the outbreak of shigellosis among swimmers in the Mississippi (Figure 6.7), the local sewage plant was identified as the cause of the outbreak (20). In the study of thyrotoxicosis described earlier, a review of the procedures used in a slaughterhouse near Luverne, Minnesota, identified a practice that caused pieces of the animals' thyroid gland to be included with beef (13). Use a camera to photograph working conditions or environmental conditions. Bring back physical evidence to be analyzed in the laboratory, such as the slabs of beef from the slaughterhouse in the thyrotoxicosis study or the mist machine from the grocery store in the Legionellosis outbreak investigation.

Step 9: Implementing Control and Prevention Measures

In most outbreak investigations, your primary goal will be control and prevention. Indeed, although we are discussing them as Step 9, you should implement control measures as soon as possible. You can usually implement control measures early if you know the source of an outbreak. In general, you aim control measures at the weak link or links in the chain of infection. You might aim control measures at the specific agent, source, or reservoir. For example, an outbreak might be controlled by destroying contaminated foods, sterilizing contaminated water,

or destroying mosquito breeding sites. Or an infectious food handler could be removed from the job and treated.

In other situations, you might direct control measures at interrupting transmission or exposure. You could have nursing home residents with a particular infection "cohorted," put together in a separate area to prevent transmission to others. You could instruct persons wishing to reduce their risk of acquiring Lyme disease to avoid wooded areas or to wear insect repellent and protective clothing.

Finally, in some outbreaks, you would direct control measures at reducing the susceptibility of the host. Two such examples are immunization against rubella and malaria chemoprophylaxis for travelers.

Step 10: Communicating the Findings

Your final task in an investigation is to communicate your findings. This communication usually takes two forms: (1) an oral briefing for local authorities and (2) a written report.

Your oral briefing should be attended by the local health authorities and persons responsible for implementing control and prevention measures. Usually these persons are not epidemiologists, so you must present your findings in clear and convincing fashion with appropriate and justifiable recommendations for action. This presentation is an opportunity for you to describe what you did, what you found, and what you think should be done about it. You should present your findings in scientifically objective fashion, and you should be able to defend your conclusions and recommendations.

You should also provide a written report that follows the usual scientific format of introduction, background, methods, results, discussion, and recommendations. By formally presenting recommendations, the report provides a blueprint for action. It also serves as a record of performance and a document for potential legal issues. It serves as a reference if the health department encounters a similar situation in the future. Finally, a report that finds its way into the public health literature serves the broader purpose of contributing to the knowledge base of epidemiology and public health.

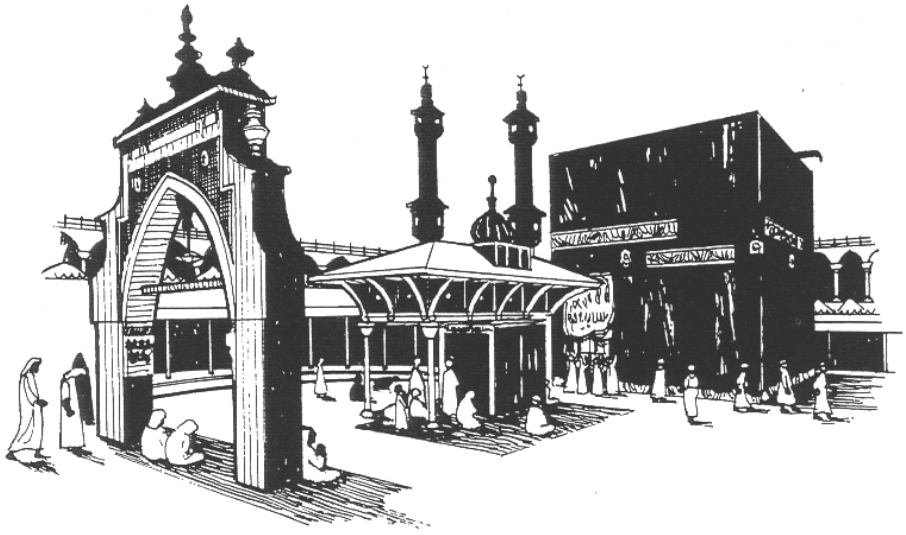
Review Exercise

Exercise 6.7

This review exercise is a case study of an outbreak of enteritis during a pilgrimage to Mecca. After reading this case study and answering all 16 imbedded questions, a student will be able to do the following:

- Define an epidemic, an outbreak, and a cluster
- Create and understand the uses of a case definition
- Draw an epidemic curve
- Calculate food-specific attack rates
- List the steps in investigating an acute outbreak

Figure 6.9
Illustration of the Kaaba in Mecca



An Outbreak of Enteritis During a Pilgrimage to Mecca

Part I

On the morning of November 1, 1979, during a pilgrimage to Mecca, the epidemiologist assigned to the Kuwaiti medical mission experienced acute onset of abdominal cramps and diarrhea at the holy mosque before the walk around the Kaaba. He subsequently learned that other members of the mission had developed similar symptoms. When he returned that evening to Muna, he initiated an investigation.

Question 1. What information do you need to decide if this is an epidemic?

The epidemiologist interviewed several ill members of the mission to better characterize the illness. On the basis of these interviews, the epidemiologist quickly prepared a questionnaire and conducted interviews with the 112 members of the Kuwaiti medical mission.

A total of 66 cases of illness were identified; 2 had onset in Kuwait prior to the beginning of the pilgrimage and 64 had onset of symptoms beginning late in the afternoon on October 31.

Question 2. Is this an epidemic? Explain your answer.

Description of the Pilgrimage

The Kuwaiti medical mission, consisting of 112 members, traveled by automobile from Kuwait to Mecca. On October 30 all members of the mission slept in Muna. At sunrise on October 31 they traveled to Arafat, where at 8:00 a.m. they had tea with or without milk for breakfast. The milk was prepared immediately before consumption by mixing powdered milk with boiled water. The remainder of the day was devoted to religious services. At 2:00 p.m., a lunch was served for all members of the mission who wished to partake. It was a typical Kuwaiti meal consisting of three dishes: rice, meat, and tomato sauce. Most individuals consumed all three dishes. The lunch had been prepared in Muna on October 30 and transported to Arafat by truck early on October 31. At sunset on October 31 the mission members returned to Muna.

Clinical Description

The investigator identified a total of 66 cases of gastroenteritis. The onset of all cases was acute, characterized chiefly by diarrhea and abdominal pain. Nausea, vomiting, and blood in the stool occurred infrequently. No case-patient reported fever. All recovered within 12-24 hours. Approximately 20 percent of the ill individuals sought medical advice. The investigator did not obtain any fecal specimens for examination.

Question 3. Develop a preliminary case definition.

Question 4. List the broad categories of diseases that must be considered in the differential diagnosis of an outbreak of gastrointestinal illness.

Note: These concepts have not been covered in this course. If you are not familiar with disease agents, review the answer to this question.

Question 5. What clinical and epidemiologic information might be helpful in determining the etiologic agent(s)?

Question 6. The Kuwaiti investigators distributed a questionnaire to all members of the mission. What information would you solicit on this questionnaire?

Part II

Investigators determined that of the 64 cases with onset during the pilgrimage, all had eaten lunch in Arafat at 2:00 p.m. on October 31. Fifteen members of the mission did not eat lunch; none became ill.

Question 7. Calculate the attack rate for those who ate lunch and those who did not. What do you conclude?

Table 6.8 (page 394-395) presents some of the information collected by the investigators. The two members who developed illness prior to October 31 have been excluded. The 15 members of the mission who did not eat lunch are not included in Table 6.8.

