Principles of Field Epidemiology

JANA RUSH, MPH, MA

DIRECTOR, OFFICE OF COMMUNICABLE DISEASE SURVEILLANCE & EPIDEMIOLOGY

CLEVELAND DEPARTMENT OF PUBLIC HEALTH
Learning Objectives

- Define epidemiology and its’ application to public health
- General knowledge of the history of epidemiology
- General knowledge of the basic steps involved in an outbreak investigation
- Understand common mathematical measures used by epidemiologists
Epidemiology is the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems.
Concept Check #1

- Graph the number of cases of congenital syphilis by year for the country
  - A. Distribution
  - B. Determinants
  - C. Application

- Recommend that close contacts of a child recently reported with meningococcal meningitis receive Rifampin
  - A. Distribution
  - B. Determinants
  - C. Application

- Compare food histories between persons with *Staphylococcus* food poisoning and those without
  - A. Distribution
  - B. Determinants
  - C. Application
History of Epidemiology

- Circa 400 B.C.-Hippocrates
  - Proposed how behaviors might influence the development of disease
- 1662-John Graunt
  - Published the 1st study quantifying birth, death, and disease occurrence
- 1800-William Farr
  - Considered the father of modern vital statistics and surveillance
- 1854-John Snow
  - Published a study linking cholera outbreaks to local water sources
- 19th and 20th Century
  - Studies extended to include chronic disease, injury, and violence
Primary Purposes of Epidemiology

- Assessing the community’s health
- Identify new and emerging diseases
- Monitor and track existing diseases
- Evaluate the effectiveness of control measures
Core Epidemiological Functions

- Public health surveillance
- Field investigation
- Analytic studies
- Evaluation
- Linkages
- Policy development
Concept Check #2

- Reviewing reports of test results for *Chlamydia trachomatis* from public health clinics
  - A. Public health surveillance
  - B. Field investigation
  - C. Analytic studies
  - D. Evaluation
  - E. Linkages
  - F. Policy development

- Conducting an analysis of patient flow at the public health clinic to determine waiting times for clinic patients
  - A. Public health surveillance
  - B. Field investigation
  - C. Analytic studies
  - D. Evaluation
  - E. Linkages
  - F. Policy development
The Epidemiological Triad

- Agent
- Environment
- Host

Agent

Environment
The Epidemiological Approach

- **Counts**
  - Health events in terms of person, place, and time

- **Divides**
  - The number of health events by the appropriate denominator to calculate rates

- **Compares**
  - Rates over time or for different groups of people
Descriptive Epidemiology

- The 5 W’s
  - What = what health issue or concern
  - Who = person
  - Where = place
  - When = time
  - Why/How = causes, risk factors, modes of transmission

Lung Cancer Rates in the U.S., 1930-99
Analytic Epidemiology

- **Experimental**
  - Clinical trials
  - Community trials

- **Observational**
  - Cohort studies
  - Case-control studies
  - Cross-sectional studies
Subjects were children enrolled in a health maintenance organization. At 2 months, each child was randomly given one of two types of a new vaccine against rotavirus infection. Parents were called by a nurse two weeks later and asked whether the children had experienced any of a list of side-effects.

- A. Experimental
- B. Observational cohort
- C. Observational case-control
- D. Observational cross-sectional
- E. Not an analytical or epidemiologic study
<table>
<thead>
<tr>
<th>Descriptive Epidemiology</th>
<th>Analytic Epidemiology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search for clues</td>
<td>Clues available</td>
</tr>
<tr>
<td>Formulate hypotheses</td>
<td>Test hypotheses</td>
</tr>
<tr>
<td>No comparison group</td>
<td>Comparison group</td>
</tr>
<tr>
<td>Answers: How much, who, what, when, where</td>
<td>Answers: How, why</td>
</tr>
</tbody>
</table>
Prevalence

- The number of affected persons present in the population divided by the number of people in the population

Prevalence = \frac{\text{# of cases}}{\text{# of people in the population}}
Prevalence

- Useful for assessing the burden of disease within a population
- Valuable for planning
- **Not** useful for determining what caused disease
In 1999, a US state reported an estimated 253,040 residents over 20 years of age with diabetes. The US Census Bureau estimated that the 1999 population over 20 in that state was 5,008,863.

\[
\text{Prevalence} = \frac{253,040}{5,008,863} = 0.051
\]

- In 1999, the prevalence of diabetes was 5.1%
  - Can also be expressed as 51 cases per 1,000 residents over 20 years of age
Incidence

- The number of new cases of a disease that occur during a specified period of time divided by the number of persons at risk of developing the disease during that period of time

\[
\text{Incidence} = \frac{\# \text{ of new cases of disease over a specific period of time}}{\# \text{ of persons at risk of disease over that specific period of time}}
\]
Incidence

- High incidence represents diseases with high occurrence; low incidence represents diseases with low occurrence

- Can be used to help determine the causes of disease

- Can be used to determine the likelihood of developing disease
A study in 2002 examined depression among persons with dementia. The study recruited 201 adults with dementia admitted to a long-term care facility. Of the 201, 91 had a prior diagnosis of depression. Over the first year, 7 adults developed depression.

