Session II
Study Design

Session Overview

• Developing and testing hypotheses
• Study Designs:
  – Selection
  – Implementation
• Sampling
Learning Objectives

• Understand the differences in methodology between various study designs
• Be able to describe the advantages and disadvantages of alternative study designs
• Know how to assess which study design to apply during an outbreak investigation
• Understand how to select cases and controls in a case-control study design
• Understand sampling

Basic Steps of an Outbreak Investigation

1. Verify the diagnosis and confirm the outbreak
2. Define a case and conduct case finding
3. Tabulate and orient data: time, place, person
4. Take immediate control measures
5. Formulate and test hypothesis
6. Plan and execute additional studies
7. Implement and evaluate control measures
8. Communicate findings

Exposure and Outcome

A study considers two main factors: exposure and outcome

• **Exposure** refers to factors that might influence one’s risk of disease
  – Smoking
  – Eating at a particular restaurant
• **Outcome** refers to case definitions
  – Individuals who do and do not have the disease/condition of interest
Developing Hypotheses

• A hypothesis is an educated guess about an association that is testable in a scientific investigation

• Descriptive data provide information to develop hypotheses

• Hypotheses tend to be broad initially and are then refined to have a narrower focus

Example

• Hypothesis: People who ate at the church picnic were more likely to become ill
  – Exposure is eating at the church picnic
  – Outcome is illness - diarrhea and fever, where diarrhea is defined as at least 3 soft stools in a 24 hour period

• Hypothesis: People who ate the egg salad at the church picnic were more likely to have laboratory-confirmed Salmonella
  – Exposure is eating egg salad at the church picnic
  – Outcome is laboratory confirmation of Salmonella

Analytic Studies

• Used to test the current hypothesis:
  – Is there an association between exposure and disease?
  – How strong is the association?
Analytic Studies

Two types used in outbreak investigations
  – Cohort
  – Case-control

Definition of a Cohort

In epidemiology, “Any designated group of individuals who are followed or traced over a period of time.”


Cohort Study Types

A cohort study analyzes an exposure / disease relationship within the entire cohort.

• Prospective
  – The Framingham Study

• Retrospective
  – Usually used in outbreak investigations
### Cohort Studies

#### Study Population
- Exposed
- Non-exposed

- Exposure is self selected
- Follow through time

<table>
<thead>
<tr>
<th>Disease</th>
<th>No Disease</th>
<th>Disease</th>
<th>No Disease</th>
</tr>
</thead>
</table>

### Cohort Study

- Identify cohort
  - Do not select cohort so that either everyone is exposed or everyone is diseased

### Cohort Studies: Prospective vs. Retrospective

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prospective</td>
<td>Assessed at beginning of study</td>
</tr>
<tr>
<td>Retrospective</td>
<td>Assessed at some point in the past</td>
</tr>
</tbody>
</table>
Cohort Study

• Preferred study design when:
  – Members of cohort are easily identifiable
  – Members of a cohort are easily accessible
  – Exposure is rare
  – There may be multiple diseases involved

Cohort Study Example

• Recent norovirus outbreaks on cruise ships
• Attempt to interview all passengers
• Collect food history information

Cohort Study Examples

• Shigellosis among swimmers in a Georgia park
  – Used park registry to identify park visitors

• Whirlpools and Methicillin-Resistant Staphylococcus aureus
  – Occurred on a college football team
Case-Control Study

- Sometimes, identifying a cohort is difficult
  - Members of cohort can’t be identified / contacted
- Case-control study is alternative

---

### Case-control Studies

<table>
<thead>
<tr>
<th></th>
<th>Had Exposure</th>
<th>No Exposure</th>
<th>Had Exposure</th>
<th>No Exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Study Population**

---

**Case-Control Study**

Steps in a Case-Control Study:

1. Identify the source population
2. Establish a case definition and select cases
3. Select controls
Case-Control Study

Step 1 - Identify source population
• Represents the population that the cases came from; is similar to the cohort in a cohort study

Case-Control Study

Step 2 – Establish a case definition and select cases
• A standard set of criteria for deciding disease status
  – Clinical criteria, time, place, and person

Case-Control Study

Step 3 – Select controls
• Represent source population
• Collect same exposure information as for cases
**Case-Control Study**

Step 3 – Select controls (cont’d.)