Question 8. Using appropriate time periods, draw an epidemic curve.

Question 9. Are there any cases for which the time of onset seems inconsistent? How might they be explained?

Question 10. Modify the graph you have drawn (Question 8) to illustrate the distribution of incubation periods.

Question 11. Determine or calculate the minimum, maximum, mean, median, mode, range, and standard deviation of the incubation periods.

Question 12a. Calculate the frequency of each clinical symptom among the cases.

Question 12b. How does the information on the symptoms and incubation periods help you to narrow the differential diagnosis? (You may refer to the attached "Abbreviated Compendium of Acute Foodborne Gastrointestinal Diseases" in Appendix E).

Question 13a. Using the food consumption histories in Table 6.8, complete item 7 of the "Investigation of a Foodborne Outbreak" report form in Appendix F.

Question 13b. Do these calculations help you to determine which food(s) served at the lunch may have been responsible for the outbreak?

Question 14. Outline further investigations which should be pursued. List one or more factors that could have led to the contamination of the implicated food.

Table 6.8
Selected characteristics of Kuwaiti medical mission members who ate lunch at Arafat, Saudi Arabia, October 31, 1979

			Onset of Illness		Foods			Signs and Symptoms*					
Id #	Age	Sex	Date	Hour	Rice	Meat	TS*	D	C	BS	N	V	F
31	36	M	Oct. 31	5 p.m.	X	X	X	D	C	BS			
77	28	M	Oct. 31	5 p.m.	X	X		D	C				
81	33	M	Oct. 31	10 p.m.	X	X	X	D	C				
86	29	M	Oct. 31	10 p.m.	X	X	X	D	C				
15	38	M	Oct. 31	10 p.m.		X		D		BS	N		
17	48	M	Oct. 31	10 p.m.	X	X		D	C				
18	35	M	Oct. 31	10 p.m.	X	X	X	D	C				
35	30	M	Oct. 31	11 p.m.	X	X	X	D	C				
88	27	M	Oct. 31	11 p.m.	X	X	X	D	C				
76	29	M	Oct. 31	11 p.m.	X	X	X	D	C	BS			
71	50	M	Oct. 31	12 MN	X	X	X	D					
1	39	F	Nov. 1	1 a.m.	X	X	X	D	C			V	
27	36	M	Nov. 1	1 a.m.	X	X	X	D	C		N		
28	44	M	Nov. 1	1 a.m.	X	X	X	D	C				
29	48	M	Nov. 1	1 a.m.	X	X	X	D	C	BS			
30	35	M	Nov. 1	2 a.m.	X	X	X	D	C				
50	29	M	Nov. 1	2 a.m.	X	X	X	D	C				
59	51	M	Nov. 1	2 a.m.	X	X	X	D	C				
67	40	M	Nov. 1	2 a.m.	X	X		D					
72	58	M	Nov. 1	2 a.m.	X	X	X	D	C				
73	28	M	Nov. 1	3 a.m.	X	X	X	D	C				
60	31	M	Nov. 1	3 a.m.	X	X	X	D	C				
61	38	M	Nov. 1	3 a.m.	X	X	X	D		BS			
51	32	M	Nov. 1	3 a.m.	X	X	X	D	C			V	
52	37	M	Nov. 1	3 a.m.	X	X		D					
58	30	M	Nov. 1	3 a.m.	X	X	X	D	C				
22	35	M	Nov. 1	3 a.m.	X	X	X	D	C				
25	30	M	Nov. 1	3 a.m.	X	X		D	C				
32	50	M	Nov. 1	3 a.m.	X	X	X	D	C				
38	26	M	Nov. 1	3 a.m.	X	X	X	D	C				
79	29	M	Nov. 1	3 a.m.	X	X	X	D	C				
80	28	M	Nov. 1	3 a.m.	X	X	X	D	C				
37	30	M	Nov. 1	4 a.m.	X	X	X	D					
65	34	M	Nov. 1	4 a.m.	X	X		D		BS			
66	45	M	Nov. 1	4 a.m.	X	X		D	C				
87	41	M	Nov. 1	4 a.m.	X	X	X	D	C				
89	43	M	Nov. 1	4 a.m.	X	X	X	D	C				
90	43	M	Nov. 1	4 a.m.	X	X	X	D	C				
91	38	M	Nov. 1	4 a.m.	X	X	X	D	C				
92	37	M	Nov. 1	4 a.m.	X	X	X	D	C				
70	31	M	Nov. 1	5 a.m.	X	X	X	D	C				
2	34	F	Nov. 1	5 a.m.	X	X	X	D	C				
21	38	M	Nov. 1	5 a.m.	X	X	X	D	C				
40	38	M	Nov. 1	5 a.m.	X	X	X	D					
78	27	M	Nov. 1	5 a.m.	X	X	X	D	C				
82	39	M	Nov. 1	5 a.m.	X	X	X	D	C				
83	40	M	Nov. 1	5 a.m.	X	X	X	D	C				

*TS = Tomato sauce, D = diarrhea, C = cramps, BS= blood in stool, N= nausea, V= vomiting, F = fever

Table 6.8 (continued)
Selected characteristics of Kuwaiti medical mission members who ate lunch at Arafat, Saudi Arabia, October 31, 1979

			Onset of Illness		Foods			Signs/Symptoms					
Id #	Age	Sex	Date	Hour	Rice	Meat	TS*	D	C	BS	N	V	F
84	34	M	Nov. 1	5 a.m.	X	X		D	C				
14	52	M	Nov. 1	6 a.m.	X	X	X	D					
16	40	M	Nov. 1	6 a.m.	X	X	X	D		BS			
93	30	M	Nov. 1	6 a.m.	X	X	X	D	C				
94	39	M	Nov. 1	6 a.m.	X	X	X	D	C				
33	55	M	Nov. 1	7 a.m.	X	X	X	D	C				
34	28	M	Nov. 1	7 a.m.	X	X	X	D	C				
85	38	M	Nov. 1	7 a.m.	X	X		D	C				
43	38	M	Nov. 1	9 a.m.	X	X		D	C				
69	30	M	Nov. 1	9 a.m.	X	X	X	D	C				
4	30	F	Nov. 1	10 a.m.	X			D	C				
5	45	F	Nov. 1	10 a.m.		X			C				
3	29	F	Nov. 1	1 p.m.	X	X		D	C				
12	22	F	Nov. 1	2 p.m.	X	X	X		C				
74	44	M	Nov. 1	2 p.m.	X	X	X	D					
75	45	M	Nov. 1	5 p.m.	X	X	X	D		BS			
95	40	M	Nov. 1	11 p.m.	X	X	X	D	C				
6	38	F	WELL		X	X							
7	52	F	WELL		X	X	X						
8	35	F	WELL		X		X						
9	27	F	WELL		X	X	X						
10	40	F	WELL		X	X	X						
11	40	F	WELL		X	X	X						
13	50	M	WELL		X	X	X						
19	38	M	WELL		X	X	X						
20	38	M	WELL		X	X	X						
23	29	M	WELL		X	X	X						
24	27	M	WELL		X	X	X						
26	47	M	WELL		X	X	X						
36	60	M	WELL		X								
39	27	M	WELL		X	X	X						
41	30	M	WELL		X	X	X						
42	38	M	WELL		X	X	X						
44	50	M	WELL		X	X	X						
45	27	M	WELL		X	X	X						
46	31	M	WELL		X	X	X						
47	46	M	WELL		X	X	X						
48	38	M	WELL		X	X	X						
49	36	M	WELL		X								
53	36	M	WELL		X	X	X						
54	27	M	WELL		X	X	X						
55	40	M	WELL		X	X	X						
56	30	M	WELL		X	X	X						
57	25	M	WELL		X	X	X						
62	50	M	WELL		X								
63	44	M	WELL		X								
64	47	M	WELL		X		X						
68	31	M	WELL		X	X	X						

*TS = Tomato sauce, D = diarrhea, C = cramps, BS = blood in stool, N = nausea, V = vomiting, F = fever

Part III

The lunch which was served in Arafat at 2:00 p.m. on October 31 was prepared at 10:00 p.m. the night before in Muna. It consisted of boiled rice, chunks of lamb fried in oil, and tomato sauce prepared from fresh tomatoes which were sectioned and stewed. The cooked rice was placed in two large pots and the lamb was divided evenly on top. The tomato sauce was kept in a third pot.

