Obtaining Exposure Information

Once you have identified cases and controls, the next step is to obtain exposure information from both groups.

- Exposures occurred in the past - often many years ago.

Collecting information on past exposures

- Often collect information by interview.
- Interviewee may not know information or may not remember it well.
- Poor memory of past exposures can lead to misclassification.
Sources of Exposure Information

- Self-report:
  - Person-to-person interview; telephone interview; self administered questionnaire or a combination
- Medical records
  - (HMO; County records, MD offices, hospital records)
  - Records completed before the occurrence of outcome event preferable
- Interviews with proxies/surrogates
  - Spouses, sibs, friends

Collecting Exposure Information

- Will cases and controls remember with equal accuracy?
- Will they report accurately?
- Will proxy respondents answer as the intended respondent would have answered?

Recall Bias

- Differing recall by cases than controls is named recall bias
- Cases may be more motivated to try to remember things in their past that might have caused their disease than are (healthy) controls
- Few examples in literature showing major impact on study findings
Observer Bias

- As interviewer may know who is a case and who is not, they may treat the cases and controls differently when collecting exposure information
- Can lead to interviewer/observer bias

Advantages

- Good for studying rare diseases
- Relatively fast to conduct (months/years)
- Relatively inexpensive
- Relatively few subjects
- Existing records may be available
- Minimal risks to subjects usually
- Study multiple causes of disease (exposures)
- Evaluation of diseases with long latency periods

Disadvantages

- Selection of appropriate controls may be difficult
- Relies on recall or existing records
- Validation of information difficult
- Control of other variables may be difficult
- Cannot calculate relative risk directly or determine prevalence
- Temporal sequence between E and D difficult to establish
- Prone to Selection/Observer/Recall Bias
Measure of Association
- Need a measure of the association between disease and exposure
- In case-control studies the measure is called the Odds Ratio
- May be used on its own or as an estimate the Relative Risk

Calculating Risk in Case-Control Studies
- Cohort studies calculate relative risk
- To calculate relative risk must have incidence
- In a case-control study do not have incidence – just selected cases and controls
- Case-Control study cannot calculate relative risk directly - use **ODDS RATIO** to estimate

Odds Ratio
- Odds is the number of ways an event can occur divided by the number of way it cannot occur
- Probability event can occur = P (40%) = 0.4
- Probability event cannot occur = 1-P (60%) = 0.6
- Odds ratios can be calculated for cohort, case-control and cross-sectional studies
Calculating the Odds Ratio

<table>
<thead>
<tr>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td></td>
</tr>
<tr>
<td>Unexposed</td>
<td></td>
</tr>
</tbody>
</table>

\[
\text{Odds of being a case among exposed} = \frac{a}{b} \\
\text{Odds of being a case among unexposed} = \frac{c}{d}
\]

\[
\text{ODDS RATIO} = \frac{ad}{bc}
\]

Odds ratio
- If exposure is not related to disease odds ratio (OR) = 1
- If exposure is positively associated with disease OR is greater than 1
- If exposure is negatively associated with disease (protective) OR is less than 1

Confidence Interval of the OR
- The probable range in which a population parameter lies based on a random sample of the population.
- Most common interval used is 95%
- If the data collection and analysis was repeated many times, the confidence interval should include within it the correct value, 95% of the time (Rothman, 2002)
- More on confidence intervals in a later lecture
### R x 2 tables -- Odds ratio estimates

#### Dose-response

<table>
<thead>
<tr>
<th>Smoking</th>
<th>Bladder Ca</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>&gt; 20 yrs</td>
<td>84/138</td>
</tr>
<tr>
<td></td>
<td>10 - 20 yrs</td>
<td>72/146</td>
</tr>
<tr>
<td></td>
<td>&gt; 0 - 10 yrs</td>
<td>64/160</td>
</tr>
<tr>
<td>Never</td>
<td></td>
<td>110/460</td>
</tr>
</tbody>
</table>

Never = reference (non-exposed) group

### Hypothetical case-control study to assess effects of exposure to asbestos in the etiology of lung cancer

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Disease</th>
<th></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>NO</td>
<td>YES</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>NO</td>
<td>d</td>
</tr>
</tbody>
</table>

Odds ratio = \(a/c = 60 \times 180 / 20 \times 140 = 10800 / 2800 = 3.857\)

Interpretation: persons with the disease are 3.86 times more likely to have the exposure compared to persons without the disease.

