VOLUME 1, ISSUE 5

# FOCUS on Field Epidemiology

## Epidemic Curves Ahead


**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.**



The Center for Public Health Preparedness is funded by a cooperative agreement between the Centers for Disease Control and Prevention and the Association of Schools of Public Health, Project # A1011-21/21


An epidemiologist is hardly the first person to call in a health emergency. When someone is having a heart attack, no one cries, "Is there an epidemiologist in the house?" But epidemiologists can offer to count that person and tally him in the "Yes" column for heart attacks. And epidemiologists are really good at counting!

With infectious diseases, when we count cases in an outbreak, we like to put that tally to good use. We do this by creating an epidemic curve, or epi curve. While outbreak investigations can throw many unexpected curves, the epi curve is one that should be created in every potential outbreak situation.

This issue of FOCUS explains how epi curves are used and describes methods for making an epi curve.

### What exactly is an epi curve and how can it help in an outbreak?

An epi curve is a graphic depiction of the number of outbreak cases by date of illness onset. It is useful because it can provide information on the outbreak's (1):

- Pattern of spread
- Magnitude
- Outliers
- Time trend
- Exposure and/or disease incubation period

Each of these aspects of an epi curve will be discussed in detail.

### Outbreak pattern of spread

The overall shape of the curve can reveal the type of outbreak (common source, point source or propagated).

A **common source** outbreak is one in which people are exposed intermittently or continuously to a common harmful source. The period of exposure may be brief or long. An intermittent exposure in a common source outbreak often results in an epi curve with irregular peaks that reflect the timing and extent of the exposure (2). Figure 1 shows an example of a common source outbreak with intermittent exposure. Continuous exposure will often cause cases to rise gradually (and possibly plateau, rather than peak) (2). Figure 2 gives an example of a continuous exposure.

An epi curve with a sharp upward slope and a gradual downward slope typically describes a **point source** outbreak. A point source outbreak is a common source outbreak in which the exposure period is relatively brief and all cases occur within one incubation period. Figure 3 illustrates a point source epidemic curve.

A **propagated** outbreak is one that is spread from person to person. Because of this, propagated epidemics may last longer than common source epidemics and may lead to multiple waves of infection if secondary and tertiary cases occur. The classic propagated epi curve has a series of progressively taller peaks, each an incubation period apart, but in reality the epi curve may look somewhat different (2). Figure 4 is an example of a propagated epi curve.

Figure 1. Example of an epidemic curve from a common intermittent exposure source