These pots were covered with metal tops and placed in an open spot among some rocks near the kitchen and allowed to stand overnight. They were presumably not touched by anyone during this period. Early in the morning on October 31, the pots were transported by truck from Muna to Arafat where they stood in the truck until 2:00 p.m. The temperature in Arafat at noon that day was 35 degrees Centigrade. The food was not refrigerated from the time of preparation to the time of consumption.

Cooks and all other individuals who helped in preparing the meal were intensively interviewed regarding any illness present before or at the time of preparation. All individuals interviewed denied having any illness and knew of no illness among any other members of the group responsible for meal preparation. No specimens were obtained from any of the cooks for laboratory examination.

The following is quoted verbatim from the report prepared by the epidemiologist who investigated the outbreak:

> "This clinical picture probably suggests an infection by *Clostridium perfringens*. This organism could be detected in the food elements consumed as well as in the patient's stool. However, no laboratory diagnostic procedures were possible in the outbreak site. All the investigations conducted were based entirely on epidemiologic grounds.
>
> The incubation period as well as other data extrapolated from epidemiological analysis suggests that *Clostridium perfringens* is the causative agent. This organism is widely distributed in nature especially in soil and dust. So there is ample opportunity for contamination of the food. If cooked meat is allowed to cool slowly under suitable anaerobic conditions, spores which might have survived cooking or have subsequently come from dust may germinate and within a few hours produce large numbers of vegetative bacilli. In fact, the pilgrimage camp in Muna lacks sanitary cooking facilities. The food is usually prepared in a dusty place open to the blowing winds creating an ideal situation for *Clostridium perfringens* contamination.
>
> The type of the organism, the type of food dish it usually contaminates, its mode of spread and the differences in the attack rates for those who consumed meat and those who did not points to the meat as the probable source of infection in this outbreak.
>
> Conclusion: The acute illness of enteritis in Arafat affected many persons in an epidemic form. It was a common-source outbreak, the source being the meat consumed at the Arafat lunch. The incubation period was about 13 hours. The illness was characterized by colicky abdominal pain and diarrhea with no elevation of temperature. The responsible

agent for this outbreak is most probably *Clostridium perfringens*.

The lunch at Arafat should have been prepared in the same day of consumption, or kept refrigerated if it had to be prepared the day before. Although kitchens could not be fully equipped to fulfill the essential safety measures in a place like Muna, they should be supplied by essential measures to protect food from contamination. The remaining food in Arafat should have been condemned after the investigation, but none remained at that time.

The epidemiological investigations carried out in this epidemic could explore the nature of this epidemic and answer most of the questions raised. The laboratory investigation, although helpful to detect the causative organisms, should not replace the more efficient epidemiological methods in the exploration of such epidemics. The lack of the necessary laboratory facilities to detect the causative organisms in foodborne outbreaks should not discourage the investigative epidemiologist and make him doubtful and lose confidence in his epidemiological tools."

Question 15. In the context of this outbreak, what control measures would you recommend?

Question 16. Was it important to work up this outbreak?

Answers to Exercises

Answer–Exercise 6.1 (page 352)

One reason to investigate is simply **to determine how many cases we would expect in the community**. In a large community, nine cases of a common cancer (for example, lung, breast, or colon cancer) would not be unusual. In a very small community, nine cases of even a common cancer may seem unusual. If the particular cancer is a rare type, then nine cases even in a large community may be unusual.

If the number of cancer cases turns out to be high for that community, we might pursue the investigation further. We may have a **research** motive–perhaps we will identify a new risk factor (workers exposed to a particular chemical) or predisposition (persons with a particular genetic marker) for the cancer.

Control and prevention may be a justification. If we find a risk factor, control / prevention measures could be developed. Alternatively, if the cancer is one which is generally treatable if found early, and a screening test is available, then we might investigate to determine not why these persons developed the disease, but why they died of it. If the cancer were cancer of the cervix, detectable by Pap smear and generally treatable if caught early, we might find (1) problems with access to health care, or (2) physicians not following the recommendations to screen women at the appropriate intervals, or (3) laboratory error in reading or reporting the test results. We could then develop measures to correct the problems we found (public screening clinics, education of physicians, or laboratory quality assurance.)

If new staff need to gain experience on a cluster investigation, **training** may be a reason to investigate. More commonly, cancer clusters frequently generate **public concern**, which, in turn, may generate **political pressure**. Perhaps one of the affected persons is a member of the mayor's family. A health department must be responsive to such concerns, but does not usually need to conduct a full-blown investigation. Finally, **legal concerns** may prompt an investigation, especially if a particular site (manufacturer, houses built on an old dump site, etc.) is accused of causing the cancers.

Answer–Exercise 6.2 (page 356)

Tuberculosis does not have a striking seasonal distribution. The number of cases during August could be compared with (a) the numbers reported during the preceding several months, and (b) the numbers reported during August of the preceding few years.

Aseptic meningitis is a highly seasonal disease which peaks during August-September-October. As a result, the number of cases during August is expected to be higher than the numbers reported during the preceding several months.

To determine whether the number of cases reported in August is greater than expected, we must look at the numbers reported during August of the preceding few years.

Answer–Exercise 6.3 (page 362)

Which items to include in a line listing is somewhat arbitrary.
The following categories of information are often included:

Identifying information

- Identification number or case number, usually in the leftmost column
- Names or initials as a cross-check

Information on diagnosis and clinical illness

- Physician diagnosis
- Was diagnosis confirmed? If so, how?
- Symptoms
- Laboratory results
- Was the patient hospitalized? Did the patient die?

Descriptive epidemiology–time

- Date of onset
- Time of onset

Descriptive epidemiology–person

- Age
- Sex
- Occupation, if relevant, or other seemingly relevant characteristics

Descriptive epidemiology–place

- Street, city, or county
- Worksite, school, day care center, etc., if relevant

Risk factors and possible causes

- Specific to disease and outbreak setting

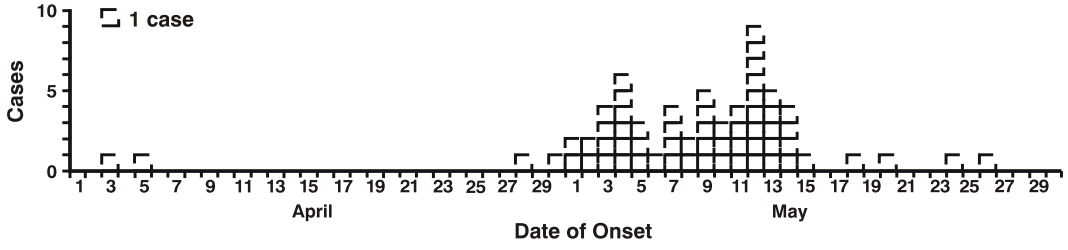
An example of a line listing from the six case report forms is shown below.

