VOLUME 3, ISSUE 1

FOCUS on Field Epidemiology

CONTRIBUTORS

Author:

Amy Nelson, PhD, MPH

Kim Brunette, MPH

FOCUS Workgroup*

Reviewers:

FOCUS Workgroup*

Production Editors:

Tara P. Rybka, MPH

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.

Cohort Studies for Outbreak Investigations

In a previous issue of FOCUS, we introduced cohort studies and talked about how to decide whether a cohort study is the best design for your situation. Remember that cohort studies are useful when there is a defined population at risk for developing the disease of interest (such as all members of the Glee Club) and when it is possible to interview all members or a representative sample of the cohort.

In outbreak investigations, cohort studies are usually retrospective because in an outbreak an exposure has already occurred, and enough cases to signal an outbreak have also occurred. The aim is to determine what exposures occurred in the past to cause these cases of disease.

In this issue we cover the basics of conducting a cohort study and discuss how to calculate measures of disease (prevalence, risks) and disease association (relative risk).

Establishing the Cohort

There are two ways to establish a cohort. One way is to choose cohort members based on characteristics that assume exposure has occurred. This is often done in occupational studies. For example, an office building, school, or neighborhood may be considered exposed if it is close to contamination sources of interest, such as landfills, nuclear plants, or factories; it can be considered unexposed if it is located far away from these sources of contamination.

To be certain that an exposure caused the disease, the unexposed group must be similar to the exposed group in all respects except the exposure. Therefore they should have similar demographic characteristics such as age, gender, race, and income. If you compare an "unexposed" elementary school to an "exposed" elementary school, both schools should be located in neighborhoods with the same types of economic and other opportunities, similar socio-economic status (SES), similar parent involvement, similar school resources, and similar student populations.

Using groups that have other differences may lead to "confounding." If you compare a low-income public school to a private boarding school, you will find many differences between the two, and you will not know whether the difference in disease outcome is due to an exposure, or whether it is due to the other differences.

Another way to establish a cohort is to identify a particular population group and then determine whether or not they were exposed.

- People who happened to be at the same place at the same time (for example, all customers at a deli during a certain period) might be considered a population group. You could characterize them by whether they ate the ham, potato salad, or any other food that you want to evaluate as an exposure.
- A population group can also be defined by attendance at an event, such as a summer concert.

These people can be characterized by whether or not they ate the beef on a stick, played in the duck pond, etc.

- Membership in a particular group, such as Rotarians, the PTA, or the triathlon club, can also define a population. For example, athletes from all over the world attended Eco-Challenge-Sabah 2000, a multi-sport race in Borneo, Malaysia. Reports of disease from health departments across the United States triggered an investigation, and exposure to the Segama River was implicated. (1)

People who all belong to the same group or attend the same event are likely to be very similar to each other. Therefore, confounding may not be a major issue in these cohorts.

Conducting the Investigation

The exposure that is causing an outbreak is not always known at the beginning of the investigation, so investigators attempt to measure a number of plausible exposures and evaluate each one.

For example, investigators might classify exposures as people who did or did not eat at a particular restaurant, people who did or did not use recreational swimming facilities, people who did or did not get ice from the hotel ice machine, or people who did or did not eat the potato salad at the church picnic.

Next, investigators will develop questionnaires and interview members of the cohort to gather demographic information and exposure to any potential risk factors for disease. (Information on questionnaire development and interviewing techniques is available in past issues of FOCUS.) Investigators will also determine which cohort members meet the case definition, and then analyze this information to determine whether there is a relationship between exposure and disease.

After you swoop down and gather the data with perfect technique and FOCUS flair, you can present your data to the world and save lives… But first, you must analyze the data to determine what caused the outbreak, how to prevent further illness, and how to prevent future outbreaks.

Analyzing the Data

Oh, no! Not math! Never fear, the basic analyses in a cohort study are simpler than those for any other study design.