### Odds Ratios for Matched Pairs

- Matched Pairs are individually matched to each other
- Each pair is considered as one unit
- If either case or control in pair does not participate in study, both members of the pair must be eliminated from analysis
**Matched Pairs**

**Concordant Pairs:**
Both case & control were exposed or neither case nor control were exposed

**Discordant Pairs:**
Case was exposed but control was not or control was exposed but case was not

---

**Odds Ratios for Matched Pairs**

\[
\begin{array}{c|c|c|c}
\text{Cases} & \text{Exposed} & \text{Unexposed} \\
\hline
\text{Exposed} & a & b \\
\text{Unexposed} & c & d \\
\end{array}
\]

\[
\frac{b}{c}
\]
Generally speaking in traditional case-control studies OR ~ RR when the following assumptions are met:

1. When the *cases* are representative, with regard to history of exposure, of all people with the disease in the population from which the cases were drawn.
2. When the *controls* are representative, with regard to history of exposure, of all people without the disease in the population from which the cases were drawn.
3. When the disease being studied does not occur frequently.

### Interpretation of OR as RR

OR ~ RR when the disease is rare:

\[
RR = \frac{a}{a+b} \frac{c+d}{c+d} \quad a+b \sim b \\
\frac{c}{c+d} \quad c+d \sim d
\]

\[
OR = \frac{a/b}{c/d} = \frac{a d}{b c}
\]

**Case-Control Studies: Scenario 1 Question**

A case-control study was conducted of pediatric pneumococcal disease in Eastern United States in children under 6 years of age. A telephone questionnaire was administered to parents of 120 cases of the disease and 120 control children. The exposure of interest was attendance in a group daycare setting. 53 cases and 26 controls attended group daycare.

Create a 2x2 table and calculate the odds ratio for this study.
Case-Control Studies: Scenario 1

Answer

A case-control study was conducted of pediatric pneumococcal disease in Eastern United States in children under 6 years of age. A telephone questionnaire was administered to parents of 120 cases of the disease and 120 control children. The exposure of interest was attendance in a group daycare setting. 53 cases and 26 controls attended group daycare.

Create a 2x2 table and calculate the odds ratio for this study.

<table>
<thead>
<tr>
<th>Exposure: Group Daycare</th>
<th>Cases</th>
<th>Controls</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>53</td>
<td>26</td>
<td>79</td>
</tr>
<tr>
<td>No</td>
<td>67</td>
<td>94</td>
<td>161</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>120</td>
<td>120</td>
<td></td>
</tr>
</tbody>
</table>

OR = \( \frac{53 \times 94}{67 \times 26} = \frac{4983}{1742} = 2.86 \)

Children under 6 years of age who attended a group daycare program had 2.86 times the risk of developing pneumococcal disease as children who did not attend group daycare.

Case-Control Studies: Scenario 2

Question

A case-control study was conducted on characteristics associated with bladder cancer risk. One characteristic of interest was educational attainment. Of 760 bladder cancer cases, 400 had less than a high school education. Of 1737 controls, 560 had less than a high school education.

Create a 2x2 table and calculate the odds ratio for this study.

<table>
<thead>
<tr>
<th>Exposure: Education &lt; High School</th>
<th>Cases</th>
<th>No Cases</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>400</td>
<td>560</td>
<td>960</td>
</tr>
<tr>
<td>No</td>
<td>360</td>
<td>1177</td>
<td>1537</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>760</td>
<td>1737</td>
<td></td>
</tr>
</tbody>
</table>

OR = \( \frac{400 \times 560}{360 \times 1177} = \frac{470,800}{20,600} = 2.34 \)

Those with less than high school education had more than twice the risk (2.34) of developing bladder cancer as those with more education.

Case-Control Studies: Scenario 2

Answer

A case-control study was conducted on characteristics associated with bladder cancer risk. One characteristic of interest was educational attainment. Of 760 bladder cancer cases, 400 had less than a high school education. Of 1737 controls, 560 had less than a high school education.

Create a 2x2 table and calculate the odds ratio for this study.

