# FOCUS on Field Epidemiology

## CONTRIBUTORS

**Author:**

   Michelle Torok, MPH

**Reviewers:**

   Amy Nelson, PhD, MPH

   FOCUS Workgroup*

**Production Editors:**

   Lorraine Alexander, DrPH

   Gloria C. Mejia, DDS, MPH

**Editor in Chief:**

   Pia D.M. MacDonald, PhD, MPH

* All members of the FOCUS workgroup are named on the last page of this issue.

**UNC SCHOOL OF PUBLIC HEALTH**

**NORTH CAROLINA CENTER FOR PUBLIC HEALTH PREPAREDNESS**

## Case Finding and Line Listing: A Guide for Investigators

It is a cool, crisp evening aboard the *S.S. Epidemiologie*. The cold inky water slaps against the ship, and the members of the Epidemiology Yacht Club pull their coats tightly about them as they gaze toward the horizon. A couple of them point out at the water, where a small, frozen white peak protrudes just above the waves. Wow, it is cold, they think, shivering at the thought of ice. Suddenly, an epidemiologist comes frantically running out onto the deck. Ill-prepared for the cold, he is wearing a t-shirt that reads "FOCUS" with a picture of a magnifying glass. He screams to his fellow passengers, "It's only the tip of the iceberg! THE TIP OF THE ICEBERG!" The others watch in amazement as he shouts his message to anyone he can find. No one is paying attention to what he is saying. As he begins to untie the lifeboats, there is a horrible grinding noise, as the ship's hull is broken by treacherous ice, hiding beneath the surface of those dark waves.
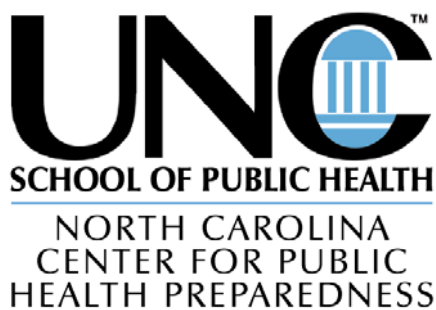
The moral of the story? Once an outbreak is suspected, some leg work needs to be done ASAP. Is the outbreak limited to a small rural town, or is it part of a national situation? Who is affected, only intravenous (IV) drug users or the general population? Answering these questions is one of the first tasks of an outbreak investigation and it requires finding as many cases as possible (see Focus Issue 3, Embarking on an Outbreak Investigation, for more information). Once cases are identified, the investigators will have demographic, risk factor, clinical and laboratory information coming out of their ears! A line listing can help them organize this crucial information and get below the "tip of the iceberg."

## Why is case finding important?

Case finding is important because in outbreak settings, the cases reported to the health department may represent only a small portion of the total number of outbreak-related cases that exist. While these reported cases are very important because they help clue investigators in to an outbreak, they may not adequately represent all of the people affected. The reported cases may represent only a subgroup of all the cases, such as students from a certain school or patients of a health care provider who is particularly conscientious about reporting communicable diseases to the local health department. To get the full scope of an outbreak, investigators need to know exactly what types of people are getting the disease, when they became symptomatic and where they may have been exposed. This information may help identify a potential exposure source and/or cause of the outbreak.

- For example, in 2001, human pulmonary blastomycosis was identified in four students attending the same high school in a small town in North Carolina. There was construction going on at the high school that could have disrupted the soil in a way that was consistent with previ-

ous blastomycosis outbreaks. Some investigators may have concluded from this that the outbreak was due to soil exposure at the school, closed the construction site and moved on to other duties. However, in this outbreak, the investigators were savvy and looked intensively for additional cases. Their activities revealed four additional human cases as well as four canine cases, none of which were related to the high school. Although a common exposure could not be identified, case finding prevented an inappropriate conclusion and intervention from occurring (1).

Another reason case finding is important at the beginning of an investigation is to start refining the case definition (refer to Focus Issue 3 for more information on case definitions). It is then essential to continue refining the case definition throughout an investigation as more detailed information becomes available. Additionally, knowing the extent of the outbreak is crucial for determining the amount of resources that should be allocated to the investigation — an important decision since most health departments have limited resources. Case finding also aids in the development of appropriate control measures by defining the population with the exposure of interest.

### How to find cases

When attempting to find cases at the beginning of an outbreak, it is best to cast a wide net. This can help determine the size and geographic boundaries of the outbreak since cases that are recognized first may represent only the "tip of the iceberg."

Cases can be identified through active or passive case finding strategies. One active case finding method involves soliciting health facilities and laboratories to identify additional cases. Another method of active case finding is to screen an exposed population with a diagnostic test.