Prevalence

The simplest descriptive measure of disease is **prevalence.** Prevalence is the number of ill people divided by the total population at risk (i.e., the cohort) at *a particular point in time.* Prevalence is often expressed as a percent.

- In 1993, the prevalence of chronic fatigue syndrome among patients attending a primary care physician was 0.3%. (2)

Risk

Another measure of disease is **risk**, which reflects the probability of acquiring disease. Risk is the number of cases divided by the total number of people in the population (including cases and non-cases).

$$\text{Risk} = \frac{\text{\# ill people}}{\text{\# people at risk}}$$

- The risk of acquiring HIV from a blood transfusion in the U.S. is approximately 0.0002%. (3)

In an outbreak setting, we like to know the risk of disease among those exposed to various risk factors. Knowing the risk helps us determine how people are getting sick and how to prevent more cases from occurring.

You can calculate risk in a cohort study, because you know the number of people who are at risk of developing the disease—they are all in your cohort. You cannot calculate risk in a case-control study because only a sample of people who are at risk of developing disease are included in the study. You may not even know the total number of people at risk in a case-control study.

In infectious disease epidemiology, risk is also called the **attack rate**, which is the percentage of people acquiring a disease among a group.

- An influenza epidemic in a nursing home had an attack rate of 65% (43 of 66 residents became ill). (4)

Risk can be calculated separately for those who are members of the group, and for those who are not members of the group. The ratio of these two numbers is the **risk ratio** (RR), or the **relative risk** (the risk of one group relative to the risk of another group).

$$\text{Risk ratio} = \frac{\text{risk in exposed group}}{\text{risk in non-exposed group}}$$

Risk and Risk Ratios: 2x2 table

	Ill	Not Ill	Total	Risk (Attack Rate)
Exposed	a	b	a+b	a/(a+b)
Not Exposed	c	d	c+d	c/(c+d)

Risk Ratio = [a/(a+b)] ÷[c/(c+d)]

To interpret the risk ratio, we compare the value of the RR to 1. If risks in both groups are the same, the ratio of the risks will be 1, indicating that there is no association between the exposure being evaluated and the risk of disease.

If RR = 1, exposure has no association with disease.
If RR > 1, exposure may be positively related to disease.
If RR < 1, exposure may be inversely related to disease.

- In an outbreak of histoplasmosis in a high school, the risk ratio for students in classrooms near the courtyard during rototilling was 1.3, meaning that the risk of illness for students near the courtyard was 1.3 times the risk of illness for students not near the courtyard. (5)

Practice Calculating Risk

To demonstrate how to calculate risk and risk ratio, let's work through an example.

Daycare X had a special activities day when all 61 children who attended the daycare were taken to the zoo. Boxed lunches were catered for the children to eat at the zoo. The night after the zoo trip, children begin getting sick. Over the next few days, several children became ill.

After 6 of the children were culture-confirmed with *Salmonella* Enteritidis, a case was defined as any child or adult attending the Zoo Day trip of Daycare X presenting with diarrhea, abdominal cramps, and/or fever within 72 hours of the trip. A total of 27 children met this case definition.

The overall risk of illness among children:

$$\frac{\#\ \text{ill}}{\text{total \# children}} = \frac{27}{61} = 0.44 = 44\%$$

All children and adults attending the Zoo Day were asked which animal exhibits they visited, whether they participated in the petting zoo, and what lunch and snack items they ate.

We obtained the following information on selected exposures from the questionnaire data:

Table 1. Selected exposures from children attending Daycare X Zoo Day

Exposure	Ill (n=27)	Not Ill (n=34)
Turkey sandwich	21	14
Fruit salad	10	30
Chips	13	17
Petting zoo	17	15

Since we can see that many of the sick children ate the turkey sandwich, let's focus on that exposure.