ID #	Initials	Dateof Onset	Diagnosis	How Confirmed	Age	Sex	County	Physician	Cleveland-McKay Wedding
1	KR	7/23	Probable Trichinosis	Not done	29	M	Columbia	Goodman	Yes
2	DM	7/27	Trichinosis	Biopsy	33	M	Columbia	Baker	Yes
3	JG	8/14	Probable Trichinosis	Not done	26	M	Columbia	Gibbs	Yes
4	RD	7/25	Trichinosis	Serologic	45	M	King	Webster	Yes
5	NT	8/4	Trichinosis	Not done	27	F	Columbia	Stanley	Yes
6	AM	8/11	R/O trichinosis	pending	54	F	Clayton	Mason	Yes

Answer–Exercise 6.4 (page 369)

The epidemic curve shown in Figure 6.10 suggests a common source outbreak. We can estimate time of exposure by starting at the peak of the epidemic and going back the mean incubation period, or by starting at the rise of the epidemic and going back the minimum incubation period. Going back 30 days (mean incubation period for hepatitis A) from the epidemic peak on May 9 puts the estimated exposure on April 9. Assuming the minimum incubation period (15 days) for the April 28 case, exposure would have occurred on April 13. So, we can estimate that exposure occurred between April 9 and April 13, give or take a few days on either side.

Figure 6.10
Epidemic curve for Exercise 6.4: Hepatitis A by date of onset, April-May



Answer–Exercise 6.5 (page 382)

A case-control study is the design of choice, since 17 persons with the disease of interest have already been identified. We would need to enroll these 17 persons as the case group. We would also need to determine what group might serve as an appropriate comparison or control group. Neighbors might be used for the control group, for example. In our case-control study we would determine whether each case and each control was exposed to electromagnetic fields (however we defined that exposure). Finally, we would compare the exposure experience of cases and controls.

The alternative to a case-control study is a cohort study. For a cohort study we would have to enroll a group of persons exposed to electromagnetic fields (however we defined that exposure), and a comparison group of persons not exposed. We would then have to determine how many in each group developed leukemia. Since leukemia is a relatively rare event, we would need rather large groups in order to have enough leukemia cases to make our study valid. Therefore, a cohort study is less practical than a case-control study in this setting.

Answer–Exercise 6.6 (page 383)

The appropriate measure of association for a cohort study is the relative risk, calculated as the ratio of attack rates.

Relative risk = 44.64/7.14 = 6.2

Chi-square = $\frac{T[|ad-bc|-(T/2)]^2}{V1\times V2\times H1\times H2}$

For the table shown above, the chi-square becomes:

$$= \frac{126\times[|25\times 65-31\times 5|-126/2]^2}{30\times 96\times 56\times 70}$$

$$= 249{,}435{,}774/11{,}289{,}600$$

$$= 22.09$$

A chi-square of 22.09 corresponds to a p-value of < 0.00001. A p-value this small indicates that the null hypothesis is highly improbable, and the investigators rejected the null hypothesis.

Answer--Exercise 6.7 (page 387)

An Outbreak of Enteritis During a Pilgrimage to Mecca

Question 1. What information do you need to decide if this is an epidemic?

Answer 1.

- Is the number of cases more than the number expected?
- Therefore, we need to know background rate.

Question 2. Is this an epidemic?

Answer 2. Yes. An epidemic can be defined as the occurrence of more cases in a place and time than expected in the population being studied. Of the 110 members without signs and symptoms of gastroenteritis prior to the pilgrimage, 64 (58%) developed such signs and symptoms during this trip. This is clearly above the expected or background rate of gastroenteritis in most populations. Gastroenteritis prevalence rates from recent surveys are approximately 5% and are consistent with this population (2/112 had such signs and symptoms at the time of the pilgrimage).

One could survey other groups of pilgrims originating from the same country to determine their rates of diarrheal illness if the existence of an outbreak was uncertain. Practically speaking, however, an attack rate of 58% is an epidemic until proven otherwise.

The term “outbreak” and “epidemic” are used interchangeably by most epidemiologists. The term “outbreak” is sometimes preferred, particularly when talking to the press or public, because it is not as frightening as “epidemic.” The term “cluster” may be defined as the occurrence of a group of cases in a circumscribed place and time. In a cluster, the number of cases may or may not be greater than expected.

Question 3. Develop a preliminary case definition.

Answer 3.

Points to consider:

- As a general rule, during the initial phase of an investigation, the case definition should be broad.
- The case definition should include four components: **time, place, person,** and **diagnosis** (or signs, symptoms). Depending on the frequency of the symptoms observed and the probable etiologic agent, a more precise case definition can be developed later.

Case definition:

Clinical:	acute onset of abdominal cramps and/or diarrhea
Time:	onset after noon on October 31 and before November 2
Place/Person:	member of the Kuwaiti medical mission in route to Mecca

Note. The Kuwaiti investigators had already decided that lunch on October 31 was the responsible meal and defined an outbreak-associated case of enteritis as a person in the Kuwaiti mission who ate lunch at Arafat at 2:00 p.m. on October 31 and subsequently developed abdominal pain and/or diarrhea prior to November 2, 1979.

However, at this point in your consideration of the outbreak you have not implicated the lunch, and it would probably be premature to limit your case definition to those who ate lunch.

Question 4. List the broad categories of diseases that must be considered in the differential diagnosis of an outbreak of gastrointestinal illness.

Answer 4.

Broad categories: Bacterial
Viral
Parasitic
Toxins

More specifically:

Differential Diagnosis of Acute Foodborne Enteric Illness

Bacteria and bacterial toxins
*Bacillus cereus**
Campylobacter *jejuni*
Clostridium botulinum (initial symptoms)
*Clostridium perfringens**
*Escherichia coli**
Salmonella, non-typhoid
Salmonella typhi
Shigella
Staphylococcus aureus
Vibrio cholerae O1
Vibrio cholerae non-O1
Vibrio parahemolyticus
Yersinia enterocolitica

Viruses
Norwalk-like agents (i.e., 27 nm viruses)
Rotavirus*

Toxins
Heavy metals (especially cadmium, copper, tin, zinc)
Mushrooms
Fish & shellfish (e.g., scombroid, ciguatera)
Insecticides

Parasites
Cryptosporidium
Entamoeba histolytica
Giardia lamblia

*These agents are most compatible with the following characteristics of this outbreak:

- acute onset
- lower GI signs and symptoms
- no fever
- appreciable proportion seeking medical advice
- no mention of non-enteric (dermatologic, neurologic) manifestations

However, you have not yet reached the point in your investigation to consider the most likely etiologic possibilities for the illness.

Question 5. What clinical and epidemiologic information might be helpful in determining the etiologic agent(s)?

Answer 5.
Incubation period
Symptom complex
Duration of symptoms
Severity of symptoms
Seasonality
Geographic location
Biologic plausibility of pathogens

Question 6. The Kuwaiti investigators distributed a questionnaire to the persons who ate the implicated lunch. What information would you solicit on this questionnaire?

Answer 6.

- **Identifying information**
- **Demographics (age, sex, race)**
- **Clinical information**
  - Symptoms
  - Date & time of onset of symptoms
  - Duration of symptoms
  - Medical intervention, if required
- **Information on possible causes**
  - Exposure information regarding foods consumed, including amounts
  - Other potential exposures
  - Other factors that may modify risk of diarrhea (e.g., antacids, antibiotics)

Question 7. Calculate the attack rate for those who ate lunch and those who did not. What do you conclude?

Answer 7.