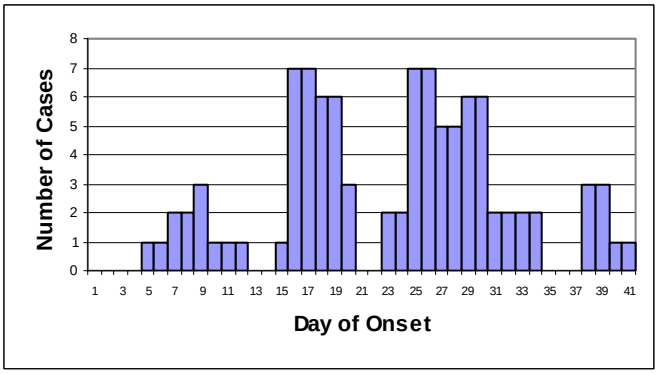



Figure 2. Example of an epidemic curve from a common continuous exposure source

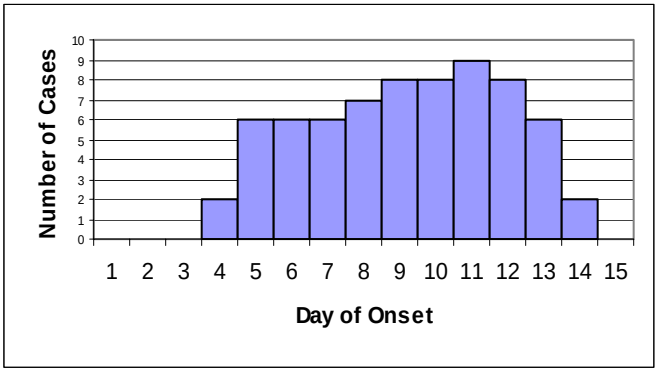



Figure 3. Example of a point source epidemic curve

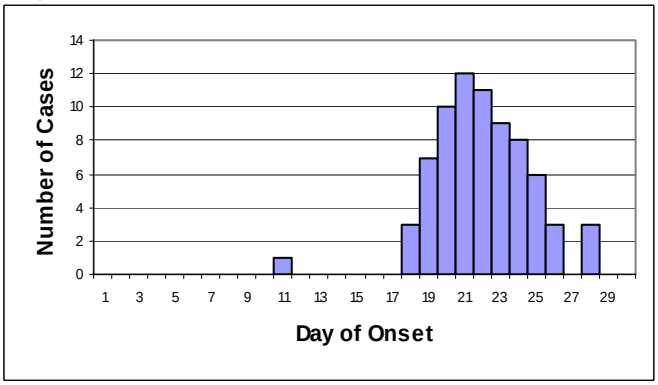
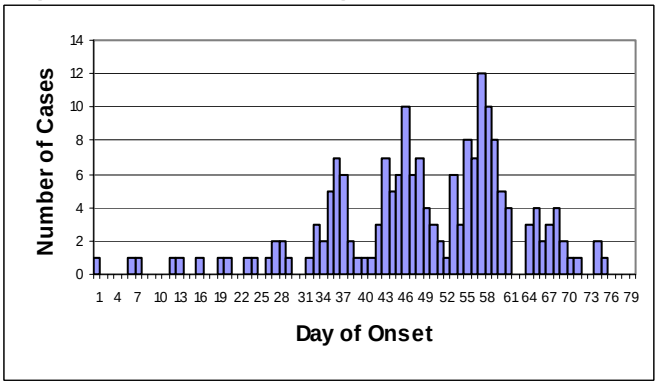



Figure 4. Example of a propagated epidemic curve




### Magnitude of the outbreak

An epidemic curve can provide a sense of the magnitude of the outbreak as well. For example, there were 73 cases reported in the point source outbreak shown in Figure 3—a fairly large outbreak for certain diseases in a small geographical area. Additional information about the magnitude of the outbreak within subpopulations can be obtained by stratifying the epi curve, that is, separating the sample into several subsamples according to specific criteria, such as gender, age, clinical symptoms or geographic location.

### Outbreak time trend

Again, using the point source outbreak (Figure 3) as an example, the epi curve allows us to glean useful information about the time trend involved. Illness onset for the first case patient was on Day 11, and cases continued to be reported for the rest of the month. The outbreak peaked on Day 21 and then began to decline. No new cases were reported after Day 28. Unless there has been secondary spread (cases of disease acquired from a primary case), based on the curve, this outbreak appears to be over.

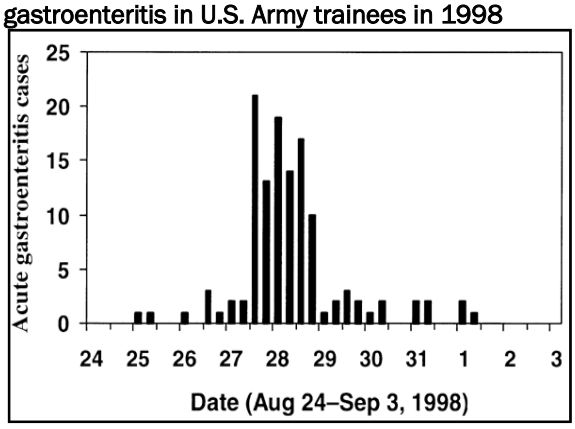
### Outbreak outliers

Cases at the very beginning or end that do not appear to be related to the outbreak are referred to as "outliers." The first thing that should be done when considering outliers is to make sure they are not mistakes due to miscoding or data entry error. Assuming they are not errors, important information can be deducted from outliers. For example, an early case may not be part of the outbreak; it may represent the baseline level of illness. However, it may also represent the source of the outbreak, such as an infected food handler, or it may be a case that was exposed earlier than the others. A late case may not be part of the outbreak; but alternatively, a late case may represent an individual who had a long incubation period, who was exposed later than the other cases, or who was a secondary case (acquired the disease from a primary case) (2).