- 35 of the 61 children reported eating at least part of their turkey sandwich. This is the EXPOSED group.
- The other 26 children reported NOT eating any of the sandwich. This group is the UNEXPOSED group.
- 21 of 35 exposed children became ill, and 6 of the 26 unexposed children became ill. This kind of information is often gathered into a table, such as Table 2, to show the ill and the exposed.

Table 2. 2x2 table showing exposure to the turkey sandwich by illness status.

Exposure	Ill	Not Ill	Total
Turkey Sandwich	21 (60%)	14 (40%)	35
No Turkey Sandwich	6 (23%)	20 (77%)	26
Total	27	34	61

Now we can calculate the risk among the children who were exposed and not exposed to the turkey sandwich. Each of these risks can also be referred to as the attack rate.

The risk of illness among those exposed to turkey:

$$\frac{\#\ \text{ill exposed}}{\text{total \# children exposed}} = \frac{21}{35} = 0.60 = 60\%$$

The risk of illness among those NOT exposed to turkey (risk among the unexposed):

$$\frac{\#\ \text{ill unexposed}}{\text{total \# children unexposed}} = \frac{6}{26} = 0.23 = 23\%$$

We can see that the risk (or attack rate) of illness among those exposed to turkey (60%) was greater than among those unexposed to turkey (23%). To put a number on how much greater the risk was in the exposed, we obtain the risk ratio.

The risk ratio:

$$\frac{\text{risk among the exposed}}{\text{risk among the unexposed}} = \frac{0.60}{0.23} = 2.61$$

A RR of 2.61 shows that the risk of acquiring *Salmonella* among those who ate turkey was 2.61 times the risk of acquiring *Salmonella* among those who did not eat turkey.

Is a RR estimate of 2.61 a strong association? The strength of an association can be determined using confidence intervals and other statistical methods that will be discussed in a future issue of FOCUS.

But, wait…If the turkey sandwich was responsible for disease, why were there cases among the not exposed?

Several factors may have contributed to the illness among those not exposed to the turkey sandwich. First, since we conducted interviews after the fact, people (and especially children) may have forgotten that they ate the turkey sandwich. Also, cross-contamination may have occurred either during food preparation or while the children were eating. We could also examine when children became ill. If children who were exposed became ill earlier in the outbreak, while unexposed children became ill later in the outbreak, secondary transmission could have occurred between children at the daycare. This could appear to reduce the apparent role of the turkey sandwich. Finally, unexposed children could have become ill by chance; that is, they would have gotten sick regardless of Zoo Day.

So have you found the culprit?

Even after we find an association between an exposure and disease, we should examine other potential exposures to see if there are other significant associations.

Next, we should attempt to find the source of contamination. If you can trace a foodborne bacteria trail to its source, you have a more definitive cause of infection and can intervene to promote safer food practices. A future issue of FOCUS will describe how to conduct a traceback investigation.

Examples of Cohort Studies for Outbreak Investigations

Gastroenteritis at a tourist resort

"In July 2000, an outbreak of gastroenteritis occurred at a tourist resort in the Gulf of Taranto, southern Italy. Illness [was identified] in 344 people, 69 of whom were staff members. . . .Because of the high number of cases [that occurred] in resort staff members, a retrospective cohort study was performed to assess risk factors associated with illness in this group." Laboratory testing and an environmental investigation were also undertaken. (6)

Persons eligible for the [cohort] study were staff members employed at the resort from July 1 to 31. Standard questionnaires were sent to all 224 staff members in the first week of August. Information requested included. . . [demographic information], date of onset and type of symptoms, and water and food preferences." Investigators did not inquire about specific food histories or exposures, as a month had elapsed between onset of symptoms and distribution of the questionnaire. (6)

Questionnaires were returned by 181 of the 224 staff members. "The attack rate in this group was 69 (38.1%) of 181." The highest attack rates were observed in waiters, sports trainers, entertainers, and cleaning staff. Relative risks were significant only for exposure to beach showers (RR=1.8) and consuming drinks with ice (RR=1.8). (6)

Laboratory testing was conducted on 30 samples (28 stool and 2 vomit specimens). Norwalk-like virus was found in 22 of 28 stool specimens. "[The] environmental investigation identified a breakdown in the water system of the resort, and tap water samples from different places in the resort showed contamination with fecal bacteria." (6)

In this study, the investigators did not try to identify and interview all the persons who had spent time at the resort; instead, they chose to focus on the staff cohort, which was a more manageable group to study.