112 members of the mission
-15 members who didn't eat lunch
- 2 members sick before pilgrimage
95 *at risk of developing illness*
64 became ill among those who ate lunch
0 *became ill among those who didn't eat lunch*

Attack rate for those who ate lunch:
64 ill/95 at risk = 67%

Attack rate for those not eating lunch:
0 ill/15 at risk = 0%

Conclusion: Lunch is strongly associated with disease.

Question 8. Using appropriate time periods, draw an epidemic curve.

Answer 8.
Points for consideration about epi curves:

1. The epi curve is a basic tool of epidemiologists to
   a. establish existence of epidemic vs. endemic illness
   b. delineate time course and magnitude of an epidemic
   c. develop inferences about transmission, e.g., common source, person to person, intermittent exposure. Note that changing the interval on the *x*-axis can significantly alter the shape of an epi curve.
   d. predict future course of an epidemic: when it will end, that a second wave is underway, that secondary cases are occurring, etc.
2. With common source outbreaks, the width of the curve is determined by the incubation period, varying doses, and host susceptibility.
3. Often a few cases don't fit into the body of an epi curve. Such exceptions may be quite important--as index cases or other special situations.
4. A rule of thumb: When the incubation period is known, the maximum time period on the *x*-axis should not usually exceed 1/4 - 1/3 of the incubation period.

Summary of the temporal distribution (see Figure 6.11a).

a. Onsets of cases occurred over a period of 31 hours extending from 5 p.m. on October 31 to 11 p.m. on November 1.

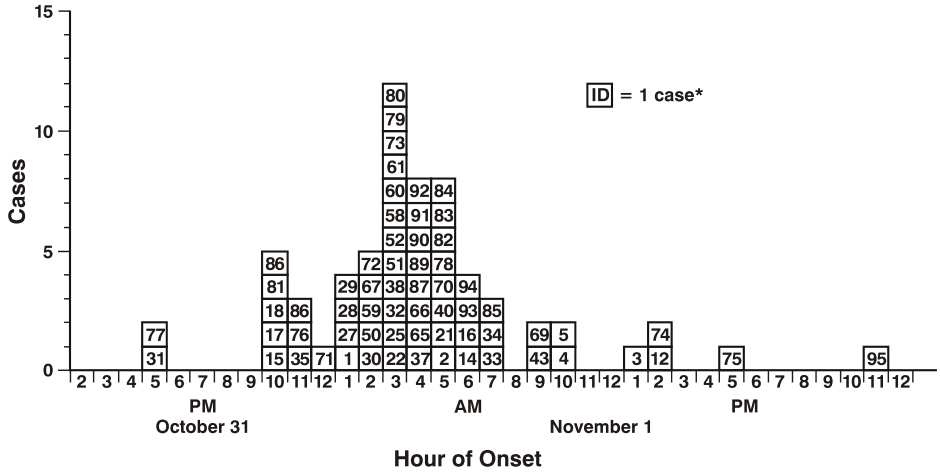
b. Onsets of 53 (82.8%) of the cases occurred throughout the 10 hour interval from 10 p.m. on October 31 through 7 a.m. on November 1.

c. The peak (12 cases) occurred at 3 a.m. on November 1.

d. The median hour of onset = 3:30 a.m. November 1 (actual middle rank = 32.5 which falls between the 3 and 4 a.m. measurement intervals).

e. It is likely that the way the questionnaire was designed forced the interviewees to give a rounded time for onset of symptoms.

Figure 6.11a
Outbreak associated cases of enteritis by hour of onset of illness, Kuwaiti Mission, Arafat, Saudi Arabia, October 31 – November 1, 1979



*ID# included for reference only.

Question 9. Are there any cases for which the times of onset seem inconsistent? How might they be explained?

Answer 9.

1. The two cases (#31 and 77) with onsets at 5 p.m. on October 31
   a. Illnesses unrelated to the outbreak?
   b. Earlier exposures to food items? Cooks?
   c. Short incubation periods? Large doses? Enhanced susceptibility?
   d. Times of onset incorrect?

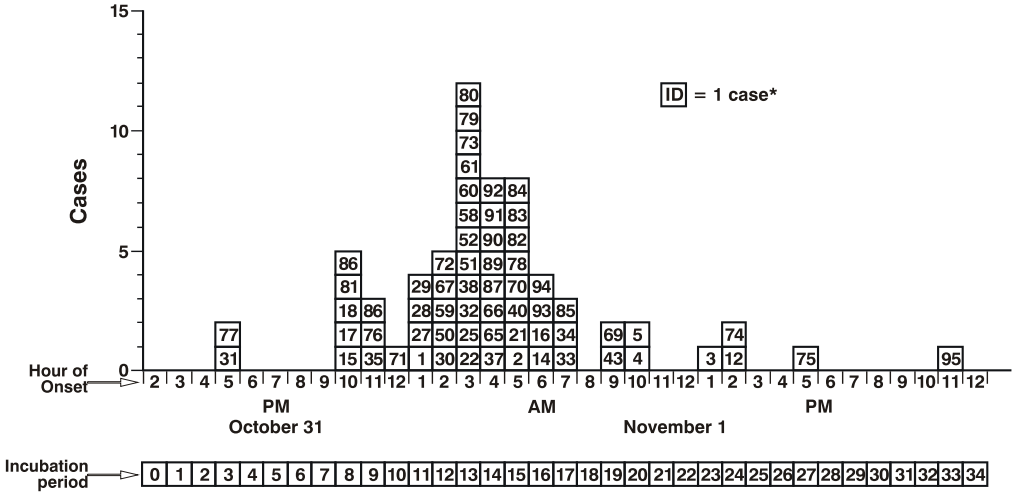
2. The two cases (#75 and 95) occurring late on November 1
   a. Illnesses unrelated to the outbreak?
   b. Foods eaten at later time?
   c. Secondary cases?
   d. Times of onset incorrect?
   e. Long incubation periods? Small doses? Enhanced resistance?

Question 10. Modify the graph you have drawn (Question 8) to illustrate the distribution of incubation periods.

Answer 10.

Since all meal participants were served at 2:00 p.m. the distribution of onsets and incubation periods is the same. Therefore, to illustrate the distribution of incubation periods, you need only to show a second label for the *x*-axis, as in Figure 6.11b.

Figure 6.11b
Outbreak associated cases of enteritis by hour of onset of illness and incubation period, Kuwaiti Mission, Arafat, Saudi Arabia, October 31 – November 1, 1979



Question 11. Determine or calculate the minimum, maximum, mean, median, mode, range, and standard deviation of the incubation periods.

Answer 11.

Minimum = 3 hours

Maximum = 33 hours

Mean = 14 hours

Median = 13.5 hours (middle rank = 64 + 1)/2 = 32.5, which falls between the intervals for 13 and 14 hours.)

Mode = 13 hours

Range = maximum - minimum = 30 hours

Standard deviation = 5 hours

Note: The range in which roughly 95% of the observations fall = $\bar{x}$ ± 1.96 (rounded to 2) standard deviations = 4 to 24 hours (see Lesson 3 for calculation steps).

Comment

The incubation period (though not necessarily the clinical features) are about right for *Clostridium perfringens*, *Salmonella*, *Vibrio parahemolyticus*, and *Bacillus cereus*. The incubation period is a bit short for enterotoxigenic E. Coli and *Vibrio cholerae* non-O1. Too long for staph enterotoxin, heavy metals, chemicals, and most toxins produced by fish, shellfish, and mushrooms. Illnesses that have upper GI signs and symptoms, such as nausea and vomiting, and intoxications due to chemicals, metals, etc., usually have short incubation periods, while illnesses with predominately lower GI signs and symptoms, such as diarrhea, have longer incubation periods.