### Period of exposure/incubation period for the outbreak

If the timing of the presumed exposure is known, epi curves can be used to estimate the incubation period of the disease, and this may facilitate identification of the causative agent. The period between the known or hy-

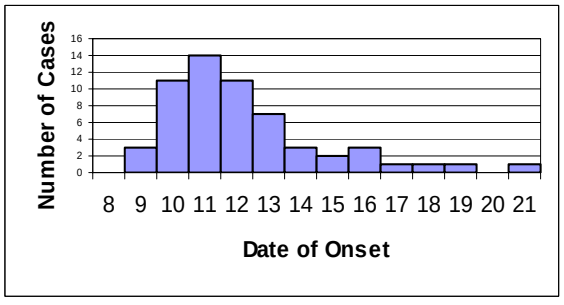
Figure 5. Epidemic curve from outbreak of norovirus gastroenteritis in U.S. Army trainees in 1998




pothesized exposure time and the peak of the epi curve represents the hypothesized median incubation period (3). Figure 5 was created from a U.S. Army trainee outbreak of norovirus gastroenteritis in 1998 (4). In this case, the exposure was thought to have occurred August 26 or 27. Based on the epi curve, the median hypothesized incubation period was very short, 24-36 hours. This is consistent with noroviruses, which have an average incubation period of 12-48 hours (5).

In common source outbreaks involving diseases with known incubation periods, epi curves can help determine the probable period of exposure (2). This can be done by looking up the average incubation period for the organism and counting back from the peak case the amount of time of the average incubation period. To estimate the minimum incubation period, count back the minimum incubation period time from the earliest case on the epi curve.

Figure 6. Epidemic curve to estimate exposure period




Ideally, the minimum and average outbreak incubation periods should be close, and the time between them will represent the probable period of exposure. Since this technique is not precise, it is recommended to widen the identified exposure period by 10% to 20% on either side so as not to miss a potential exposure (6).

Using Figure 6 as an example, suppose the causative organism is *Vibrio cholerae* and the average and minimum incubation periods are approximately 3 and 1 days, respectively. Three days back from the peak case (the 11th) would be the 8th of the month. Counting 1 day back from the first case would also be the 8th. Thus, the hypothesized exposure date is the 8th of the month. Since this technique is not precise, the exposure period should be widened several days on either side, which would give an approximate probable exposure period from the 5th to the 11th. Now potential exposures during this time frame can be investigated in the hope of finding the source of the outbreak.

## How to make an epi curve

As shown in the epi curves above, the structure of an epi curve is straight forward. Simply plot the number of cases of disease reported during an outbreak on the y-axis (the vertical line) and the time/date of illness onset on the x-axis (the horizontal line). Here are some technical tips:

- One of the trickier aspects of creating an epi curve is choosing the unit of time for the x-axis. This choice is usually based on the incubation period of the illness and the time interval of the outbreak. In general, a

Figure 7. Epidemic curve using one day as time unit on x-axis

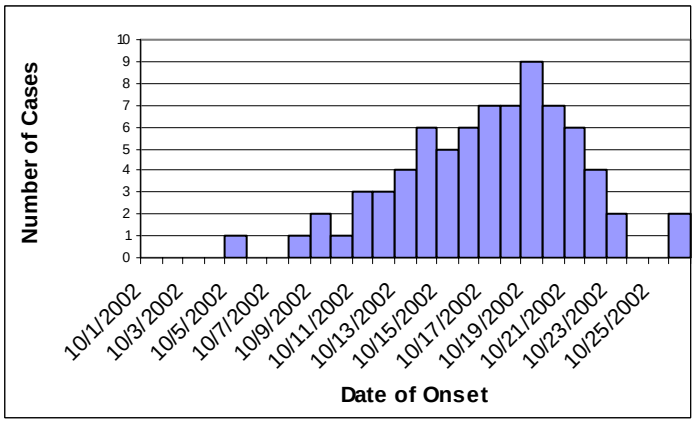
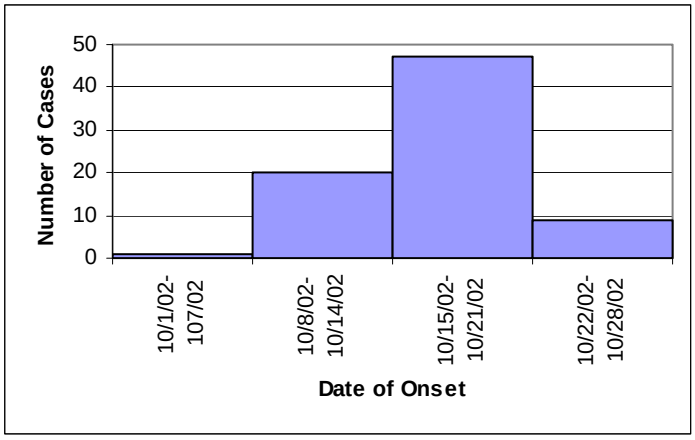


