VOLUME 2, ISSUE 4

FOCUS on Field Epidemiology

Selecting a Study Design

A young epidemiologist stood with his old mentor, surveying the lunch buffet as if the fate of the world were in the balance. "You must choose wisely," the stoop-shouldered old man advised his companion. The young epidemiologist stepped toward the buffet counter and reached slowly toward the roast beef and then the green beans. He checked for the old man's reaction and received approval. Emboldened, the young man went on to take fried okra, mashed potatoes, macaroni and cheese, Jell-o salad, two dinner rolls, and banana cream pie. An hour later, he sat at the table clutching his stomach with empty dishes haphazardly strewn about.

"We have come to the lunch buffet for a reason, grasshopper," the old mentor's voice reached the young epidemiologist through the haze. "Choosing a good meal is like choosing a good study design: it must be well balanced, with good variety, but not so much variety that you feel nauseated. A good meal gives you the nutrients you need to live, and a good study design gives you the tools to test your hypothesis."

A previous FOCUS issue explained how to use information to generate a hypothesis (or educated guess) about the cause of an outbreak. But just as suspicion is not enough to warrant a conviction in a courtroom, you need evidence to support your hypothesis. In science, we use analytic studies to test hypotheses.

This issue of FOCUS describes different types of study designs that can be used to test a hypothesis. Which one to choose is up to you....

In field epidemiology, we call the specific cause of an outbreak the "exposure." We want to know whether there is an association between our hypothesized exposure and disease, how strong the association is, what proportion of cases are due to the exposure, and whether there is an increased risk of disease with increased exposure (also known as a dose-response relationship). These questions and others can be answered with the results of a well designed analytic study. Two commonly used types of analytic studies are the cohort study and the case-control study.

Cohort study

The term "cohort" comes from the Roman Empire; it was used to describe a group of about 300 to 600 soldiers, or the tenth part of a Roman Legion. In epidemiology, we think of a cohort as a group of people who have something in common. A cohort can represent the "source population," or the population from which the cases of disease arise. For example, a cohort could be all the employees in an office building, or everyone who attended a football game at a local high school, or all the residents of a neighborhood.

In outbreak situations, cohort studies tend to be "retrospective" stud-

CONTRIBUTORS

Author:

Amy Nelson, PhD, MPH

FOCUS Workgroup*

Reviewers:

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

The North Carolina Center for Public Health Preparedness is funded by Grant/Cooperative Agreement Number U90/CCU424255 from the Centers for Disease Control and Prevention. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the views of the CDC.

ies, which look at exposures that occurred in the past in relation to a disease that has already happened.

With a cohort study, the occurrence of disease in the exposed group (or the "risk" to the exposed) is compared to the occurrence of disease in the unexposed group (or the "risk" to the unexposed). This measure, known as a risk ratio, tells us not only whether the disease is associated with the exposure, but also how strong that association is.

To conduct a cohort study of an outbreak you must be able to identify every person in the cohort. This is possible if the group of people at risk for disease is small and well defined, like those at a wedding reception, on a cruise ship, in a school, or in a prison. You generally have the option of interviewing all members of the cohort, or if this group is too large, a sample of the cohort.

In certain situations, however, it may be difficult to define a suitable cohort. For example:

- For an investigation of disease outbreak related to a restaurant, can you find every single person who ate at the Main Street Deli on January 10-20?

  Probably not, though you might be able to identify everyone who ate there and paid by credit card. The problem is that customers who pay by cash and those who pay by credit card might order different items.

- For an investigation of a disease outbreak related to contaminated lunch meat distributed through a local supermarket chain, how would you locate every person buying and/or eating meat from the store?

  You might not be able to locate every person, but you might consider all customers who use special discount cards as a subset of the grocery store customers. These cards have the added benefit of electronically recording each item that was purchased.

As these examples show, identifying the entire cohort in an outbreak may be difficult in some instances. But thinking about the group that is affected (the source population) may help.