Incidence = \( \frac{7}{110} = 0.064 \)

- The one year incidence of depression among adults with dementia is 6.4%
  - Can also be expressed as 64 cases per 1,000 persons with dementia
Prevalence can be a useful measure for assisting with determining risk factors associated with a disease?

- True
- False
Cohort Studies

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

- Cohort studies
  - A cohort study analyzes an exposure / disease relationship within the entire cohort

- Cohort study types
  - Prospective
    - The Framingham Study
  - Retrospective
    - Usually used in outbreak investigations
Cohort Studies

Study Population

Exposed

Non-exposed

Disease
No Disease
Disease
No Disease

Exposure is self selected
Follow through time
Cohort Studies

- 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 Studies: Prospective v. Retrospective

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prospective</strong></td>
<td>Assessed at beginning of study</td>
</tr>
<tr>
<td></td>
<td>Followed into the future for outcome</td>
</tr>
<tr>
<td><strong>Retrospective</strong></td>
<td>Assessed at some point in the past</td>
</tr>
<tr>
<td></td>
<td>Outcome has already occurred</td>
</tr>
</tbody>
</table>
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 an alternative
Case-Control Studies

Study Population

Had Exposure
No Exposure

Cases
Had Exposure
No Exposure

Controls
Case-Control Study

Steps in a Case-Control Study:

1. **Identify the source population**
   1. Represents the population that the cases came from; is similar to the cohort in a cohort study

2. **Establish a case definition and select cases**
   1. A standard set of criteria for deciding disease status clinical criteria, time, place, and person

3. **Select controls**
   1. Represent source population
   2. Collect same exposure information as for cases
# Cohort v. Case-Control

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

<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,</td>
</tr>
<tr>
<td></td>
<td>4. Can examine rare exposures</td>
<td>yielding invalid 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>
Basic Outbreak Investigation Steps

1. Prepare for field work
2. Verify the diagnosis and confirm the outbreak
3. Define a case and conduct case finding
4. Tabulate and orient data: time, place, person
5. Take immediate control measures
6. Formulate and test hypotheses
7. Plan and execute additional studies
8. Implement and evaluate control measures
9. Communicate findings
<table>
<thead>
<tr>
<th>Case initials</th>
<th>Sex</th>
<th>County</th>
<th>Symptom onset date</th>
<th>Fever Y/N</th>
<th>Vomiting Y/N</th>
<th>Diarrhea Y/N</th>
<th>Abdominal Cramps Y/N</th>
<th>Duration</th>
<th>Physician Seen Y/N</th>
<th>Antibiotic Y/N</th>
<th>Antidiabetic Y/N</th>
<th>Medication Y/N</th>
<th>Treatment</th>
<th>Lab Testing Lab Name</th>
<th>Collect Date</th>
<th>Type of test</th>
<th>Result</th>
<th>chicken</th>
<th>lettuce</th>
<th>mac. salad</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF</td>
<td>M</td>
<td>Wash.</td>
<td>9/10</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>3d</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>stool</td>
<td>9/12</td>
<td>cult.</td>
<td>+ salm.</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>GF</td>
<td>M</td>
<td>Askance</td>
<td>9/11</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>4d</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>stool</td>
<td>9/14</td>
<td>cult.</td>
<td>+ salm.</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>TE</td>
<td>F</td>
<td>Wash.</td>
<td>9/10</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>2d</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>---</td>
<td>--</td>
<td>---</td>
<td>---</td>
<td>y</td>
<td>y</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>LL</td>
<td>M</td>
<td>Askance</td>
<td>9/10</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>2d</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>Y</td>
<td>N</td>
<td>y</td>
<td></td>
</tr>
<tr>
<td>MJ</td>
<td>F</td>
<td>Askance</td>
<td>9/10</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>3d</td>
<td>y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>stool</td>
<td>9/13</td>
<td>cult.</td>
<td>+ salm.</td>
<td>y</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>M</td>
<td>Askance</td>
<td>9/11</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>4d</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>stool</td>
<td>9/13</td>
<td>cult.</td>
<td>+ salm.</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>PZ</td>
<td>F</td>
<td>Askance</td>
<td>9/10</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>3d</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>MZ</td>
<td>F</td>
<td>Troy</td>
<td>9/10</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>5d</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>blood</td>
<td>9/11</td>
<td>cult.</td>
<td>+ salm.</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>SK</td>
<td>M</td>
<td>Wash.</td>
<td>9/9</td>
<td>y</td>
<td>y</td>
<td>y</td>
<td>Y</td>
<td>3d</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>stool</td>
<td>9/1</td>
<td>cult.</td>
<td>+ salm.</td>
<td>y</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>
Epidemic Curve

Epi Curve for E.Coli Outbreak, n=108

- Number of cases vs Date of onset
- Data points show the number of cases from 10/11 to 11/10.
Attack Rates

