Epidemiological Methods

Duanping Liao, MD, Ph.D

Professor and Vice Chair for Research
Chief, Division of Epidemiology
Department of Public Health Sciences
Penn State University College of Medicine

Email: Dliao@phs.psu.edu
Phone: 531-4149
What is Epidemiology?
Epidemiology is the study of the distribution and determinants of health-related states or events in specified populations.

Epidemiologists develop and evaluate hypotheses about the effects of genetic, environmental, behavioral, and health care factors on human health/disease.

Epidemiologists also apply epidemiological study results to the prevention and control of health problems.

The field is interdisciplinary and has a methodology distinctive from, but also dependent on, biostatistics.
Clinical Epidemiology and Clinical Research

Primary Prevention

Risk Factors (Etiology)

Medical intervention

Disease (Outcome)

Progression / Regression

Diagnostic Tests
Se, Sp, FP, and FN
## Classical Epidemiological Studies

<table>
<thead>
<tr>
<th>Categories by Objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Descriptive Studies</strong></td>
</tr>
<tr>
<td>1. Correlational (Ecological) Studies</td>
</tr>
<tr>
<td>2. Case Reports</td>
</tr>
<tr>
<td>Case Series</td>
</tr>
<tr>
<td>3. <em>Cross-sectional</em></td>
</tr>
<tr>
<td>Studies</td>
</tr>
</tbody>
</table>
DISTRIBUTION OF DISEASE BY PLACE

Breast Cancer Incidence by Dietary Fat Intake by Country

Conclusion: Fat intake may be associated with breast cancer risk.
Ecological fallacy!!
It assumes that individual members of a group have the average characteristics of the group at large. However, statistics that accurately describe group characteristics do not necessarily apply to individuals within that group.
CROSS-SECTIONAL STUDY
(Prevalence study)
An epidemiological study in which exposure and/or outcome are assessed simultaneously among individuals in a well-defined population.

CDC: 1 in 13 pregnant women say they drink alcohol (AP, July 19, 2012):

ATLANTA (AP) — A government survey shows 1 in 13 pregnant women drink alcohol and some even go on binges.

The Centers for Disease Control and Prevention says drinking — especially binge drinking — can damage fetal brain development.
Prevalence: The total number of cases of a disease/condition in a given population at a specific time.

**CDC: 1 in 13 pregnant women say they drink alcohol (AP, July 19, 2012):**

The CDC numbers are from national telephone surveys that included 14,000 pregnant women. They were asked whether they drank alcohol the previous month and how much. Of those who said they drank, nearly 1 in 5 said they went on at least one binge — downing four or more drinks. Pregnant women ages 35 to 44 were the biggest drinkers.

The good news: Binge drinking among pregnant women is down slightly since a similar study was done in the early 2000s.

The survey results from 2006 through 2010 were released Thursday.

**Prevalence_drinking:** N of Yes / total number of women surveyed = n / 14000 = 1 / 13
n of drinker = 1077, or the prevalence of 7.7%.

**Prevalence_binge** = n of women drink 4 or more drinks / total number of drinkers
= n / 1077 = 1 / 5, thus, n= 215, or prevalence of 20%
CROSS-SECTIONAL STUDY
(Prevalence study)

Use cross-sectional study to generate hypothesis of the etiological relationship between a risk factor and a disease.

Defined Pop

Collect data on exposure, outcome and covariates

E+ & D+

E+ but D-

E- but D+

E- & D-

D +

D -

E +

P_1 (D|E+)= a / (a+b) vs. P_0 (D|E-)=c / (c+d)

E -

P_1 (E+|D+)= a / (a+c) vs. P_0 (E+|D-)= b / (b+d)

Data Structure 2 x 2 Table
Example: Smoking and prevalent CHD in the baseline examination of a sub-sample of the ARIC Study.

Data structure

\[
P_1 (D|E+) = \frac{89}{89+1453} = 5.8\% \\
P_1 (E+|D+) = \frac{89}{89+28} = 76\% \\
P_0 (D|E-) = \frac{28}{28+986} = 2.8\% \\
P_0 (E+|D-) = \frac{1453}{1453+986} = 24\%
\]
Table 2. The Exposure x Disease (2 x 2)

<table>
<thead>
<tr>
<th></th>
<th>CHD +</th>
<th>Non-CHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoker</td>
<td>89 - a</td>
<td>1453 - b</td>
</tr>
<tr>
<td></td>
<td>28 - c</td>
<td>986 - d</td>
</tr>
</tbody>
</table>

**CHD prevalence:**
- Among Smokers: $\frac{89}{89+1453}=5.7\%$
- Among non-Smokers: $\frac{28}{28+986}=2.8\%$

**Smoking prevalence:**
- Among CHD +: $\frac{89}{89+28}=76\%$
- Among CHD -: $\frac{1453}{1453+986}=24\%$

From $X^2$ test:
- $X^2 = 12.69$, $p = 0.001$
- Prevalent Odds Ratio:
  - $OR = \frac{(a \times d)}{(b \times c)}$
  - $= 2.16$
Cross-Sectional Association between Sleep and Hypertension  (Sleep 2009, 32: 491-497)

Population of 1741: Sleep Variables and Hypertension Diagnosis

<table>
<thead>
<tr>
<th>Hyp +</th>
<th>Hyp -</th>
</tr>
</thead>
<tbody>
<tr>
<td>121</td>
<td>78</td>
</tr>
<tr>
<td>837</td>
<td>705</td>
</tr>
</tbody>
</table>

Prevalence of Hypertension: 958 / 1741=55%, weighted 35%
Prevalence of Insomnia: 199/1741=11%, weighted 8%

\[
P_1 (\text{D|E+}) = \frac{121}{121 + 78} = 61\%
\]

\[
P_0 (\text{D|E-}) = \frac{837}{837 + 705} = 54\%
\]

Prevalence ratio = 61 / 54 = 1.13
Odds Ratio = \( \frac{121 * 705}{78 * 837} = 1.31 \)

After accounting for sampling weight:

Prevalence of Hypertension: 35%
Prevalence of Insomnia: 8%

\[
P_1 (\text{Prevalence of Hyp among insomnia}) = 55\%
\]

\[
P_0 (\text{Prevalence of Hyp among non-insomnia}) = 34\%
\]

X² = 25,  p < 0.01

The prevalence of hypertension is significantly higher among persons with chronic insomnia than among persons without.

Prevalence ratio = 55 / 34 = 1.62
Odds Ratio = 2.45,  95% CI 1.71,  3.51,  p < 0.01

The odds of hypertension is significantly higher among chronic insomnia patients than among non-insomnia patients.