<table>
<thead>
<tr>
<th>Exposure: Education &lt; High School</th>
<th>Cases</th>
<th>No Cases</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>400</td>
<td>560</td>
<td>960</td>
</tr>
<tr>
<td>No</td>
<td>360</td>
<td>1177</td>
<td>1537</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>760</td>
<td>1737</td>
<td></td>
</tr>
</tbody>
</table>

Those with less than high school education had more than twice the risk (2.34) of developing bladder cancer as those with more education.
Case-Control Studies: Scenario 3

Question

A matched pairs case control study was conducted of colon cancer risk and eating high fat foods. 200 colon cancer cases and 200 control individually pair matched by age (+/- 2 years) were identified. Of the study participants: 30 colon cancer cases and their matched controls ate high fat foods every day; 85 colon cancer cases and their matched controls ate high fat foods less than every day; 60 colon cancer cases ate high fat foods daily & their control ate them less often; 25 colon cancer cases ate high fat foods less than every day & their control ate high fat foods every day.

Create a 2x2 table and calculate the matched pairs odds ratio for this study.

Answer

A matched pairs case control study was conducted of colon cancer risk and eating high fat foods. 200 colon cancer cases and 200 control individually pair matched by age (+/- 2 years) were identified. Of the study participants: 30 colon cancer cases and their matched controls ate high fat foods every day; 85 colon cancer cases and their matched controls ate high fat foods less than every day; 60 colon cancer cases ate high fat foods daily & their control ate them less often; 25 colon cancer cases ate high fat foods less than every day & their control ate high fat foods every day.

Create a 2x2 table and calculate the odds ratio for this study.

\[
\begin{array}{c|cc}
 & \text{Exposed: Yes: Daily high fat foods} & \text{No: Less than daily high fat foods} \\
\hline
\text{Cases} & 30 & 60 \\
\text{Controls} & 25 & 85 \\
\end{array}
\]

\[
\text{OR} = \frac{30 \times 85}{25 \times 60} = 2.4
\]

Individuals who ate high fat foods every day had 2.4 times the risk of developing colon cancer compared to those who ate high fat foods less often.

Nested Case-Control Studies

- A case-control study conducted within a cohort study
- The cohort study provides the roster for control selection
Nested Case-Control Studies

- Cases consist of all incident cases generated by the source population (cohort) over the study period.

For each case, a set of controls is selected from subjects at risk at the time of the disease occurrence.

- Same individual can randomly be selected as a control for more than one case.

- A participant is eligible to be a control and when they develop the disease become a case.

Selecting controls in nested case-control studies: risk set sampling

- Risk set sampling for controls—controls selected from population at risk as cases are diagnosed.

- No need for rare disease assumption for OR to estimate RR.
Three Approaches for Selecting Controls

- Summer Sampling
  - Start of follow-up
  - End of follow-up
  - Controls selected from survey

- Base Sampling
  - Start of follow-up
  - End of follow-up
  - Controls sampled from population at risk

- Case Set Sampling
  - Start of follow-up
  - End of follow-up
  - Controls sampled from population at risk as cases are diagnosed

---

Case-Cohort Studies

- An alternative to the nested case-control study without matching on time
- Controls selected from those at risk at the beginning of the study period, usually a random sample of all members of the cohort
- As controls are a random sample of the study base, they can serve as controls for multiple diseases
Three Approaches for Selecting Controls

Case-Crossover Studies

- Developed for situations “when brief exposure causes a transient change in risk of a rare acute onset disease”

- Were exposures immediately preceding the event different from those exposures which usually occur?”

Case-Crossover Studies

- Each person serves as own control
- A variant of the matched-pairs design
- A case’s person time is divided into an index period (case) and a reference period (control)
- The brief period of increased risk following the transient exposure is called the hazard period
- The exposure frequency during the hazard period is compared to a control period
Case-Crossover Studies

- Used in studies looking at events such as risk of acute MI immediately following heavy physical exertion, risk of car accidents while using a cell phone, risk of occupational accidents

Case-Crossover Studies

- Cell phone use and auto accidents*
  - 699 drivers studied who had cell phones and were involved in accidents
  - Interviewed subjects and reviewed records of the accident and cell phone billing records
  - Hazard period was 10 minutes before the accident
  - Control period was one week before the collision


Case-Crossover Studies

- Risk of collision was 4 x higher when using a cell phone than when not used
  - Excess risks seen in both men and women, all ages and driving experience
For Additional Practice

For additional practice, go to:
http://www.cdc.gov/excite/classroom/index.htm
and try the exercise: "Cigarette Smoking and Lung Cancer". There are several concepts you have not yet learned, but you can answer most of the questions. When you are done, check your answers against the answer key.

NO AUDIO ON THIS SLIDE. CLICK NEXT TO CONTINUE