Passive case finding, which is less aggressive and requires less resources, may involve examining county or state surveillance data to identify cases reported through the communicable disease reporting system. In an outbreak situation, some cases can be identified through passive case finding, but it is imperative to also perform active case finding. Multiple sources should be used to find cases, and the best methods may require creativity on the part of the investigator. Here are some examples:

- Health care provider offices, clinics, hospitals and laboratories can be asked for information. For ex-

ample, an investigator might visit a local hospital emergency room and ask to review the records of all patients seen with the illness, or ask clinicians to request specimens from all patients who meet a clinical case definition, or ask infection control practitioners to review medical records of patients with a particular diagnosis.

- In some situations it may be appropriate to query the community through local television, radio or newspapers, particularly if the outbreak involves a contaminated food product or a potential act of bioterrorism. For example, in 1989, the New Mexico Department of Health and Environment was notified of three patients with eosinophilia and severe myalgia who had been taking oral preparations of the amino acid L-tryptophan (LT), an essential amino acid normally ingested as a constituent of dietary protein. Media publicity helped generate reports of 154 additional potential cases of a similar illness in 17 states and the District of Columbia within 2 weeks (2).

- Investigators can look at records such as wedding invitation lists, guest books, credit card receipts and customer lists maintained by establishments involved in an outbreak. For example, in 2001, the first case of bioterrorism-related inhalational anthrax had traveled to the state of North Carolina 3 days before he became ill, raising the possibility that exposure to *Bacillus anthracis* spores could have occurred in that state. Hospital intensive-care units, microbiology laboratories, medical examiners, veterinarians, and site investigations at locations visited by the index patient were checked to identify the naturally occurring or bioterrorism-related source of the patient's exposure. At one of these sites, a rural tourist park, credit card receipts and records of annual pass holders were held to potentially track approximately 700 park patrons who had visited the park on the same day as the index case (3). (Remember, however, that if credit card receipts are used, the credit card holder must be told how and why the information was obtained, and how this information will be handled. State and federal laws vary in what information is to be held confidential in outbreak investigations (4). )

- When an exposure has occurred in a defined setting in a defined population, it may be effective to ask every person in the population about symptoms. For example, everyone who was at a church picnic, wedding, or school function or on a cruise ship may be queried. Or a school nurse might be called to request the names of all students who

were seen with a particular illness. It may also be helpful to ask cases if they know of anyone else who has become ill.

### Common challenges of case finding

Even with active case finding attempts, several factors make it difficult to identify or confirm all of the cases. First, not all specimens are routinely tested for certain pathogens. For example, *Escherichia coli* O157:H7 testing would need to be requested by a referring health care provider because it is often not in the standard panel of stool screens. If the health care provider did not request *E. coli* O157:H7 testing, the person examined could not be classified as a laboratory-confirmed case. Of course, the patient could be a clinically confirmed case (if this definition is used), assuming that he or she met the clinical case definition. Second, diseases cause a spectrum of symptoms, ranging from mild to severe. If someone, for example, experiences only slight gastrointestinal symptoms (such as abdominal cramping) the person may not seek medical attention thereby removing the chance of being identified through this avenue. Even if the outbreak is publicized in the media, the person may not connect his or her mild symptoms to the publicized outbreak. Finally, the exposed population might not be well defined. Say, for example, a family vacationing from out of state ate at a restaurant involved in a foodborne disease outbreak. If that family paid with cash and no one in the community knew who they were, the investigators would not know to look for that family.

Fortunately, not every single case is needed to glean useful information in an outbreak. However, it is important to identify as many cases as possible and to avoid identifying only a subgroup of the cases because they may not be representative of all cases.

### Information to collect during the case finding process

The information collected depends on the outbreak, but it can be grouped into four categories (5) :

- Identifying information
- Demographic information
- Clinical information
- Risk factor information

**Identifying information of the case** may include the name, address, phone number, date of birth and contact information for the person who reported the case.

These data enable the investigator to obtain further details about the case or to tell the case about the investigation. It is important to remember that all identifying information should be kept confidential. One way to do this is to assign each case a number. The file containing personal identifiers and the assigned case number should be password protected. This number can be used throughout the investigation instead of the case's name.

**Demographic information** that may be relevant includes age, gender, race, occupation, place of occupation and travel history. This information depends on the nature of the outbreak. For example, if there were an outbreak of diarrheal disease among preschool aged children, it would be important to ask whether or not the child attended day care and, if so, the name of the facility. Demographic information also enables the investigator to describe at-risk group(s) of individuals.