Foodborne outbreak at a restaurant

In December 2000 and January 2001, local health authorities in southwest Germany were contacted by several ill persons about "protracted, sometimes relapsing gastroenteritis symptoms." The callers were part of four independent parties of 6, 7, 7, and 20 persons who attended luncheons at a particular restaurant. Laboratory testing identified *Cyclospora cayetanensis* in 9 specimens. (7)

All 40 attendees at the luncheons were asked to participate in a cohort study to identify the cause of the illness. A questionnaire that included "questions about age, gender, travel history, food items and beverages consumed at the luncheons, onset and duration of symptoms, physician consultation, examination of stool samples, antibiotic treatment, and days absent from work ." Thirty questionnaires were returned. Of the 30, 26 persons met the clinical case definition, for an attack rate of 87%. (7)

"Frequencies, relative risks, and 95% confidence intervals were calculated for 12 main courses, 3 side dishes, 12 beverages, and 2 desserts. . . . The only food item that showed a statistically significant association with disease" was a side salad dish (RR=5). No samples of the salad were available for laboratory testing. A traceback investigation of the sources of the salad contents were conducted, but did not identify any source of contamination. (7)

Investigators restricted the cohort study to the four independent parties who ate at the restaurant. Had the investigators not focused on these four parties, they would have needed to conduct a case-control study, because it is unlikely that they would have been able to identify every person who had eaten at the restaurant.

Gastroenteritis on a cruise ship

"On July 16, 2004, DEC [Alaska Department of Environmental Conservation] notified the Alaska Section of Epidemiology of several cases of gastroenteritis among passengers on a cruise ship (78-passenger capacity) that was sailing in Prince William Sound." Investigators received an additional report of a Nevada resident whose laboratory-confirmed case of *Vibrio parahaemolyticus* gastroenteritis had started while on the same ship. The patient reported having eaten raw oysters served on board the ship the day before the onset of the illness. . . .[Officials] performed a retrospective cohort study on passengers from four July 2004 cruises to determine the burden of gastrointestinal illness among passengers and risk factors for illness." (8)

"[Investigators] administered a questionnaire by telephone or in person to all the passengers [who could be] contacted, recording demographic information, characteristics of the illness, and information about the food consumed on board the ship." Attack rates and risk factors for illness were calculated. "Of 189 passengers in the cohort, 132 (70%) were interviewed." Twenty-two of the 132 passengers met the case definition of the illness, an attack rate of 17%. "The attack rate for persons who ate oysters was 29% (of 48 persons who consumed oysters, 14 became ill)." (8)

Laboratory testing confirmed *V. parahaemolyticus* in 18 isolates (8 human and 10 oyster). Environmental investigation found that rising temperatures of ocean water seem to have contributed to this outbreak, as this investigation extended by 1000 km the northernmost documented source of oysters that caused illness due to *V. parahaemolyticus*. (8)

This study was interesting in that investigators chose to include 4 cruises in their cohort as opposed to just the cruise where ill passengers were reported. This was probably due in some part to the report of vibrio illness in a passenger from a previous cruise.

