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

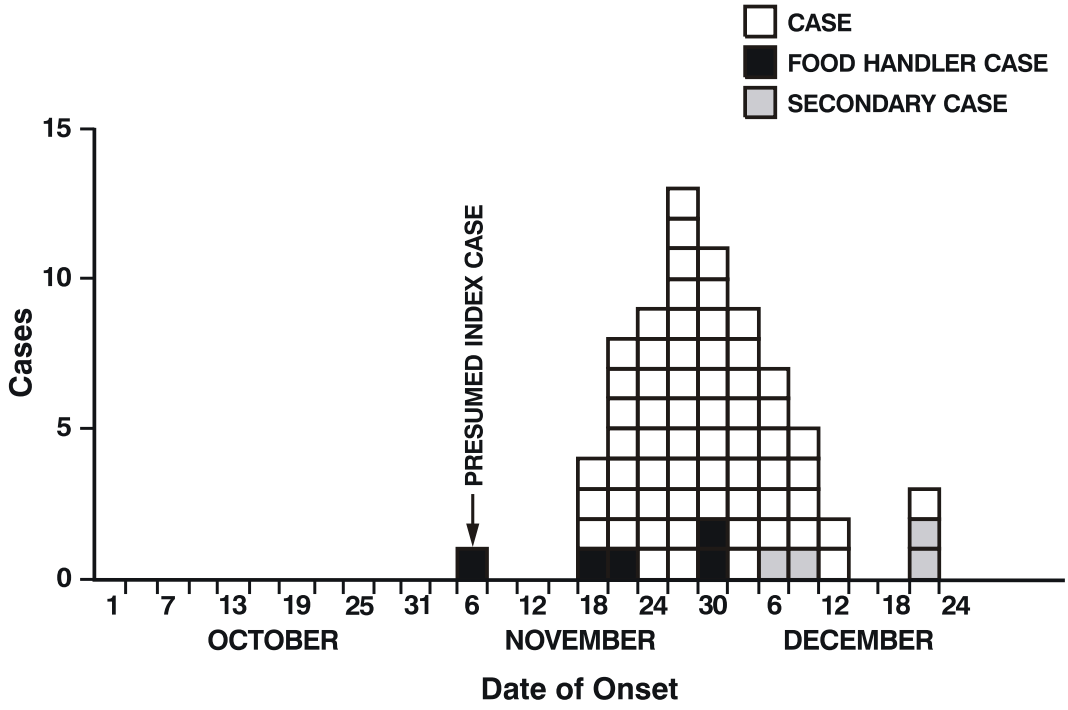
Traditionally, we depict the time course of an epidemic by drawing a histogram of the number of cases by their date of onset. This graph, called an **epidemic curve**, or **epi curve** for short, gives us a simple visual display of the outbreak's magnitude and time trend. Figure 6.2 shows a typical epidemic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

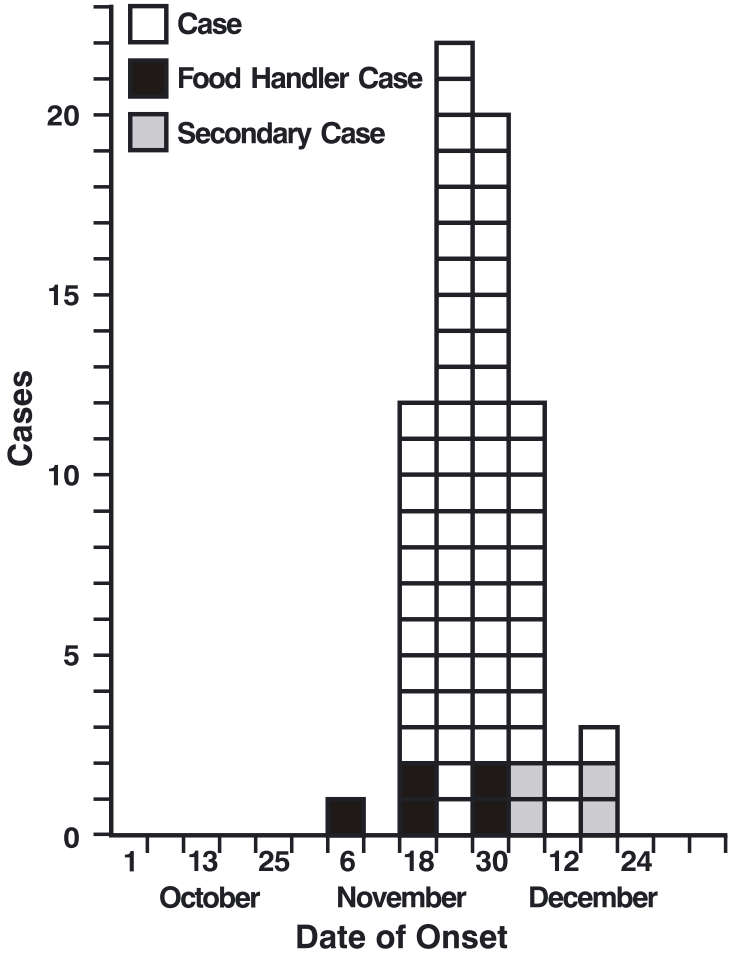
disease and/or its incubation time. In that circumstance, it is useful to draw several epidemic curves with different units on the *x*-axis to find one that seems to portray the data best. For example, Figure 6.3 shows an epidemic curve of the same data as in Figure 6.2; in Figure 6.2 the *x*-axis unit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