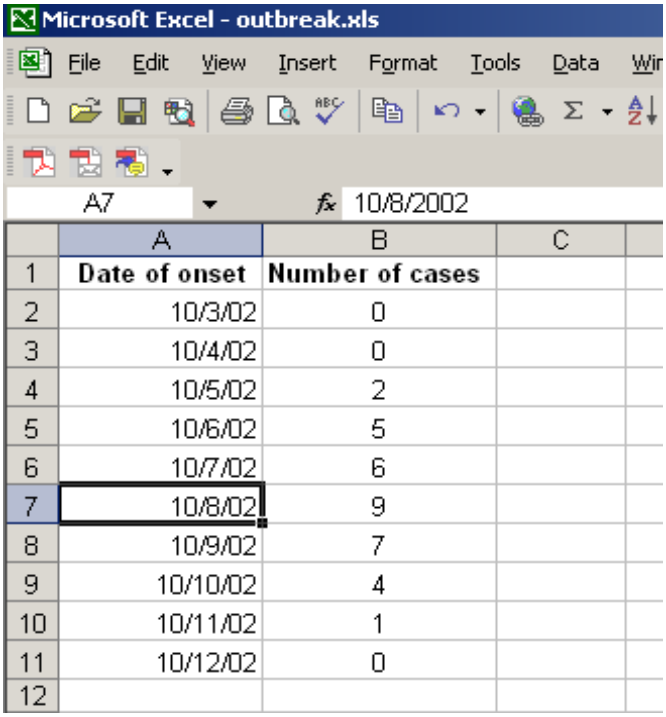

Figure 8. Epidemic curve using one week as time unit on x-axis



time unit that is approximately a quarter of the incubation period is usually a good place to start. For example, the mean incubation period for influenza is 36 hours. Therefore, for an outbreak of influenza lasting for several days, it may be useful to begin the epi curve with 9-hour intervals on the x-axis.

- If the incubation period of the illness (or the illness itself) is not known, several epi curves with different time intervals on the x-axis should be examined to see which one best represents the data. Suppose one is investigating an outbreak of a foodborne illness at a restaurant, but the causative agent has not yet been identified. Figure 7 shows an epi curve using one day as the unit of time for the x-axis, while Figure 8 shows the same data using one week as the x-axis time interval. It is obvious from the two epi curves that using one day on the x-axis is a better choice. It distributes the cases more evenly and highlights a potential index case on 10/5/02 (perhaps a food handler).
- For most diseases, the date of onset is appropriate for the x-axis, but for illnesses with very short incubation periods (for example, *Staphylococcus aureus* food poisoning), the hour of onset may be preferable. In contrast, for diseases with long incubation periods, such as tuberculosis, the best time interval may be days, weeks, or even months (6).

