Case-Control Studies

Marc Schenker M.D., M.P.H
Dept. of Public Health Sciences
UC Davis
Case-Control Studies

OBJECTIVES

After this session, you will be familiar with:

• The basic design features of a case-control study
• Rationale for applying case-control designs
• Limitations of case-control studies
• Example applications applying case-control designs
I. Overview

A. Design of Case-Control Studies

The investigator selects cases with the disease, and appropriate controls without the disease, and obtains data regarding past exposure to possible etiologic factors in both groups. The investigator then compares the frequency of exposure of the two groups.
Design of a Case-Control Study

Exposed | Not Exposed
---------|------------
Disease   |            

Exposed | Not Exposed
---------|------------
No Disease|            

“CASES”   | “CONTROLS”
**Figure 1**

<table>
<thead>
<tr>
<th></th>
<th>CASES (With Disease)</th>
<th>CONTROLS (Without Disease)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIRST: Select</td>
<td></td>
<td></td>
</tr>
<tr>
<td>THEN: Were exposed</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>Measure Exposure</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td></td>
<td>a + c</td>
<td>b + d</td>
</tr>
<tr>
<td>Proportions Exposed</td>
<td>a/(a+c)</td>
<td>b/(b+d)</td>
</tr>
</tbody>
</table>
Figure 1 (continued)

Odds Ratio = \[
\frac{\frac{a}{c}}{\frac{b}{d}} = \frac{ad}{bc}
\]

Risk
\[
= \frac{a}{a + b} \quad \text{E+} \quad \text{Case} \\
\leq \quad \frac{c}{c + d} \quad \text{E-} \quad \text{Control}
\]

\[
\text{Case} \\
\begin{array}{cc}
a & b \\
c & d
\end{array}
\]
I. Overview

B. When to use a case-control approach
   1. Rare disease: Case-control approaches are the most efficient for rare diseases, e.g. idiopathic pulmonary fibrosis, most cancers. Cohort approaches would require large populations and prohibitive expense and follow-up time. Case-control designs may also be appropriate for more common diseases, such as COPD.
B. **When should a case-control approach be used**

2. **Case ascertainment system in place:** The conduct of a case-control study may be facilitated by the availability of a case-ascertainment system.
   - a) Population-based cancer registry
   - b) Hospital-based surveillance systems
   - c) Mandated disease reporting systems

3. **When funding and time constraints are not compatible with a cohort study.**
II. Issues in Case-Control Studies

A. Issues in Ascertainment of Cases
   1. Diagnostic criteria for case studies
      a) Specificity
         e.g. lung cancer vs wheezing
      b) Diagnostic bias
      c) Validation
   2. Sources (hospital, general population)
   3. Incident or prevalent cases
II. Issues in Case-Control Studies

B. Issues in Selection of Controls

1. General questions
   a) Conceptual
      (i) Should the controls be comparable to the cases in all respects other than having the disease?
      (ii) Should the controls be representative of all non-diseased people in the population from which the cases are selected?
Figure 2

“Total” Population

Reference Population

Cases

Controls
B. Issues in Selection of Controls

b) Practical Questions
   (i) Is the approach selected for control selection feasible?
   (ii) Can this approach be used given the funds available?
2. Sources of controls

a) Population of defined area
b) Hospital patients
c) Probability sample of total population
d) Neighbors
  (i) walk (door to door)
  (ii) phone (random digit dialing)
  (iii) letter carrier routes
e) Friends or associates of cases
f) Siblings, spouses or other relatives
g) Other
C. Methodologic Issues

1. Handling potential confounding factors
   a) In the process of selecting controls:
      Matching
      The process of selecting controls so that they are similar to the cases in regard to certain characteristics such as age, sex and race.
      (i) Group matching (frequency matching, stratification)
      (ii) Individual matching (matched pairs)
C. Methodologic Issues

Handling potential confounding factors in matching:

(iii) Problems with matching:
- Matching on many variables may make it difficult or impossible to find an appropriate control.
- Cannot explore possible association of disease with any variable on which cases and controls have been matched.
C. Methodologic Issues

Handling potential confounding factors in matching:

b) In the process of selecting controls:
   Restriction

c) In the data analysis:
   (i) Stratification
   (ii) Adjustment
C. Methodologic Issues

2. Evaluating Information on Exposure
   a) Problems of recall in case-control studies
      (i) Limitations in human ability to recall
      (ii) Recall bias (cases may remember their exposure with a higher or lower accuracy than controls do)
2. Evaluating Information on Exposure

b) Avoiding other biases
   (i) Selection bias
   (ii) Information bias
   (iii) Non-response bias
   (iv) Analysis bias

c) Validity testing (reliability, sensitivity and specificity)
3. Using Multiple Controls in Case-Control Studies

a) Multiple controls of a similar type (e.g. 2 controls per case)

b) Different types of controls (e.g. hospital and neighborhood controls)
III. Nested Case-Control Studies

Figure 3

- Study Population
  - Develop Disease
    - CASES
  - Do Not Develop Disease
    - CONTROLS

Obtain interviews, bloods, urines, etc.