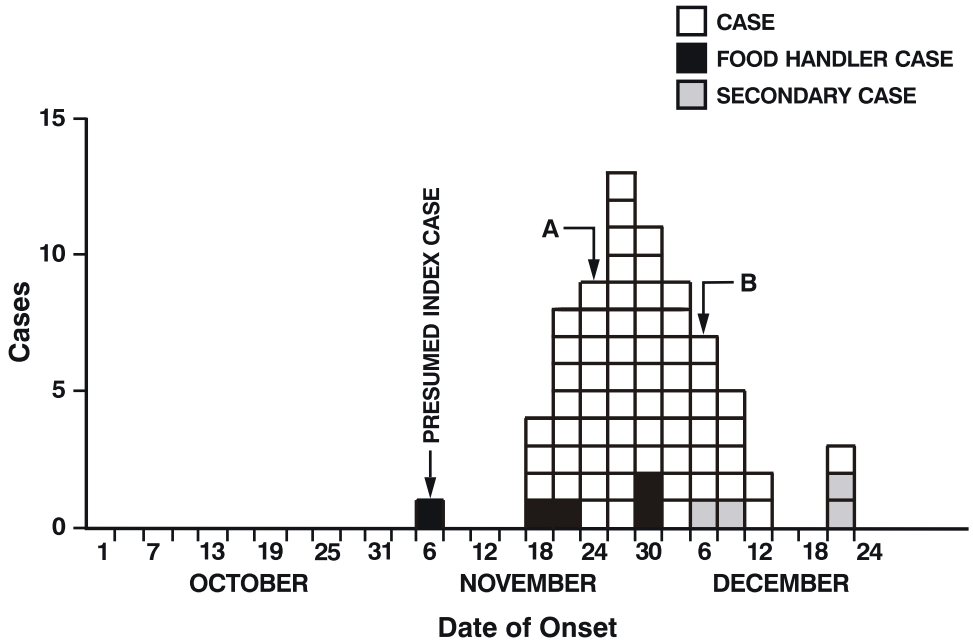
exposure, and the number of persons exposed. Person-to-person spread – a **propagated** epidemic – should have a series of progressively taller peaks one incubation period apart, but in reality few produce this classic pattern.

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



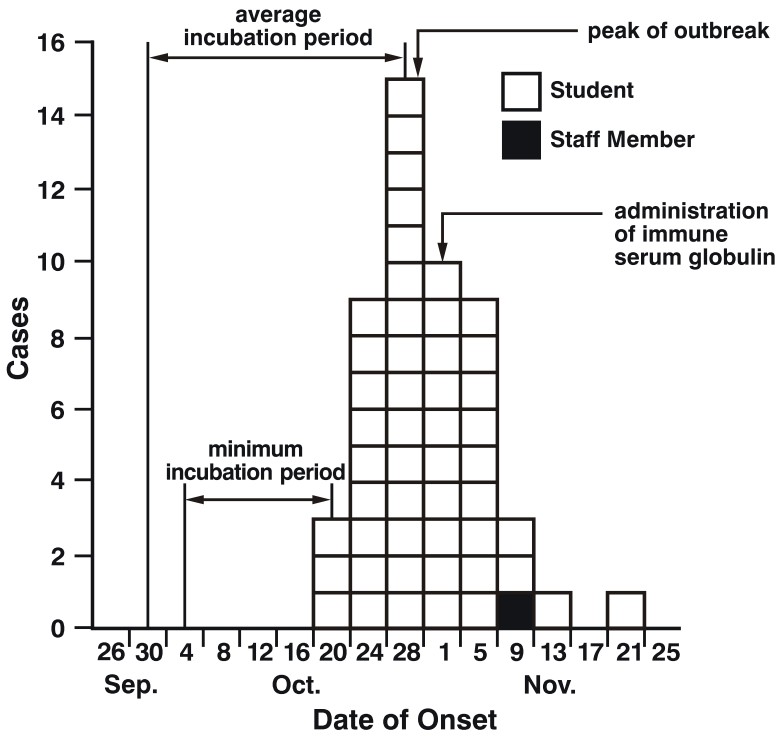
To identify the likely period of exposure from an epidemic curve,

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.