As an alternative, many investigators use the case-control study design when the cohort is not identifiable or would be difficult to identify and contact all members.

Case-control study

The most frequently used study type in outbreak situations is the case-control study (1). This is because they can be quickly implemented, and they can be used when a cohort study might be large and time-consuming.

With a case-control study, you identify people with the disease (case-patients) and then identify people without the disease (controls); then you ask everyone the same questions about past exposures. Case-control studies are valuable in outbreak situations because you already know who the sick people are — they were diagnosed by a doctor, had a positive lab culture, or identified by the health department. With a case-control study, you calculate an odds ratio to measure the strength of the association between the illness and the exposure(s). Odds ratios compare the odds of exposure among case-patients to the odds of exposure among controls to see whether and to what extent exposure levels differ. Risk ratios cannot be calculated in a case-control study.

The most difficult part of a case-control study is selecting the controls. Defining the source population (the population that gave rise to the cases) may help to narrow down potential control selection. Do the cases live in the same city, or did they attend the same event? Are they of a particular race or ethnicity? Understanding where the cases came from will help you select your controls.

Case-control studies should always be thought of in the context of the source population. This makes identification of controls easier, because controls are a sample of people from the source population.

- Referring to our earlier example, an outbreak of gastrointestinal illness was linked to eating at the Main Street Deli during January 10-20. A case-control study was conducted to identify which food product caused the outbreak. Cases were recruited from people who ate at the Deli and experienced vomiting. Controls were recruited from people who ate at the Deli but did not experience vomiting. While all known cases were recruited into the study, only a portion of the healthy controls were contacted because there was no way to identify every single person who had eaten at the restaurant during these 10 days.

In this example, the source population was the group of people who ate at the Main Street Deli during January 10-20. We want to know what foods the case-patients ate and what the controls (those who did not get sick) ate. When selecting controls, we should select them from the customers who ate at the Deli during the time period of interest.

Case-control or cohort: Which one is right?

Cohort studies and case-control studies each have distinct advantages and disadvantages (see Table 1). The choice of study design depends on the situation. A good rule of thumb is always to think about the source popula-

Table 1. Selected characteristics of cohort studies and case-control studies

Characteristic	Cohort Study	Case-Control Study
Relationship to source population	Recruits the entire cohort or takes a representative sample	Recruits those in the source population who became ill (all or a sample) and a sample of those who did not
Definition of the study subjects	Based on exposure status	Based on disease status
Exposure/disease measurement	Measures disease status among the exposed, and compares this to disease status among the unexposed	Measures exposure status among the diseased, and compares this to exposure status among the non-diseased
Measure of association	Risk ratio	Odds ratio
Cost	Can be higher	Can be lower
Advantages	Can estimate risk of disease in the source population and among subgroups	Can be used when the source population is difficult to define, when it is difficult to contact all members of a cohort, or when a cohort study would be too large and costly
Disadvantages	Can be time and labor intensive	Cannot estimate risk of disease in the population

tion. If the members of the group are easily identifiable, such as attendees to a luncheon or wedding reception, and you are able to interview all of them or a representative sample of them, then a retrospective cohort study may be the best approach. If there is no obvious connection among the cases, making the cohort difficult to identify, or the cohort is too large to contact all members, then a case-control study may be the best choice.

In the situation depicted in Figure 1 below, a retrospective cohort study is the most appropriate study design. Generally, if there are fewer than 200 people involved, you should consider interviewing everyone.

When the cohort gets much larger but is still easily identifiable, a case-control study can be used, as in Figure 2. Another strategy might be to identify cohorts within the larger cohort.

- For example, in an outbreak of food poisoning on a college campus, rather than interviewing everyone on campus consider a cohort study within a single dormitory.

