Epidemiology Workshop
Oak Ridge Reservation
Health Effect Subcommittee
June 12, 2001

Sherri Berger, MSPH
Lucy Peipins, PhD
Division of Health Studies
Workshop Outline

I. What is Epidemiology and How is it Used?

II. Measures of Disease Frequency and Association

III. Measuring Exposure and Adverse Health Effects

IV. Judging a Cause-Effect Relationship
Workshop Objectives

• Basic epidemiology concepts.
• The importance of good measurement of exposure and disease.
• When is an epidemiologic study appropriate?
What is Epidemiology?

The study of the distribution and determinants of disease in populations.

(J. Last, Dictionary of Epidemiology, 1983)
Epidemiology: The Science of Public Health

Three key aspects:
• groups of people,
• measurement,
• comparison,
What Can Epidemiology Do?

• Determine the impact of disease in groups of people.
• Detect changes in disease occurrence in groups of people.
• Measure relationships between exposure and disease.
• Evaluate the efficacy of health interventions and treatments.
What Can’t Epidemiology Do?

• Cannot tell an individual the cause of his or her disease.
• One study cannot prove a particular exposure caused an illness.
• Should not be conducted without good measurement of exposure and disease.
Is there a relationship between a particular exposure and a disease?

How Science Works

Every epidemiologic study and every analysis is built on a body of knowledge and is done to expand that body of knowledge

Epi Study

Animal studies

Case Reports

Epidemiologic study

Epidemiologic study

Toxicologic Models
Common Steps in Establishing a Relationship Between Exposure and Disease

- Physician reports series of cases
- Descriptive analyses describe what is the problem, who is affected and where disease is occurring
- Analytic studies test the exposure-disease hypothesis in a study group
- Disease experimentally reproduced by exposure in animal studies
- Observation that removing exposure lowers disease
Measures of Disease Frequency

The most basic measure of disease frequency is a simple count of affected individuals.

However, counting is not enough!
Why is a simple count not enough?

- 3 cases of cancer per year from a city of 1,000 people is very different than 3 cases per year from a city of 100,000 people
So, in epidemiology we must know:

- the size of the population from which the affected individuals come, and
- the time period the information was collected.
Basic Measure of Disease Frequency

- **Rate**
  - an example: Incidence
Rate

- basic measure in epidemiology
- the frequency with which an event occurs in a group of people
- used to compare the occurrence of disease in different groups
Example: Rates

Rate = \textbf{Number of events in a specified time period} \\
Average population during the time period

***the measure of time is a critical part of a rate!

Such as, the number of newly diagnosed cases of breast cancer per 100,000 women during 1999.
Common Rates

• mortality (death) rate is the number of deaths in a defined group of people during a specified time period.

• birth rate is the number of live births in a defined group of people over a specified time period.
Incidence

• a type of rate

• the number of new cases that develop in a group of individuals during a specific time period
Incidence (cont.)

\[ I = \frac{\text{number of new events during a period of time}}{\text{number of persons at risk during this time period}} \]
Example: Incidence

If there were 150,000 new cases of lung cancer in the United States during 1997, the incidence rate would be:

\[
\frac{150,000}{260,000,000} = 0.000058
\]

0.000058 x 100,000 people

= 58 cases per 100,000 people, per year
Measures of Association

- How much greater the frequency of disease is in one group compared with another.

- Often presented in the form of a two-by-two table.
# Two-By-Two Table

<table>
<thead>
<tr>
<th></th>
<th>Disease</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>No</td>
<td>a+c</td>
<td>b+d</td>
</tr>
</tbody>
</table>

**Legend:**
- **a**: Number of exposed cases with disease
- **b**: Number of exposed cases without disease
- **c**: Number of unexposed cases with disease
- **d**: Number of unexposed cases without disease
- **a+b**: Total exposed cases
- **c+d**: Total unexposed cases
- **a+b+c+d**: Total cases
## Hypothetical Two-By-Two Table