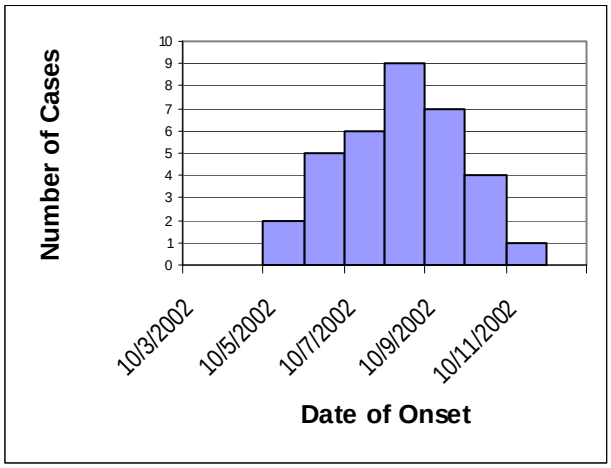
**Table 1. Example of how to enter data in MS Excel* to make an epidemic curve**




| | A | B | C |
|---|---|---|---|
| 1 | Date of onset | Number of cases | |
| 2 | 10/3/02 | 0 | |
| 3 | 10/4/02 | 0 | |
| 4 | 10/5/02 | 2 | |
| 5 | 10/6/02 | 5 | |
| 6 | 10/7/02 | 6 | |
| 7 | 10/8/02 | 9 | |
| 8 | 10/9/02 | 7 | |
| 9 | 10/10/02 | 4 | |
| 10 | 10/11/02 | 1 | |
| 11 | 10/12/02 | 0 | |
| 12 | | | |

* Microsoft® Excel Copyright© Microsoft Corporation

**Figure 9. Example of how to make an epi curve using MS Excel**




- Epi curves are a type of histogram, so there should not be any space between the x-axis categories.
- A simple but important point is this: remember to label the axes correctly and include a descriptive title with each epi curve. The epi curve, with its title and axes, should provide enough information to be completely self explanatory, as shown in Figure 8.
- The pre-epidemic period should always be included on the graph to illustrate the baseline number of cases.

### Using software packages to create epi curves

Epi curves can be easily made by hand or with a software package such as Microsoft Excel, Microsoft PowerPoint or Epi Info. If Excel is used, the easiest way to set up the data on a spreadsheet is shown in Table 1. Then follow these steps:

1. Click the "Chart wizard" on the tool bar.
2. Choose "Column" as the chart type.
3. Click "Next" twice and specify the chart options (e.g., label axes and title the graph).
4. Click "Next" and "Finish."

**Useful resource on basic principles of health-related investigations:**

- Print-based self-study course covering basic principles of epidemiology used in surveillance and investigation of health-related events can be downloaded at: http://www.phppo.cdc.gov/phtn/catalog/3030g.asp

At this point, the bars may not be touching each other as they should be, since this is a histogram. Right click on one of the bars, choose "Format data series" then "Options" and set the "Gap width" to "0" to get an epi curve that looks like Figure 9.

The idea is the same in Microsoft PowerPoint. Click on the "Insert Chart" command on the tool bar, clear the sample data shown on the Excel spreadsheet and insert your data. In PowerPoint it is easiest to format the data so that the first row in the spreadsheet is the date of onset and the second row is the number of cases.

Epi curves can also be made in Epi Info, a free software package developed by the CDC which can be downloaded for free at: http://www.cdc.gov/epiinfo. In Epi Info 2002, the most recent version, an epi curve can be made in the following way:

1. Click on the "Analyze data" button on the main menu, opening or importing the outbreak data.
2. Click on "Graph" from the menu bar under the "Statistics" section.
3. At this point, a graph box will pop up. Choose "Histogram" from the "Graph Type" option, the "Date/Time of illness onset variable" for the x-axis main variable, "Sum" from the "Show value of" option beneath the y-axis option and the "Number of cases variable" for the "Weight variable" beneath the y-axis option.
4. Type in the graph title where it says "Page title."
5. Click on "Ok."

An epi curve can also be created without the graph box, by typing in the command "graph" and various graph options. The Epi Info code necessary for doing this can be found by looking up "analysis code index" in the Epi Info help menu, and choosing "graph" from the options.

## Conclusion

Epi curves are useful for identifying the pattern of spread, magnitude, time trends, and exposure period for an outbreak. Because they are so useful, be sure to make an epi curve (or more than one) to help put the pieces of the puzzle together during your next outbreak investigation or during routine review of surveillance data.