**Clinical information**, such as clinical symptoms, the date of symptom onset, lab findings and severity of illness, allows the investigator to verify that the case definition has been met, to characterize the disease, and to create an epidemic curve (a future FOCUS issue will discuss creating epidemic curves).

**Risk factor information** allows the epidemiologist to focus an investigation. Since it is collected in the preliminary stages of an investigation, risk factor information for a line listing is usually confined to general potential risk factors and well established risk factors. Once hypotheses have been generated (upcoming FOCUS issues will deal with this subject), further details are collected. The relevant risk factor information depends upon the outbreak. For example, in an outbreak of *E. coli* O157:H7, the investigator would ask about ground beef, lettuce, alfalfa sprouts, recreational water consumption, and day care attendance for the line listing. Then, if preliminary analyses implicated beef consumption as a potential exposure, details such as the brand name of the beef would be collected as part of an a hypothesis testing study.

### What is a line listing and why is it important?

Using the information described above, a line listing is produced by epidemiologists in outbreak investigations. A line listing allows information about time, person, and place to be organized and reviewed quickly. It is also a good way to keep track of different categories of cases. For example, cases can be entered into the line listing as possible, probable or confirmed (laboratory-confirmed, clinically confirmed or both). The designation can easily be updated as the investigation pro-

gresses (for example, when a "probable" case is confirmed by the laboratory).

### How to create and manage a line listing

Line listings can be created on paper (hard copy) or on a computer (electronic version). If a computer is available, commercial programs such as Microsoft Excel, and Microsoft Access or freeware such as EpiInfo (http://www.cdc.gov/epiinfo/index.htm) can be used. An advantage of creating an electronic line listing is that frequency distributions and epidemic curves can be generated quickly.

Information that goes into a line listing is generally collected on a questionnaire or standard case form. The critical components of these questionnaires are then used to create a line listing (5). To set up a line listing, create a table in which each row represents a case and each column represents a variable of interest (variables of interest depend on the nature of the outbreak, see Table 1). New cases should be added to the list as they are identified, and all cases should be updated throughout the investigation as new information is obtained.

Table 1 gives an example of a line listing that might be used in an investigation of an outbreak of hepatitis A (5).

The number of variables to include in a line listing will vary depending on the type of line listing and the outbreak. Line listings that contain only the basic critical information have the advantage of providing a quick visual assessment of different aspects of the outbreak. However, a line listing with additional information may be more useful for assessing and characterizing the outbreak (this is much easier with an electronic version).

All line listings should include the components of the case definition.

- For example, in investigating an outbreak of acute hepatitis A, the CDC defines a case in the following manner (6):

  Clinical Description: an acute illness with a) discrete onset of symptoms and b) jaundice or elevated serum aminotransferase levels.

  Laboratory Criteria for Diagnosis: immunoglobulin M (IgM) antibody to hepatitis A virus (anti-HAV) positive.

  In this case, a line listing should include columns indicating the presence/absence of the discrete onset of symptoms, jaundice and elevated aminotransferase levels. It should also include a column in which to indicate the presence/absence of IgM antibody to hepatitis A virus (anti-HAV) positive.

### Table 1. Example of a line listing for acute Hepatitis A*

| | | | | Signs/Symptoms | | | | | | Labs | | Demographics | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Case # | Report Date | Onset | Physician Diagnosis | N | V | A | F | D | J | HAIgM | Other | Sex | Age |
| 1 | 10/12/02 | 10/5/02 | Hepatitis A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | Low SGOT | M | 37 |
| 2 | 10/12/02 | 10/4/02 | Hepatitis A | 1 | 0 | 1 | 1 | 1 | 1 | 1 | Low Alt | M | 62 |
| 3 | 10/13/02 | 10/4/02 | Hepatitis A | 1 | 0 | 1 | 1 | 1 | 1 | 1 | Low SGOT | M | 38 |
| 4 | 10/13/02 | 10/9/02 | NA | 0 | 0 | 1 | 0 | ? | 0 | NA | NA | F | 44 |
| 5 | 10/15/02 | | Hepatitis A | 1 | 1 | 1 | 1 | 1 | 0 | 1 | Hbs/Ag- | M | 17 |
| 6 | 10/16/02 | 10/6/02 | Hepatitis A | 0 | 0 | 1 | 1 | 1 | 1 | 1 | SGOT=24 | F | 43 |

N=nausea  V=vomiting  A=elevated aminotransferase  F=fever  D=discreet onset  J=jaundice  HAIgM=hepatitis AIgM antibody test  SGOT=serum glutamic oxaloacetic transaminase  ALT=alanine aminotransferase  Hbs=hepatitis B surface antigen  Ag-=antigen negative

1="yes", 0="no"

*  This table illustrates a line listing that might be used during an outbreak of hepatitis A. It was adapted from the CDC's "Excellence in Curriculum Integration through Teaching Epidemiology" program. Additional variables that might be helpful to include are drug use, occupation, meal at restaurant X, neighborhood of residence and sexual orientation.