Figure 1: Easily identifiable cohort (e.g., a church picnic, wedding, or luncheon)

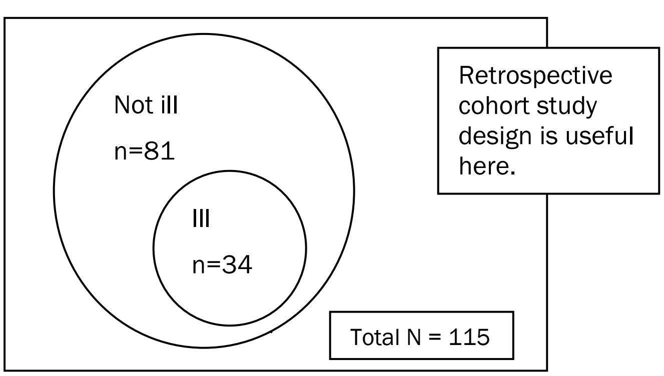
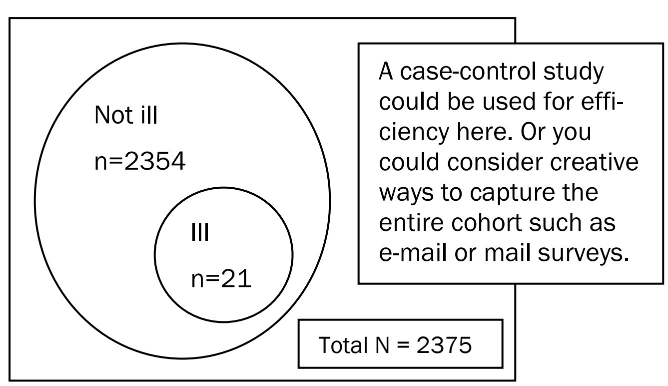
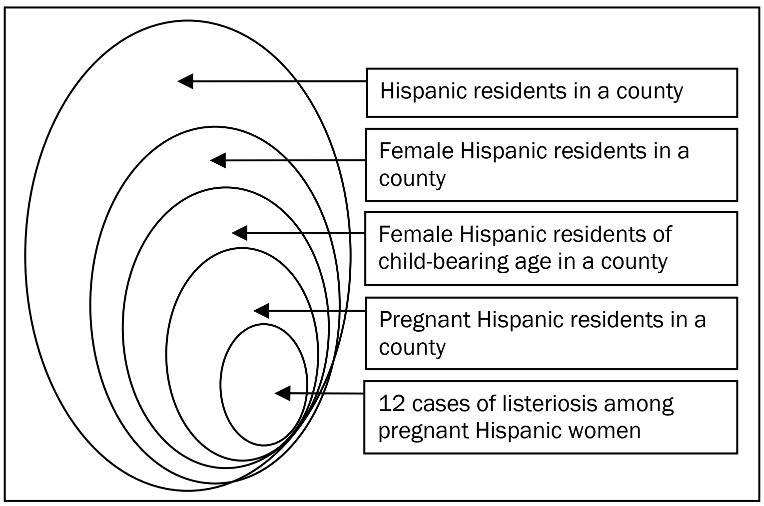


Figure 2: Easily identifiable but large cohort (e.g., a cruise ship or college campus)



The scenario in Figure 3 requires more thought, but is still based on identifying cohorts. Consider an outbreak of listeriosis in North Carolina from 2000 (2), in which pregnant Hispanic women in one county were primarily affected. If we sampled controls who were not Hispanic, foods that are associated with a Hispanic diet would be implicated. If we selected controls who were not pregnant, factors related to the immunocompromising state of pregnancy would be implicated. Drilling down to the smallest practical identifiable source population is the most efficient approach to the selection of controls. In this case, pregnant Hispanic women (during a specific time frame at a defined place) were the source population.

Figure 3: Selecting controls for a rare disease in a large cohort: case-control design and efficient selection of controls.