Question 12a. Calculate the frequency of each clinical symptom among the cases.

Answer 12a.

Frequency distribution of signs and symptoms among outbreak-associated cases of enteritis, Kuwaiti Mission, Arafat, Saudi Arabia, October 31 – November 1, 1979 (N = 64)

Sign or Symptom	Number of Cases	Percent
Diarrhea	62	96.9
Abdominal Pain	52	81.3
(Diarrhea + abdominal pain)	(50)	(78.1)
Blood in stool	8	12.5
(Diarrhea + blood in stool)	(5)	(7.8)
(Diarrhea + abdominal pain + blood in stool	(3)	(4.7)
Nausea	2	3.1
Vomiting	2	3.1
Fever	0	0

The distribution of signs and symptoms are given in the table above. Diarrhea occurred among all but two of the cases, with 78.1% experiencing both diarrhea and abdominal pain.

Blood in the stool was reported by 8 (12.5%) of the cases. Symptoms of upper GI distress occurred among 4 (6.3%) of the cases (2 persons experienced nausea while two others reported vomiting). No temperature elevations were recorded.

Question 12b. How does the information on the symptoms and incubation period help you to narrow the differential diagnosis? (You may refer to the attached compendium in Appendix F, which describes a number of acute foodborne gastrointestinal diseases.)

Answer 12b.

The clinical findings, including an apparent absence of malaise, myalgias, chills, and fever, are more consistent with an intoxication resulting from the presence of toxin in the lower GI tract than with an invasive infectious agent. The recovery of all cases within 24 hours is also consistent with such an intoxication. The absence of dermatologic and neurologic signs and symptoms in conjunction with the incubation period (the median was 13.5 hours and the mean was 14 hours) would lessen the likelihood of heavy metals, organic and inorganic chemicals, and toxins produced by fish, shellfish and mushrooms. The incubation period and clinical features help narrow the list to the following: *Clostridium perfringens*, *Bacillus cereus*, *Vibrio parahemolyticus*, and, less likely, *Vibrio cholerae* non-O1, and enterotoxin producing *E. coli.*

Question 13a. Using the food consumption histories in Table 6.8, complete item 7 of the "Investigation of a Foodborne Outbreak" report form in Appendix F.

Answer 13a.

	# persons who ATE specified food				# who DID NOT EAT specified food			
	Ill	Well	Total	Attack Rate	Ill	Well	Total	Attack Rate
Rice	62	31	93	66.7%	2	0	2	100.0%
Meat	63	25	88	71.6 %	1	6	7	14.3%
T.S.	50	26	76	65.8%	14	5	19	73.7%

You may analyze these data with 2 x 2 tables:

		ILL	WELL	TOTAL	Attack Rate	
Exposed?	Yes	a	b	a + b	AR1 = a/a + b	RR = AR1/AR2
	No	c	d	c + d	AR2 = c/c + d	
		a + c	b + d	T = a + b + c + d		

		ILL	WELL	TOTAL	Attack Rate	
Ate	Yes	62	31	93	62/93 = 66.7%	RR = 66.7/100
Rice	No	2	0	2	2/2 = 100.0%	= 0.67

		ILL	WELL	TOTAL	Attack Rate	
Ate	Yes	63	25	88	63/88 = 71.6%	RR = 72.6/14.3
Meat	No	1	6	7	1/7 = 14.3%	= 5.0
		64	31			

		ILL	WELL	TOTAL	Attack Rate	
Ate	Yes	50	26	76	50/76 = 65.8%	RR = 65.8/73.7
Tomato	No	14	5	19	14/19 = 73.7%	0.89
Sauce		64	31			

Question 13b. Do these calculations help you to determine which food(s) served at the lunch may have been responsible for the outbreak?

Answer 13b. Attack rates were high for those who ate rice, meat, and tomato sauce. However, meat is the likely culprit because it was the only food associated with a high attack rate among those who ate it, but a low attack rate among those who did not. Almost all (63/64) who ate meat also ate the other items, which probably accounts for the high attack rates for those items, too.

One of the cases did not admit to eating meat and could be explained in any number of ways:

- Unrelated illness
- Cross-contamination, e.g., common server, spoon, dish, counter, etc., or from meat to rice
- Reporting error (e.g., forgot or purposely denied eating meat)
- Transcription error (e.g., misrecorded response)

NOTE: Epidemiologic evidence shows an association between exposure and subsequent disease but **does not prove causal relationship.**

Question 14. Outline further investigations which should be pursued. List one or more factors that could have led to the contamination of the implicated food.

Answer 14.

A. Detailed review of ingredients, preparation, and storage of incriminated food. For bacterial food poisoning need:

1) initial contamination (point of origin vs point of consumption)

2) improper time-temperature relationships with respect to preparation, cooking, serving, and storage

B. Specific things about which one might inquire:

1) Origin of the meat – some sources may be at higher risk than others. Animal meats are often contaminated at time of slaughter. This aspect is usually quite difficult to control.

2) Storage of meat to time of cooking (should be kept frozen or refrigerated). This usually doesn't pose problems and since most meat is **not** eaten raw, subsequent cooking would considerably lessen the risk of disease.

3) Cooking procedures – often difficult to control both in public/private sectors. Temperatures attained and duration of optimum cooking temperatures poorly monitored. Failure to reach adequate cooking temperatures associated with diseases other than *C. perfringens* for the most part.

4) Cross-contamination – a factor difficult to control since knives, counter space, cutting boards, and pots or pans, are often used for both raw foods and cooked foods without interim cleansing.

5) Inadequate refrigeration of cooked foods – common in *C. perfringens* outbreaks. Cooked foods essentially allowed to incubate for several hours during cooling process. Not easy to correct as may involve expenditures for additional refrigeration appliances and use of shallow pans.

6) Inadequate reheating of cooked foods – as with 3).

7) Improper holding temperatures while serving – Here again, difficult to control, but commonly associated with disease outbreaks including *C. perfringens*. The food was essentially held at temperatures that permitted the growth of contaminating organisms rather than at 140 degrees Fahrenheit or above which would have prevented their multiplication.

Question 15. In the context of this outbreak, what control measures would you recommend?

Answer 15.

1. After collecting appropriate specimens for laboratory analysis, destroy remaining foods to prevent their consumption.

2. Prevent recurrence of similar event in the future.

a. Educate food handlers in proper techniques, stressing importance of time-temperature relationships.

b. Acquire necessary equipment for properly cooking, cooling, serving, and storing foods.

c. When applicable, eliminate sources of contaminated food.

3. Basic principles in prevention of *C. perfringens*.

a. Cook all foods to minimum internal temperature of 165 degrees Fahrenheit.

b. Serve immediately or hold at > 140 degrees Fahrenheit.

c. Any leftovers should be discarded or immediately chilled and held at < 40 degrees Fahrenheit using shallow pans.

d. All leftovers should be reheated and held at temperatures given above for cooked foods.

Question 16. Was it important to work up this outbreak?

Answer 16.

Reasons why it was important:

1. To identify factors associated with its occurrence in order to institute the necessary measures to prevent future recurrences.

2. To provide reassurance that a deliberate act of poisoning was not involved.

3. To demonstrate that public health officials can react promptly to a problem and identify causative factors utilizing epidemiologic methods.

Self-Assessment Quiz 6

Now that you have read Lesson 6 and have completed the exercises, you should be ready to take the self-assessment quiz. This quiz is designed to help you assess how well you have learned the content of this lesson. You may refer to the lesson text whenever you are unsure of the answer, but keep in mind that the final is a closed book examination. Circle ALL correct choices in each question.