A line listing always includes the patient's name or identifying number and the date of symptom onset or date of specimen collection (depending on what information is available). Line listings also include demographic information such as age, gender, race and occupation, as well as risk factor information. The information needed depends upon the nature of the outbreak. For example, in the case of the hepatitis A outbreak, relevant risk factor information includes drug use or sexual behavior. In some other outbreaks, this would not be relevant. Finally, if questions about the entered information arise, it can be helpful, in situations when more than one person is entering data, to include the initials of the person who entered the data.

### Using Information from a Line Listing

Frequency distributions of demographic factors such as age, race and gender are important because they provide further information about potential outbreak exposures and risk of disease.

- For example, in a 1997 outbreak of *E. coli* 0157:H7, over 50% of the cases occurred in women. This was quite different from other *E. coli* 0157:H7 outbreaks found to be associated with beef products. In this outbreak, the surprising proportion of female cases probably clued the investigators in to the possibility of a non-hamburger reservoir, which was found to be alfalfa sprouts (7).

Frequency distributions of potential risk factors such as occupation, sexual behavior or recreational activities or hobbies may clue the investigators in to the outbreak source or transmission route.

- For example, in an outbreak of multi-drug resistant *Salmonella* Typhimurium, cases were employees of a veterinary facility. Although most cases of *Salmonella* infections in the U.S. are transmitted by food, they can also be transmitted through farm animals, reptiles and pets (8).

Spot maps/Geographic Information Systems (GIS) can be used to plot locations such as residence or place of employment (A future FOCUS issue will discuss GIS). This information can provide clues to potential exposure patterns in the outbreak. If the overall population varies in different areas on the map, the investigator should plot the attack rate in each area (instead of number of cases) because plotting only the number of cases can be misleading.

- For example, in an investigation of nosocomial surgical site infections in a large medical facility, a spot map of the hospital may show the clustering of cases by operating room.

### Conclusion

Finding cases and creating a line listing are two crucial parts of any outbreak investigation. Finding cases is necessary to characterize the outbreak as accurately as possible, and a line listing enables the investigator to quickly summarize, visualize and analyze the key components of the outbreak.

**Glossary:**

**Active case finding:** Case finding in which the investigator collects case information from providers/ institutions.

**Passive case finding:** Case finding that relies on providers/ institutions to report cases.

**Attack rate:** The cumulative incidence of infection in a group observed over a period during an epidemic.

**Line listing:** A table in which critical information from an outbreak is listed. Each column represents an important variable (e.g. identifier, age, sex) and each row represents a different case.

## REFERENCES

1. Personal communication, Dr. Pia MacDonald, 2003.

2. CDC. Epidemiologic notes and reports eosinophilia-myalgia syndrome -- New Mexico. MMWR 1989; 38 (45): 765-767.

3. Maillard JM et al. First case of bioterrorism-related inhalational anthrax, Florida, 2001: North Carolina investigation. Emerg Infect Dis 2002; 8 (10): 1035-1038.

4. Dwyer DM and Groves C. Outbreak Epidemiology. In: Nelson KE, Masters Williams C, Graham NMH, eds. Infectious Disease Epidemiology: Theory and Practice. Gaithersburg, MD: Aspen Publishers, 2001: 119-148.

5. CDC. EXCITE: Epidemiology in the Classroom. How to Investigate an Outbreak: Steps of an Outbreak Investigation, 2002. (http://www.cdc.gov/excite/classroom/outbreak_steps.htm#step6).

6. CDC. Summary of notifiable diseases-United States, 2001. MMWR 2003; 50 (53): 100-108.

7. Breuer T, Benkel DH, Shapiro RL et al. A multistate outbreak of *Escherichia coli* 0157:H7 infections linked to alfalfa sprouts grown from contaminated seeds. Emerg Infect Dis 2001;7 (6): 997-982.

8. CDC. Outbreaks of multidrug-resistant *Salmonella* Typhimurium associated with veterinary facilities --- Idaho, Minnesota, and Washington, 1999. MMWR 2001; 50 (33): 701-4.

## UPCOMING TOPICS!

- Epidemic Curves Ahead
- Hypothesis Generation during Outbreaks
- Designing Questionnaires for Outbreaks
- Interviewing Techniques for Epi Studies
- Introduction to Forensic Epidemiology

We are on the web!
http://www.sph.unc.edu/nccphp