<table>
<thead>
<tr>
<th>Smoking</th>
<th>Lung cancer</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>70</td>
<td>370</td>
</tr>
<tr>
<td>No</td>
<td>15</td>
<td>715</td>
</tr>
<tr>
<td>Total</td>
<td>1,000</td>
<td>1,085</td>
</tr>
</tbody>
</table>
Relative Risk (RR)

- Measures how likely the exposed group will develop a disease compared to the unexposed group.

\[ RR = \frac{\text{incidence in the exposed}}{\text{incidence in the unexposed}} = \frac{a}{a+b} \]
\[ \frac{c}{c+d} \]
## Example: Hypothetical Study

<table>
<thead>
<tr>
<th>Smoking</th>
<th>Lung cancer</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>70</td>
<td>300</td>
<td>370</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>15</td>
<td>700</td>
<td>715</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>85</td>
<td>1,000</td>
<td>1,085</td>
</tr>
</tbody>
</table>

**TSDR**
Relative Risk = \( \frac{70}{70+300} = 9.0 \) 
\( \frac{15}{15+700} \)

Which means… participants who smoked were 9 times more likely to develop lung cancer than those who did not smoke.
**Interpreting Measures of Association**

RR of 1.0 indicates that the occurrence of disease in the exposed and unexposed groups are identical:

- No association observed between exposed and unexposed groups.
Interpreting Measures of Association (Continued)

• RR greater than 1.0 indicates a positive association, or an increased risk among the exposed.

• RR less than 1.0 means that there is a decreased risk among the exposed group.
**Standardized Mortality Ratios (SMRs)**

- the ratio of the observed number of deaths to the expected number of deaths

- a standard group of people is used to determine the expected number of deaths
  - the standard is often the US population, a state, or a county.
  - serves as the comparison group
**SMRs (Continued)**

Calculation:

\[
\text{SMR} = \frac{\text{observed deaths}}{\text{expected deaths}}
\]
Example: Study of Smoking and Lung Cancer

Our hypothetical study found 58 lung cancer deaths between 1948 and 1963:

- based on US population rates, we know that 42.9 cancer deaths were expected in a similar population.

\[
\text{SMR} = \left( \frac{58}{42.9} \right) = 1.35
\]
Our study group had a risk of cancer mortality approximately 35% greater than those in the general population.
Measuring Exposure and Measuring Outcomes in Environmental Epidemiology

- Contaminants in air, water, food, or soil come in contact with people through swallowing, breathing, or skin contact

- Once the contaminant gets in the body, it can act at that point of entry or can be distributed throughout the body where it can act upon a target organ or organs
# Measuring Exposure

We can measure exposures or their surrogates along the entire path from emissions to body burden

<table>
<thead>
<tr>
<th>Types of Data</th>
<th>Approximation to actual exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Residence in a defined geographical area (e.g. county) of a site</td>
<td>Poorest</td>
</tr>
<tr>
<td>2) Residence in a geographical area in proximity to a site where exposure is assumed</td>
<td></td>
</tr>
<tr>
<td>3) Distance <strong>or</strong> duration of residence</td>
<td></td>
</tr>
<tr>
<td>4) Distance <strong>and</strong> duration of residence</td>
<td></td>
</tr>
<tr>
<td>5) Quantified surrogate of exposure (e.g. estimate of drinking-water use)</td>
<td></td>
</tr>
<tr>
<td>6) Quantified area measurements in vicinity of the residence</td>
<td></td>
</tr>
<tr>
<td>7) Quantified personal measurement</td>
<td>Best</td>
</tr>
</tbody>
</table>

Baker D, et al., Environmental Epidemiology: A Textbook on Study Methods and Public Health Applications, 1999
What is an Outcome or Adverse Health Effect?

• Any change in health status or body function.
  