**Things to remember when making an epi curve:**

1. Plot the number of reported cases on the y-axis.
2. Plot the date/time of symptom onset on the x-axis.
3. Choose an appropriate time interval for the x-axis, or try several to see which best represents the data.
4. Include pre-outbreak time on the x-axis to show the "baseline" disease level and to show visually when the outbreak began.
5. Label the x and y-axes clearly.
6. Give the epi curve a descriptive, self-explanatory title.
7. Include more detailed information, such as cases by geographic location or by symptom, if helpful.
8. To be technically correct, make the bars touch each other (unless there are periods of time with no cases, in which case there will be space between the bars).

**Glossary:**

**Epidemic curve:** A histogram that shows the course of a disease outbreak by plotting the number of cases by time of onset.

**Primary case:** The individual who introduces the disease into the family or group under study by acquiring the disease from the original exposure; not necessarily the first person diagnosed.

**Stratifying:** The process of separating a sample into several subsamples according to specific criteria, such as age groups or gender.

**Secondary cases:** Cases of infection that occur after exposure to the primary case.

**Index case:** The first case patient in a family or other defined group to come to the attention of the investigator.

**Incubation period:** A period of subclinical or inapparent pathologic changes following exposure, ending with the onset of symptoms of infectious disease.

**Outlier:** Observations differing so widely from the rest of the data as to lead one to suspect that a gross error may have been committed, or to suggest that these values come from a different population.

**Histogram:** A graphic representation of the frequency distribution of a continuous variable. Rectangles are drawn in such a way that their bases lie on a linear scale representing different time intervals and their heights are proportional to the frequencies of the values within each of the intervals.

**THE UNIVERSITY OF NORTH CAROLINA**

The North Carolina Center for Public Health Preparedness
The University of North Carolina at Chapel Hill
Campus Box 8165
Chapel Hill, NC 27599-8165

Phone: 919-843-5561
Fax: 919-843-5563
Email: nccphp@unc.edu

FOCUS Workgroup:

- Lorraine Alexander, DrPH
- Jill Koshiol, MSPH
- Pia D.M. MacDonald, PhD, MPH
- Gloria C. Mejia, DDS, MPH
- Sally B. Mountcastle, PhD, MSPH
- Amy Nelson, PhD, MPH
- E. Danielle Rentz, MPH
- Tara P. Rybka, BA
- Cheryl R. Stein, MSPH
- Michelle Torok, MPH
- Nicole Tucker, MPH
- Drew Voetsch, MPH

**REFERENCES:**

1. CDC. Botulism in Argentina foodborne outbreak investigation. http://www.phppo.cdc.gov/phtn/casestudies/computerbased/botarg.htm.

2. CDC. Principles of epidemiology, 2nd edition. Atlanta, GA: Public Health Practice Program Office, 1992.

3. Dwyer DM, Groves C. Outbreak epidemiology. In: Nelson KE, Masters Williams C, Graham NMH, eds. Infectious disease epidemiology: theory and practice. Gaithersburg, MD: Aspen Publishers, Inc, 2001: 119-148.

4. Arness MK, Feighner BH, Canhan ML. Norwalk-like viral gastroenteritis outbreak in U.S. Army trainees. Emerg Infect Dis 2000; 6 (2):204-207.

5. CDC. Norwalk-like viruses: public health consequences and outbreak management. MMWR 2001; 50 (RR09): 1-18.

6. Weber DJ, Menajovsky LB, Wenzsel R. Investigation of outbreaks. In: Weber D, Thomas J, eds. Epidemiologic methods for the study of infectious diseases. NY, NY: Oxford University Press, Inc, 2001: 291-310.

## UPCOMING TOPICS!

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

If you would like to receive electronic copies of Focus on Field Epidemiology, please fill out the form below:

- NAME: _____
- DEGREE (S): _____
- AFFILIATION: _____
- E-MAIL ADDRESS: _____
- May we email any of your colleagues? If so, please include their email here:

  _____

  _____

  _____

Please fax to: (919) 919-843-5563

or mail to: North Carolina Center for Public Health Preparedness
The University of North Carolina at Chapel Hill
Campus Box 8165
Chapel Hill, NC 27599-8165

Or go on-line: http://www.sph.unc.edu/nccphp/focus/

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