Case Studies

Yersinia and chitterlings

From November 15, 2001, through February 15, 2002, 12 cases of *Yersinia enterocolitica*, an infection that causes gastroenteritis, were identified at a large urban pediatric emergency department in Tennessee (3). All case-patients were black infants, but a cohort study of all black infants in the city would not have been useful (or possible). The source population was black infants who had access to medical care, who were from the population served by that urban hospital. Therefore, the controls chosen were black infants who presented to the emergency department of the same urban hospital with any chief complaint other than gastroenteritis. This represents one method of sampling from among the cohort of black infants in Tennessee. Previous seasonal increases in yersiniosis among this population had been found to be associated with household exposure to chitterlings (pork intestines). The results of the analysis showed that chitterlings had been prepared in 100% of the case households but only 35% of the control households, and parents were usually able to identify ways that the kitchen might have become contaminated (e.g., chitterlings were cleaned in the sink). Thus, the case-control study clearly implicated the source of the outbreak.

Pseudomonas from ear piercing

In September 2000, an Oregon physician reported treating 2 patients on 2 consecutive days with infections of the cartilage of the ear (4). Both patients had received ear piercings at the same kiosk. Because the investigators could contact all patrons of the kiosk, they approached this outbreak from a cohort perspective. Using records from the kiosk, the investigators constructed a cohort of customers who received piercings from August 1 through September 15. In total, 118 people had experienced 186 piercings. Seven piercings (4%) resulted in laboratory-confirmed *Pseudomonas aeruginosa* and 18 piercings (10%) resulted in a suspected case (localized pain or swelling and drainage of blood or pus for more than 14 days). Risk of infection was increased if piercing was in the cartilage rather than the lobe of the ear. The investigators were able to 1) determine the risk of infection among the entire population, 2) determine that the risk was different based on site of piercing, and 3) identify practices that might have led to contamination of equipment and subsequent infection.

Conclusion

Cohort studies and case-control studies are both options for determining the cause of an outbreak. In both types of studies, the investigator is studying the source population that gives rise to the cases of disease. In a cohort study, the entire population may be used to investigate the cause of disease, or a representative sample of the population may be used. In a case-control study, all or most cases of disease are captured, and controls are sampled. Both types of studies are effective; which one you choose will depend on the circumstances of the outbreak you are investigating.

CONTACT US:

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

REFERENCES:

1. Dwyer DM, Strickler H, Goodman RA, Armenian HK. Use of case-control studies in outbreak investigations. *Epidemiol Rev.* 1994;16(1):109-123.

2. MacDonald PM, Whitwam RE, Boggs JD, et al. Outbreak of Listeriosis among Mexican Immigrants as a Result of Consumption of Illicitly Produced Mexican-Style Cheese. *Clin Infect Dis*. 2005; 40:677-682.

3. Jones TF. From pig to pacifier: chitterling-associated yersiniosis outbreak among black infants. *Emerg Infect Dis.* 2003;9(8):1007-1009.

4. Keene WE, Markum AC, Samadpour M. Outbreak of Pseudomonas aeruginosa infections caused by commercial piercing of upper ear cartilage. *Jama.* 2004;291(8):981-985.

FOCUS Workgroup:

- Lorraine Alexander, DrPH
- Anjum Hajat, MPH
- Pia D.M. MacDonald, PhD, MPH
- Gloria C. Mejia, DDS, MPH
- Sandi McCoy, MPH
- Amy Nelson, PhD, MPH
- E. Danielle Rentz, MPH
- Tara P. Rybka, MPH
- Cheryl R. Stein, MSPH
- Michelle Torok, MPH
- Drew Voetsch, MPH

UPCOMING TOPICS!

- Introduction to Forensic Epidemiology
- Differences between Public Health and Law Enforcement Investigations
- Environmental Health Assessment
- Traceback Investigations
- Risk Communication

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 e-mail any of your colleagues? If so, please include their e-mail addresses 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 online: http://www.sph.unc.edu/nccphp/focus/

We are on the web!

http://www.sph.unc.edu/nccphp