Epidemiologic Principles
Causality
Confounding
Bias

GOALS

◆ Apply elements of causality to assessment of data
◆ Identify potential confounders in research designs and studies
◆ Recognize sources of bias in published research reports

Surgical Site Infection Rate

◆ All surgeons: 2.3%
◆ Dr. H: 4.5%
Why?
◆ Sees highest risk patients (confounding)
◆ Caused by factor associated with both Dr. H and infections (confounding)
◆ Collects better data (bias)
◆ Sample size is too small (statistical artifact)
◆ Chance

Wound Infection Rates

Did Dr. H “cause” more infections?
◆ Temporal sequence: surgery before infection
◆ Strength of association: High relative risk
◆ Consistency: present over several risk categories
◆ Statistical significance: Events unlikely to be chance
Associations Between Variables

- None
- Artifactual
  - Chance
  - Bias
- Indirect (confounding, extraneous)
- Causal

Evaluating Causality

- Koch’s Postulate: An organism (cause) is always found with the disease (effect): SPECIFICITY
- Exception: Many different “causes” can result in the same effect (e.g., pneumonia is caused by different organisms)

- Exception: The same “cause” can have many different effects (e.g., Strep. may cause sore throat, impetigo, scarlet fever)
Evaluating Causality

◆ Koch’s Postulate:
The organism (cause) when isolated from a diseased person will induce the same disease (effect) in another person

◆ Exception:
Some “causes” may not produce any effect (eg. Colonization with an organism with no disease)

ELEMENTS OF CAUSALITY

Temporal Relationship

‘Cause’ must precede ‘effect’
**Strength of Association**

- Risk of the outcome ‘effect’ among those exposed to the ‘cause’ must be greater than the risk among unexposed.

**Strength of Association Measured by Relative Risk**

<table>
<thead>
<tr>
<th>Disease</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>A+C</td>
<td>B+D</td>
<td>A+B+C+D</td>
</tr>
</tbody>
</table>

**Calculating Relative Risk**

\[
\frac{A}{A+B} \quad \text{vs.} \quad \frac{C}{C+D}
\]

Incidence in exposed vs. Incidence in unexposed

\[
\frac{A}{A+B} \text{ divided by } \frac{C}{C+D}
\]
Specificity of the Association

One ‘cause’ is specifically and only associated with one ‘effect’
(e.g. HIV and AIDS)

Plausability

Association between ‘cause’ and ‘effect’ makes biological or psychological sense

Consistency of Association

The same ‘cause’ is associated with the same ‘effect’ in a variety of circumstances
Example:
Smoking and Lung Cancer
◆ Temporal: Did smoking precede lung cancer?
◆ Strength: Large relative risk?
◆ Specificity: Lung cancer only occurs in smokers?
◆ Plausibility: Biologic rationale?
◆ Consistency: Lung cancer in men/women smokers? Several brands? Various study designs?

Why Was It Easy to Determine Causal Association Between Smoking and Lung Cancer?
◆ Exposure is easily, accurately assessed
◆ ‘Cause’ (smoking) is common and present in otherwise similar people
◆ Large relative risk and clear dose response
◆ Lung cancer (‘effect’) comparatively uncommon in non-smokers

Nurse Accused of Murder
Old Age and Confusion: Relevant Questions?

- Temporal Relationship?
- Strength of Association?
- Specificity?
- Plausability?
- Consistency?

Catheterization and UTI: Relevant Questions?

- Temporal Relationship
- Strength of Association
- Specificity
- Plausability
- Consistency

Three Factors That Interfere With Causal Inference

- Chance
- Confounding
- Bias
Did It Occur By Chance?

Statistical significance?
Adequate statistical power?
Replicated studies?
Statistical tests to control for multiple comparisons?

Confounding (Extraneous) Variable

Variable that has an irrelevant or unwanted effect on the relationship between the variables being studied, causing a distortion of the ‘true’ relationship

Confounding

Exposure ➔ Outcome

Confounder
Example

- **Exposure** ('cause') = type of needle (plastic or steel)
- **Outcome** ('effect') = phlebitis
- **Confounder** = time in place

Example

- **Exposure** ('cause') = hours of study
- **Outcome** ('effect') = class grades
- **Potential confounders** =
  - Health
  - Intelligence

Crude mortality rates in US are higher than in Nicaragua, despite the fact that death rates in Nicaragua in every age category are higher.