AR

# of cases of a disease
# of people at risk (for a limited period of time)

Food-specific AR

# people who ate a food and became ill
# of people who ate that food
## Food Specific Attack Rates

<table>
<thead>
<tr>
<th>Item</th>
<th>Consumed Item</th>
<th>Did Not Consume Item</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ill</td>
<td>Total</td>
</tr>
<tr>
<td>Chicken</td>
<td>12</td>
<td>46</td>
</tr>
<tr>
<td>Cake</td>
<td>26</td>
<td>43</td>
</tr>
<tr>
<td>Water</td>
<td>10</td>
<td>24</td>
</tr>
<tr>
<td>Green Salad</td>
<td>42</td>
<td>54</td>
</tr>
<tr>
<td>Asparagus</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>
Measures of Association

- Assess the strength of an association between an exposure and the outcome of interest

- Two widely used measures:
  - Risk ratio (a.k.a. relative risk, RR)
    - Used with cohort studies
  - Odds ratio (a.k.a. OR)
    - Used with case-control studies
2x2 Tables

- Used to summarize counts of disease and exposure in order to do calculations of association

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>$a$</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>$b$</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>$a+b$</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>$c$</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>$d$</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>$c+d$</td>
</tr>
<tr>
<td>Total</td>
<td>Yes</td>
<td>$a+c$</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>$b+d$</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>$a+b+c+d$</td>
</tr>
</tbody>
</table>
2x2 Tables

\[ a = \text{number who are exposed and have the outcome} \]
\[ b = \text{number who are exposed and do not have the outcome} \]
\[ c = \text{number who are not exposed and have the outcome} \]
\[ d = \text{number who are not exposed and do not have the outcome} \]

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>(a)</td>
<td>(b)</td>
<td>(a + b)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>(c)</td>
<td>(d)</td>
<td>(c + d)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>(a + c)</td>
<td>(b + d)</td>
<td>(a + b + c + d)</td>
</tr>
</tbody>
</table>
2x2 Tables

\[ a + b = \text{total number who are exposed} \]
\[ c + d = \text{total number who are not exposed} \]
\[ a + c = \text{total number who have the outcome} \]
\[ b + d = \text{total number who do not have the outcome} \]
\[ a + b + c + d = \text{total study population} \]

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>a</td>
<td>b</td>
<td>a+b</td>
</tr>
<tr>
<td>No</td>
<td>c</td>
<td>d</td>
<td>c+d</td>
</tr>
<tr>
<td>Total</td>
<td>a+c</td>
<td>b+d</td>
<td>a+b+c+d</td>
</tr>
</tbody>
</table>
## Risk Ratio

<table>
<thead>
<tr>
<th></th>
<th>III</th>
<th>Not III</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>A</td>
<td>B</td>
<td>A+B</td>
</tr>
<tr>
<td>Unexposed</td>
<td>C</td>
<td>D</td>
<td>C+D</td>
</tr>
<tr>
<td>Risk Ratio</td>
<td>[A/(A+B)]</td>
<td>[C/(C+D)]</td>
<td></td>
</tr>
</tbody>
</table>
Interpreting a Risk Ratio

- RR=1.0 = no association between exposure and disease
- RR>1.0 = positive association
- RR<1.0 = negative association / protective effect
Risk Ratio Example

\[
RR = \left( \frac{43}{54} \right) / \left( \frac{3}{21} \right) = 5.6
\]

<table>
<thead>
<tr>
<th></th>
<th>III</th>
<th>Well</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ate alfalfa sprouts</td>
<td>43</td>
<td>11</td>
<td>54</td>
</tr>
<tr>
<td>Did not eat alfalfa sprouts</td>
<td>3</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>46</td>
<td>29</td>
<td>75</td>
</tr>
</tbody>
</table>
# Odds Ratio

<table>
<thead>
<tr>
<th></th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Unexposed</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>Odds Ratio</td>
<td>$(A/C)/(B/D) = (A<em>D)/(B</em>C)$</td>
<td></td>
</tr>
</tbody>
</table>
The odds ratio is interpreted in the same way as a risk ratio:

- OR=1.0 = no association between exposure and disease
- OR>1.0 = positive association
- OR<1.0 = negative association
Odds Ratio Example

<table>
<thead>
<tr>
<th></th>
<th>Case</th>
<th>Control</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ate at restaurant X</td>
<td>60</td>
<td>25</td>
<td>85</td>
</tr>
<tr>
<td>Did not eat at</td>
<td>18</td>
<td>55</td>
<td>73</td>
</tr>
<tr>
<td>restaurant X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>78</td>
<td>80</td>
<td>158</td>
</tr>
</tbody>
</table>

\[ \text{OR} = \frac{60}{18} / \frac{25}{55} = 7.3 \]
Summary

- Outbreaks occur frequently
- Outbreaks are almost always unexpected events
  - Systematic investigation is crucial
  - Preparation is key
Questions??

Contact Information:

Cleveland Department of Public Health
Office of Communicable Disease Surveillance & Epidemiology

75 Erieview Plaza, Cleveland, OH 44114
216-664-EPIS (3747)