  – Wide spectrum of effects such as wheezing, change in immune function, changes in blood chemistry, DNA adducts, adverse birth outcomes, developmental disabilities, clinical disease and death
These changes can be...
What Outcomes Do We Look For?
Biologically plausible outcomes

- Contaminant → Target Organ System
- Route of exposure
- Toxicity or level → Severity of Outcome
- Alternative explanations → Confounding
Health Effects are Not Uniquely Caused by Environmental Exposures

- 25% of all pregnancies result in spontaneous abortions
- 2% of infants are born with developmental deformities
- Among men, 1 in 2 will develop cancer in their lifetime (1 in 3 for women)
- 8-10% of children have asthma

ATSDR
Cancer

- Smoking is the leading cause of cancer in the U.S.
  - 25% of Americans currently smoke

- Can be a very long time between exposure and the development of disease – many exposures over that time

- Cancer is not 1 disease – it is thought to be over 100 different disease in 40 anatomic sites

ATSDR
Measuring Adverse Health Effects

• Goal: to count *all* the cases in a particular exposed group or population and compare it with cases in an unexposed group or population

• Where do we get this information?
Data Sources

- death certificates
- birth certificates
- medical exams
- hospital discharge data
- questionnaires
- disease registries
- laboratory tests or biomarkers
**Data Sources: Four Considerations**

- Severity of illness
- Accuracy of measuring disease-misclassification of the outcome
- Availability of information on confounders or other risk factors
- Cost and complexity and invasiveness
Death Certificates

- Readily available
- Not all conditions of interest result in death
- Considerable inaccuracy in diagnosis (error in major disease category is 25%, error within categories is 25%)
- No data on other risk factors
Registries

- Registries of reportable diseases such as TB, cancer, birth defects.
- Cancer incidence data is available for all states.
- No information on other risk factors.
- Completeness and timeliness may be an issue.
Medical Exams and Biologic Tests

- Generally regarded as the “gold standard”

- However, physician diagnosis may vary for outcomes with no standard case definition
Questionnaires

• May miss most severe outcomes

• Some outcomes may only be measured through a questionnaire

• Self-reports of illness or symptoms
Questionnaires, (cont.)

• Wording or type of administration may influence responses

• Can collect data on other risk factors

• Subject to recall and response bias
Measuring Outcomes: Other Considerations

- Timing and latency: especially important in chronic disease
- Individual variability by: age, gender, pre-existing illnesses, genetic predisposition
- Confounding and bias: What are alternative explanations for the exposure-disease association?
Summary: Measuring Health Outcomes and Exposures

• The closer we can get to a measure of dose in the body, the more precise our exposure estimate will be.

• Characteristics such as age or pre-existing illness modify or determine the health effects we might see.

• Sources of data on health effects and exposures are varied.
Basic Criteria for Considering a Health Study

We need:

1. An exposure that can be measured.
2. A completed pathway.
3. An exposed population.
4. A measurable effect that is plausibly related to the exposure.
Basic Criteria for Considering a Health Study

• Will the epidemiologic study we are planning advance our knowledge about the relationship between exposure and disease?
How to Judge a Cause-Effect Relationship

You’re given a paper that reports the results of a study, how do you interpret this study?
Strength of Association

• example: Relative Risk (RR)

• the larger the risk, the more likely the relationship exists
  – less likely it is due to other factors, such as confounding
Consistency of Findings

Has this association been seen with other studies, with other study designs, and in different groups of people?

- If so, this strengthens the findings
Dose-Response Relationship

• As exposure increases, does the risk of disease increase?
**Time Sequence**

- the exposure MUST occur before the disease develops

- this is an essential criteria
Biologic Plausibility

• does this make biologic sense?

• is there a known mechanism by which an exposure can lead to disease?
Biologic Plausibility (cont.)

Example: daily consumption of fatty foods increases risk of heart disease

- there is a biologic mechanism
- fatty foods raises LDL cholesterol, which increases the risk of heart disease
In Conclusion...

• these are not hard and fast rules (except time sequence)
  – there can be exceptions!

• These tools help us judge the scientific literature – have the authors convinced us?
Epidemiology is an excellent tool if used correctly!

Joseph Mangano ⊗ Cancer Mortality near Oak Ridge, Tennessee