Why?
Relationship Between Cholesterol Level and CHD

<table>
<thead>
<tr>
<th>Serum Cholesterol (mgm%)</th>
<th>Men Ages 30-49</th>
<th>Men Ages 50-62</th>
<th>Women Ages 30-49</th>
<th>Women Ages 50-62</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;190</td>
<td>1.0</td>
<td>2.8</td>
<td>0.3</td>
<td>4.1</td>
</tr>
<tr>
<td>190-219</td>
<td>1.2</td>
<td>4.9</td>
<td>0.2</td>
<td>2.3</td>
</tr>
<tr>
<td>220-249</td>
<td>2.5</td>
<td>5.3</td>
<td>0.6</td>
<td>2.5</td>
</tr>
<tr>
<td>250+</td>
<td>4.1</td>
<td>7.0</td>
<td>1.3</td>
<td>3.2</td>
</tr>
</tbody>
</table>

To Look for Confounding….

◆ Is the factor related to exposure? Disease? (must be related to both)
◆ Stratify by the variable (e.g. age groups). Is the relative risk different?

Examples of Confounders?

◆ Effect of breathing exercises on post-operative respiratory complications
◆ Effect of training course for pediatric nurses on nurturing behaviors of nurses
◆ Effect of type of nursing education on involvement in professional organization and politics
## Is Drinking Alcohol Associated with Increased Risk of Lung Cancer?

<table>
<thead>
<tr>
<th>Lung cancer patients who drink</th>
<th>Lung cancer patients who did not drink</th>
<th>Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>200/250=80%</td>
<td>50/250=20%</td>
<td>80/20=4</td>
</tr>
</tbody>
</table>

## Same Subjects, Stratified by Smoking

<table>
<thead>
<tr>
<th>Among smokers, # with lung cancer</th>
<th>Among non-smokers, # with lung cancers</th>
<th>Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>220/1000=22%</td>
<td>10/1000=1%</td>
<td>22</td>
</tr>
</tbody>
</table>

## Same Subjects, Stratified by Smoking

<table>
<thead>
<tr>
<th>Among smokers, # who drank</th>
<th>Among non-smokers, # who drank</th>
<th>Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>900/1000=90%</td>
<td>100/1000=10%</td>
<td>9</td>
</tr>
</tbody>
</table>
**Same Subjects, Stratified by Smoking**

<table>
<thead>
<tr>
<th>Among smokers, # of drinkers with lung cancer</th>
<th>Among non-smokers, # of drinkers with lung cancer</th>
<th>Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>197/1000= 19.7%</td>
<td>3/1000= 0.3%</td>
<td>65.7</td>
</tr>
</tbody>
</table>

**Conclusion**

- Smoking was associated with lung cancer AND
- Smoking was associated with drinking
- Smoking was associated with both the dependent (lung cancer) and independent variable (drinking) and is therefore a confounding variable

**THEREFORE...** it was the smoking, not the drinking associated with lung cancer

**Age-Adjusted Esophageal Cancer Deaths by Race and Sex**

![Graph showing age-adjusted esophageal cancer deaths by race and sex]
Age-Specific Mortality by Birth Year, Esophageal Cancer

Avoiding Confounding

- Use homogeneous subjects
- Match subjects or stratify by potential confounder
- Randomize
- Statistical procedures such as analysis of covariance

BIAS

A prejudice or opinion formed before the fact. In research, usually unintentional and unknown to researcher
Selection Bias

Study population differs in a way that is likely to affect study results

Detection Bias

Knowledge about a particular exposure or characteristic of the subjects increases the search for certain effects

Investigator Bias

A preconceived notion about the outcome of a study which can influence the investigator’s evaluation
Non-Response Bias

Responders vary from non-responders with regard to relevant variables

Recall Bias

Certain subjects recall past differentially better than other subjects

Give a rival hypothesis….

◆ Nursing students and test anxiety
◆ Remedial math course
◆ Adolescent girls and pelvic exam
Minimize Bias

◆ SELECTION: strict inclusion criteria
◆ DETECTION: identify ‘effect’ equally in all subjects
◆ INVESTIGATOR: ‘blinding’/‘masking’, inter-rater reliability, explicit and objective measurement

Minimize Bias

◆ NON-RESPONSE: randomize study groups or carefully select groups for comparability, make study participation easy, followup with non-responders to identify systematic differences
◆ RECALL: structured interview or survey, reinterview a sample

Want More?