Methicillin-resistant Staphylococcus aureus *(MRSA) outbreak among player on a football team*

In September 2003, the Connecticut Department of Public Health was notified "about a cluster of skin infections due to MRSA among members of a college football team." Investigators "performed a retrospective cohort study of members of the 2003 football team" to assess skin injuries, hygienic practices, other exposures, and known risk factors for MRSA "during the interval between the arrival at football camp and the announcement of the outbreak." (9)

Ninety of the 100 players completed face-to-face interviews. Ten met the case definition for MRSA (attack rate=10%). Cornerback defensive backs (RR=17.5) and wide receivers (RR=11.7) had the highest infection risk, "accounting for 8 of 10 case patients. . . .Players who sustained turf burns had a risk of infection that was 7 times higher than that for players without turf burns" (RR=7.2). Additionally, "players who reported body shaving were 6.1 times more likely to develop MRSA infections." (9)

"[Investigators hypothesized] that MRSA was spread between players largely by frequent direct contact between cornerbacks and wide receivers during practice scrimmage and drills. . . .Infection appear[ed] to have been facilitated by interruptions of skin integrity, including turf burns and microabrasions likely sustained while body shaving." (9)

This investigation was a typical cohort study. The cohort was comprised of an easily identified population, the football team.

Glossary:

Prevalence: The number of ill people divided by the total population at risk at one point in time.

Risk: The probability of acquiring a disease; the number of cases divided by the total number of people in the population (both cases and non-cases). Also known as attack rate in infectious disease epidemiology.

Risk Ratio (RR): The ratio of risk among people who are members of a group and risk among people who are not members of the group. Also known as relative risk.

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

FOCUS Workgroup:

- Lorraine Alexander, DrPH
- Kim Brunette, MPH
- Anjum Hajat, MPH
- Pia D.M. MacDonald, PhD, MPH
- Gloria C. Mejia, DDS, MPH
- Sandi McCoy, MPH
- Amy Nelson, PhD, MPH
- Tara P. Rybka, MPH
- Michelle Torok, MPH
- Rachel A. Wilfert, MD, MPH

REFERENCES:

1. Centers for Disease Control and Prevention. Update: outbreak of acute febrile illness among athletes participating in Eco-Challenge Sabah 2000 — Borneo, Malaysia, 2000. *MMWR Morb Mortal Wkly Rep*. 2001;50:21-24.
2. Bates D, Schmitt W, Buchwald D, et al. Prevalence of fatigue and chronic fatigue syndrome in a primary care practice. *Arch Intern Med.* 1993;153:2759-2765.
3. Centers for Disease Control and Prevention. How safe is the blood supply in the United States? Available at: http://www.cdc.gov/hiv/pubs/faq/faq15.htm. Accessed December 5, 2005.
4. Infuso A, Baron S, Fauveau H, et al. Value of influenza vaccine during an outbreak of influenza A in a nursing home, Pyrénées Atlantiques, France, November-December 1995. *Euro Surveill*. 1996;1(5):35-37.
5. Chamany S, Mirza SA, Fleming JW, et al. A large histoplasmosis outbreak among high school students in Indiana, 2001. *Pediatr Inf Dis J*. 2004; 23:909-914.
6. Boccia D, Tozzi AE, Cotter B, et al. Waterborne outbreak of Norwalk-like virus gastroenteritis at a tourist resort, Italy. *Emerg Infec Dis*. 2002;8:563-568.
7. Doller PC, Dietrich K, Filipp N, et al. Cyclosporiasis outbreak in Germany associated with the consumption of salad. *Emerg Infec Dis*. 2002;8:992-994.
8. McLaughlin JB, DePaola A, Bopp CA, et al. Outbreak of *Vibrio parahaemolyticus* gastroenteritis associated with Alaskan oysters. *N Engl J Med*. 2005;353:1463-1470.
9. Begier EM, Frenette K, Barrett NL, et al. A high-morbidity outbreak of methicillin-resistant *Staphylococcus aureus* among players on a college football team, facilitated by cosmetic body shaving and turf burns. *Clin Infect Dis*. 2004;39:1446-1453.

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/

UPCOMING TOPICS!

- Case-Control Studies
- Conducting Traceback Investigations
- Conducting Environmental Health Assessments
- Basics of Data Analysis

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