TIME 1
YEARS
TIME 2
A. Advantages of Nested Case-Control Studies

1. Possibility of recall bias is eliminated, since data on exposure are obtained before disease develops.

2. Exposure data are more likely to represent the pre-illness state since they are obtained years before clinical illness is diagnosed.

3. Costs are reduced compared to those of a prospective study, since laboratory tests need to be done only on specimens from subjects who are later chosen as cases or as controls.
Specific Respiratory Disease Applications

A. Causes of disease

1. Incident asthma: Childhood or adulthood, use nested designs within exposed cohorts
2. COPD: Use to investigate factors other than cigarette smoking
3. Symptom syndromes: Tight-building syndrome and multiple chemical sensitivity as examples
4. Rare diseases, i.e. idiopathic pulmonary fibrosis
Respiratory Disease Applications

B. Course of disease
   1. Use to assess factors leading to poor/good outcomes
   2. Use for surveillance for complications of therapy

C. For disease surveillance
The Association between Tuberculosis and Cancer

From the first 7500 autopsies at the Johns Hopkins Hospital, the following were identified and selected:

Cancer cases 816
Controls (no cancer) 816

*Pearl, 1929*
### Summary of Data from Study of Tuberculosis and Cancer

<table>
<thead>
<tr>
<th></th>
<th><strong>Cancer</strong></th>
<th><strong>“Controls”</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n</strong></td>
<td>816</td>
<td>816</td>
</tr>
<tr>
<td><strong>%</strong></td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>Autopsies</strong></td>
<td>54</td>
<td>133</td>
</tr>
<tr>
<td><strong>With Tbc.</strong></td>
<td>6.6</td>
<td>16.3</td>
</tr>
</tbody>
</table>

*Pearl, 1929*
B. Childhood Asthma and Passive Smoking

How would you design a study to evaluate whether passive smoking is a risk factor for childhood asthma?

1. What are advantages and disadvantages of different study designs?
2. How would you evaluate exposure?
B. Childhood Asthma and Passive Smoking
Urinary cotinine as a biomarker of exposure*

1. Design: Two case groups of asthmatics were compared with a control group for passive smoking exposure by comparing urinary cotinine levels.

2. Cases: 72 children aged 3-14 years who were seen at a pediatric ER and walk-in clinic with an acute asthma attack comprised one of the case groups. The other case group consisted of 35 children aged 3-14 years attending an asthma clinic with a history of episodic or chronic airflow obstruction requiring some form of bronchodilator therapy.

B. Childhood Asthma and Passive Smoking

3. Controls: 121 children attending the ER for problems other than asthma.
4. Data: Questionnaires and urinary cotinine/creatinine levels were used to assess passive smoking exposure.
5. Results: Smoking by the maternal caregiver was most strongly associated with asthma in the child. 
   Odds ratio was 1.9 (95% C.I.: 1.04 - 3.35; p=.04).
Table 4. Asthma Versus Control: Exposure Variables

<table>
<thead>
<tr>
<th>Factor</th>
<th>Asthma (n - 107)</th>
<th>Control (n = 121)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any smoker at home, %</td>
<td>54</td>
<td>51</td>
</tr>
<tr>
<td>Daily cigarettes by all smokers, mean +/- SD</td>
<td>8.7 +/-12.8</td>
<td>6.1 +/-10.3</td>
</tr>
<tr>
<td>Maternal caregiver smokes, %</td>
<td>44</td>
<td>28^a</td>
</tr>
<tr>
<td>CCR &gt; 30 ng/mg, %</td>
<td>38</td>
<td>25^b</td>
</tr>
<tr>
<td>Mean CCR, ng/mg^c</td>
<td>43.6 +/-87.7</td>
<td>25.8 +/-46.5^d</td>
</tr>
</tbody>
</table>

^aOdds ratio = 2.0(1.1, 3.4), p = 0.03
^bOR = 1.9 (1.04, 3.35), p = 0.04
^cAitchison transformation, see METHODS
^dp = 0.06
B. Childhood asthma and passive smoking

6. The results could not show, however, that recent elevations in tobacco exposure triggered acute asthma attacks requiring visits to the ER.