1. The most common way(s) that a local health department uncovers outbreaks is/are by: (Circle ALL that apply.)

A. receiving calls from affected residents

B. receiving calls from health care providers

C. reviewing all case reports received each week to detect common features

D. performing descriptive analysis of surveillance data each week

E. performing time series analysis to detect deviations from expected values based on the previous few weeks and comparable time periods during the previous few years

2. In an ongoing outbreak of a disease with *no* known source and mode of transmission, the primary reason for an investigation relates to:

A. prevention and control

B. training of staff

C. learning more about the disease

D. being responsive to the concerns of the community

E. legal responsibility

1. Analyze data by time, place, and person
2. Conduct a case-control study
3. Generate hypotheses
4. Conduct active surveillance for additional cases
5. Verify the diagnosis
6. Confirm that the number of cases exceeds the expected number
7. Coordinate who will talk to the press about the investigation

3. For an investigation of an outbreak, what is the logical order of the activities listed above?

A. 1-2-3-4-5-6-7

B. 5-6-4-1-2-3-7

C. 6-5-1-3-2-4-7

D. 7-6-5-4-1-3-2

E. 5-6-1-3-2-4-7

4. If you were a state employee, the first step in the investigation of an outbreak of meningococcal meningitis 200 miles away might include: (Circle ALL that apply)

A. talking with someone knowledgeable about meningococcal meningitis

B. talking with someone knowledgeable about field investigations

C. talking with a couple of the initial case-patients

D. discussing the feasibility of mass vaccination

E. stopping your mail

5. The appropriate role for an epidemiologist from the CDC in the investigation of a local outbreak of botulism (possibly foodborne):

A. is to lead the investigation in consultation with CDC experts

B. is to provide consultation to the local staff who will conduct the investigation

C. is to lend a hand to the local staff

D. is whatever is negotiated in advance with the local health department

6. As described in this lesson, the primary distinction between the terms "outbreak" and "epidemic" is:

A. "outbreak" does not imply that the cases are all related

B. "outbreak" implies a grouping of cases but not necessarily more than expected

C. "outbreak" is limited to fewer than 20 cases, epidemic to more than 20

D. "outbreak" does not generate as much anxiety among the public

Number of cases of Disease X reported to the state health department by Counties A-D

	Week Ending					
County	12/13	12/20	12/27	1/3	1/10	1/17
A	4	3	2	2	3	1
B	12	9	0	0	24	15
C	1	0	1	2	7	9
D	1	1	0	1	0	0

7. Explanations most consistent with the pattern of case reports received from County B include: (Circle ALL that apply.)

A. changes in the case definition

B. change in the denominator

C. new physician in the county

D. change in diagnostic procedures

E. batch processing

8. Why should an investigator who has no clinical background nonetheless talk to a patient or two as an early step in the outbreak investigation? (Circle ALL that apply.)

A. To verify the clinical findings as part of verifying the diagnosis

B. To verify the laboratory findings as part of verifying the diagnosis

C. To learn more about the clinical manifestations of the disease

D. To develop hypotheses about the cause of the outbreak

E. To advise the patient about the common risk factors and usual course of the illness, after reviewing *Control of Communicable Diseases in Man*

9. A case definition during an outbreak investigation should specify: (Circle ALL that apply.)

A. clinical criteria

B. time

C. place

D. person

E. hypothesized exposure

10. A characteristic of a well conducted outbreak investigation is that:

A. every case is laboratory confirmed

B. a few cases are laboratory confirmed and the rest meet the case definition

C. a "loose" case definition is used during the analytic epidemiology phase

D. the case definition includes three categories: definite, probable, and possible

11. Common methods of identifying additional cases (expanding surveillance) as part of an outbreak investigation include: (Circle ALL that apply.)

A. sending a letter to physicians

B. telephoning the infection control nurse at the local hospital

C. advising the public through newspapers, TV, and radio to contact the local health department

D. asking case-patients who they were with at the time of exposure (if known)

E. reviewing morbidity and mortality data for the local area from the National Center for Health Statistics

12. The ultimate purpose for characterizing an outbreak by time, place, and person is to:

A. identify errors and miscodes in the data

B. provide a comprehensive description of an outbreak by portraying its time course, geographic extent, and populations most affected by the disease

C. ensure that all true cases are captured by the surveillance system

D. generate hypotheses

E. test hypotheses

13. For a disease of unknown etiology and incubation period, an epidemic curve can be used to derive which of the following? (Circle ALL that apply.)

A. Peak dates of onset of the illness

B. Peak dates of reporting of the cases to the health department

C. Probable period of exposure

D. Future direction of the epidemic

14. Which of the following apply to drawing an epidemic curve? (Circle ALL that apply.)

A. The *y*-axis is dates of onset of the illness

B. The time interval should be less than one-eighth the minimum incubation period of the disease

C. The type of graph should be a histogram

D. The graph should begin with the first case of the epidemic

15. For *Clostridium perfringens* food poisoning, the minimum incubation period is 8 hours, and the average incubation period is 10 to 12 hours. Based on the graph shown below, when is the probable period of exposure?

A. October 6, periods 1-2 (12:01 A.M. to 8:00 A.M.)

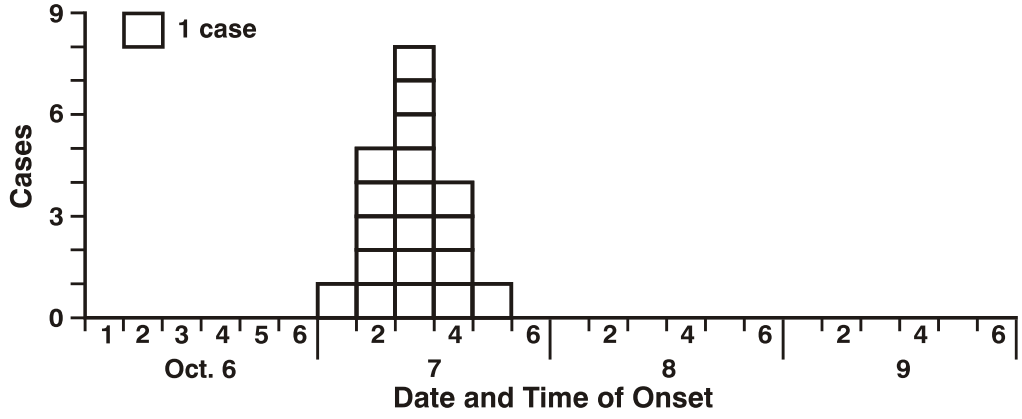
B. October 6, periods 2-3 (4:01 A.M. to noon)

C. October 6, periods 3-4 (8:01 A.M. to 4 P.M.)

D. October 6, periods 4-5 (12:01 P.M. to 8:00 P.M.)

E. October 6, periods 5-6 (4:01 P.M. to midnight)

Figure 6.12
Data and time of onset (by 4 hour periods starting at 12:01 A.M. each day)



16. The geographic distribution of cases should be tabulated or mapped according to:

A. residence of each case

B. place of usual occupation, school, or other primary daytime exposure

C. health care facility where diagnosis was made

D. location where disease onset occurred

E. variable of "place" that produces a meaningful pattern

17. Reasonable ways of generating hypotheses in an outbreak investigation include: (Circle ALL that apply.)

A. asking the local health officer what he/she thinks is the cause

B. asking the case-patients what they think is the cause

C. reviewing a textbook about the disease under investigation

D. postulating explanations for the patterns seen in the descriptive epidemiology

E. focusing on the patients who do not fit the general patterns seen in the descriptive epidemiology

18. During an investigation of an outbreak of gastroenteritis on a small college campus, the investigators confirmed the diagnosis, searched for additional cases, and characterized the cases by time, place, and person. No obvious hypotheses regarding source or mode of transmission came to mind. The investigators should next:

A. interview a few cases in depth

B. conduct a case-control study

C. conduct a cohort study

D. sample and test foods from the school dining hall for the incriminated agent

E. interview and test the dining hall foodhandlers for the incriminated agent

19. In an epidemiologic study, investigators enrolled 100 children with Kawasaki syndrome and 100 children *without* Kawasaki syndrome. Among children with Kawasaki syndrome, 50 had been exposed to compound C in the previous 3 weeks. Among those without Kawasaki syndrome, 25 had been exposed to compound C. In this study, the best estimate of the relative risk of Kawasaki syndrome associated with exposure to compound C is:

A. 1.0

B. 1.5

C. 2.0

D. 3.0

E. not calculable from the information provided

20. In the epidemiologic study of Kawasaki syndrome described in the previous question, the mean serum porcelain levels of children with Kawasaki syndrome was lower than the mean serum porcelain levels of children without Kawasaki syndrome. The difference was statistically significant at the 5% level ($p < 0.05$). This means that:

A. elevated serum porcelain causes Kawasaki syndrome

B. deficiency of serum porcelain causes Kawasaki syndrome

C. the difference between mean serum porcelain levels is unlikely to have occurred by chance alone

D. the difference between mean serum porcelain levels is likely to have occurred by chance alone

21. The report of an epidemiologic study described the association between a particular exposure and a particular disease as “a weakly positive association, but not statistically significant at the 0.05 level.” The data most consistent with this statement is:

A. odds ratio = 10.0, p-value = 0.20

B. odds ratio = 1.5, p-value = 0.03

C. relative risk = 1.8, p-value = 0.01

D. relative risk = 10.0, p-value = 0.10

E. risk ratio = 1.8, p-value = 0.20

Use the data in this table for questions 22 and 23.

	Ate specified food			Did not eat specified food		
Food item	**Ill**	**Well**	**Total**	**Ill**	**Well**	**Total**
Macaroni salad	25	15	40	20	39	59
Potato salad	17	38	55	28	16	44
Three-bean salad	43	47	90	2	7	9
Punch	40	52	92	5	4	7
Ice cream	20	1	21	25	53	78

22. After attending a retirement party for the agency director, many of the health department staff developed gastroenteritis. All attendees were interviewed by the public health nurse who had recently completed the CDC *Principles of Epidemiology* self study course. Calculate the appropriate measure of association for each of the home-made food items shown in the table above. For which food is the measure of association largest?

A. Macaroni salad

B. Potato salad

C. Three-bean salad

D. Punch

E. Ice cream

23. Which of the food items do you think is most likely to have caused this outbreak?

A. Macaroni salad

B. Potato salad

C. Three-bean salad

D. Punch

E. Ice cream

24. Control and prevention measures should be implemented:

A. as early as possible after verifying the diagnosis

B. as early as possible after performing the descriptive epidemiology

C. as early as possible after performing the analytic epidemiology (testing hypotheses)

D. as early as possible after refining the hypotheses and executing additional studies

25. For a federal investigator, which of the following communication modes should be used first to announce the findings of an outbreak investigation?

A. Written report for local authorities

B. Written report for state newsletter

C. Written report for the *Morbidity and Mortality Weekly Report*

D. Oral report for the local authorities

E. Press conference to explain findings the public

Answers in Appendix J

If you answered at least 20 questions correctly, you understand
Lesson 6 well enough to begin to prepare for the final examination.

References

1. Addiss DG, Davis JP, LaVenture M, Wand PJ, Hutchinson MA, McKinney RM. Community-acquired Legionnaires' disease associated with a cooling tower: evidence for longer-distance transport of Legionella pneumophila. Am J Epidemiol 1989;130:557-568.
2. Bender AP, Williams AN, Johnson RA, Jagger HG. Appropriate public health responses to clusters: the art of being responsibly responsive. Am J Epidemiol 1990;132:S48-S52.
3. Benenson AS (ed). Control of Communicable Diseases in Man. Fifteenth Edition. Washington, DC: American Public Health Association, 1990.
4. Caldwell GG. Twenty-two years of cancer cluster investigations at the Centers for Disease Control. Am J Epidemiol 1990;132:S43-S47.
5. Centers for Disease Control. Hepatitis–Alabama. MMWR 1972:21:439-444.
6. Centers for Disease Control. Legionnaires' disease outbreak associated with a grocery store mist machine–Louisiana, 1989. MMWR 1990;39:108-110.
7. Centers for Disease Control. Pertussis–Washington, 1984. MMWR 1985;34:390-400.
8. Devier JR, Brownson RC, Bagby JR, Carlson GM, Crellin JR. A public health response to cancer clusters in Missouri. Am J Epidemiol 1990;132:S23-31.
9. Fiore BJ, Hanrahan LP, Anderson HA. State health department response to disease cluster reports: a protocol for investigation. Am J Epidemiol 1990;132:S14-22.
10. Fraser DW, Tsai TF, Orenstein W, et al. Legionnaires' disease: Description of an epidemic of pneumonia. N Engl J Med 1977;297:1189-1197.
11. Goodman RA, Buehler JW, Koplan JP. The epidemiologic field investigation: science and judgment in public health practice. Am J Epidemiol 1990;132:9-16.
12. Gross, M. Oswego County revisited. Public Health Rep 1976;91:168-170.
13. Hedberg CW, Fishbein DB, Janssen RS, et al. An outbreak of thyrotoxicosis caused by the consumption of bovine thyroid gland in ground beef. N Engl J Med 1987;316:993-998.
14. Hertzman PA, Blevins WL, Mayer J, Greenfield B, Ting M, Gleich GJ. Association of the eosinophilia-myalgia syndrome with the ingestion of tryptophan. N Engl J Med 1990;322:869-873.
15. Hopkins RS, Juranek DD. Acute giardiasis: an improved clinical case definition for epidemiologic studies. Am J Epidemiol 1991;133:402-407.
16. Hutchins SS, Markowitz LE, Mead P, et al. A school-based measles outbreak: the effect of a selective revaccination policy and risk factors for vaccine failure. Am J Epidemiol 1990;132:157-168.
17. MacDonald KL, Spengler RF, Hatheway CL, et al. Type A botulism from sauteed onions. JAMA 1985;253:1275-1278.
18. Neutra RR. Counterpoint from a cluster buster. Am J Epidemiol 1990;132:1-8.

19. Rimland D, Parkin WE, Miller GB, Schrack WD. Hepatitis B outbreak traced to an oral surgeon. N Engl J Med 1977;296:953-958.
20. Rosenberg MD, Hazlet KK, Schaefer J, Wells JG, Pruneda RC. Shigellosis from swimming. JAMA 1976;236:1849-1852.
21. Ryan CA, Nickels MK, Hargrett-Bean NT, et al. Massive outbreak of antimicrobial-resistant salmonellosis traced to pasteurized milk. JAMA 1987;258:3269-3274.
22. Schulte PA, Ehrenberg RL, Singal M. Investigation of occupational cancer clusters: theory and practice. Am J Public Health 1987;77:52-56.
23. Swygert LA, Maes EF, Sewell LE, Miller L, Falk H, Kilbourne EM. Eosinophilia-myalgia syndrome: results of national surveillance. JAMA 1990;264:1698-1703.
24. Taylor DN, Wachsmuth IK, Shangkuan Y-H, et al. Salmonellosis associated with marijuana: a multistate outbreak traced by plasmid fingerprinting. New Engl J Med 1982;306:1249-1253.