- Sources of controls
  - Random sample
  - Friends of cases

---

**Cohort versus Case-Control**

<table>
<thead>
<tr>
<th></th>
<th>Cohort Study</th>
<th>Case-Control Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferred Study</td>
<td>Members are easily identifiable</td>
<td>Identifying cases in the cohort would be too costly or time consuming</td>
</tr>
<tr>
<td>Design When . . .</td>
<td>Members are easily accessible</td>
<td>Accessing controls is too costly or time consuming</td>
</tr>
<tr>
<td></td>
<td>Exposure is rare</td>
<td>Members are rare</td>
</tr>
<tr>
<td></td>
<td>These may be multiple diseases involved</td>
<td></td>
</tr>
<tr>
<td>Study Group</td>
<td>Exposed persons</td>
<td>Persons with illness (cause)</td>
</tr>
<tr>
<td></td>
<td>Comparison Group</td>
<td>Unexposed persons</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Persons without the illness (controls)</td>
</tr>
</tbody>
</table>

---

**Study Design**

<table>
<thead>
<tr>
<th>Study Design</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohort</td>
<td>1. Least prone to selection bias</td>
<td>1. Prospective can be expensive, time-consuming</td>
</tr>
<tr>
<td></td>
<td>2. Can reasonably conclude that cause preceded disease</td>
<td>2. Prospective can lead to loss to follow up</td>
</tr>
<tr>
<td></td>
<td>3. Can study several diseases at once</td>
<td>3. Exposed may be followed more closely than unexposed, yielding invalid</td>
</tr>
<tr>
<td></td>
<td>4. Can examine rare exposures</td>
<td>conclusions about causality</td>
</tr>
<tr>
<td></td>
<td>5. Retrospective can be low-cost</td>
<td></td>
</tr>
<tr>
<td>Case-Control</td>
<td>1. Less expensive and quicker than cohort</td>
<td>1. Inefficient for studying rare exposures</td>
</tr>
<tr>
<td></td>
<td>2. Can examine the effect of multiple exposures</td>
<td>2. Susceptible to selection bias</td>
</tr>
<tr>
<td></td>
<td>3. Require a smaller sample population</td>
<td>3. Cannot directly estimate the risk of disease</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Cannot study several diseases at once</td>
</tr>
</tbody>
</table>
Matching in Case-Control Studies

- Makes one or more case and control attributes similar (e.g., age, gender, residence)
- An unmatched study design is usually preferred

Matching: Points to Consider

- More complex data analysis required
- Inability to assess role of matching factor on disease status
  - Do not match on exposure factor
- Potential for over-matching

Sampling

Sampling is the systematic selection of a portion of the larger source population. A sample should be representative of the larger source population.
Sampling

Why sample?

Because it is more efficient – saves time and money!

Sampling

Sample size

Is the purpose of the study to determine the source of the outbreak?
- A small number of cases and controls can reveal risk factors for infection.

Is the purpose of the study to determine the number of persons who become sick over a specific period of time?
- A cohort study would require a larger sample.

Sampling

Types of sampling

Simple random sample (SRS)
Randomly select persons to participate in study. There are many variations of SRS.

Convenience sample
Choose those individuals who are easily accessible.
Sampling
Problems with convenience sampling

- Based on subjective judgment
- Cases may or may not be representative of the total population
- May lead to biased results

Session II Summary

- An analytic study is used to test scientific hypotheses that may help support actions for specific control measures and to help prevent recurrence of a problem.

- A case definition with specific criteria helps you select your study population, as long as it does not include the hypothesis.

- Case-control studies, when conducted properly, are generally adequate and usually more efficient than cohort studies.

Session II Summary

- Cohort studies may be preferable when you work with confined (e.g., easily identifiable and accessible) study populations such as on a cruise ship or at a wedding reception.

- Case-control study controls need to be representative of the source population, and not matched on the exposure factor if matching is used.
References and Resources


- Centers for Disease Control and Prevention "Gastroenteritis at a University in Texas" http://www.phppo.cdc.gov/casestudies/classroom/gastro.htm


