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.