Strategies for data analysis: Case-control study

Postgraduate Research Training in Reproductive Health
Faculty of Medicine, University of Yaounde

Pierre Marie Tebeu (M.D)
pmtebeu@yahoo.fr
Specific objectives

- Identify the even or parameter for including in the study.
- Identify the exposure parameter differentiating case to control.
- Identify baseline characteristics of the population study
- Calculate the odd, the odds ratio,
- Interpreting the odd ratio,
- Identify sub-groups for analysis
- Identify confounders
What is a case? Control?(1)

- Tebeu et al. In process (Data from CHU Yaounde 1980-2000).

- Aim: To test the hypothesis stating that women aged 40 and above are at risk of delivery by caesarean section compared to those aged 20-29
What is a case? Control?(2)

- Caesarean section is the event to be investigated in the two age groups.
- Then a woman delivered by C/S in any of both groups is a case.
- And a woman delivered vaginally in any of both group is a control.
Illustration of the definition of case and control

- **NB:** 2 by 2 table for summarising data from case-controls study

<table>
<thead>
<tr>
<th></th>
<th>Case (Even)</th>
<th>Controls (No even)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>a</td>
<td>b</td>
<td>a+b</td>
</tr>
<tr>
<td>Not exposed</td>
<td>c</td>
<td>d</td>
<td>a+d</td>
</tr>
<tr>
<td>Total</td>
<td>a+c</td>
<td>b+d</td>
<td>-</td>
</tr>
</tbody>
</table>

Risk-of exposure

- $R1 = \frac{a}{a+c}$
- $R2 = \frac{b}{b+d}$

If been exposed increases the risk of even, then $R1 > R2$
Example of calculation of the risk to have been exposed (Tebeu et al.)

<table>
<thead>
<tr>
<th></th>
<th>C/S =Cases</th>
<th>No C/S =controls</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed to the risk (40-46)</td>
<td>29</td>
<td>153</td>
<td>182</td>
</tr>
<tr>
<td>Non exposed to the risk (20-29)</td>
<td>55</td>
<td>494</td>
<td>549</td>
</tr>
<tr>
<td>Total</td>
<td>84</td>
<td>647</td>
<td>-</td>
</tr>
</tbody>
</table>

Risk-of exposure
R1 = 29/84 = 34.52%
R2 = 153/657 = 23.29%

R1 > R2: Aged 40-46 is associated with the risk of delivery by caesarean section
### What is Odds?(1)

<table>
<thead>
<tr>
<th></th>
<th>case</th>
<th>Control</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed (E+)</td>
<td>a</td>
<td>b</td>
<td>a+b</td>
</tr>
<tr>
<td>Unexposed(E-)</td>
<td>c</td>
<td>d</td>
<td>c+d</td>
</tr>
</tbody>
</table>

- **Odds of exposure among cases**: $\frac{\text{probability (P) of being exposed}}{\text{probability of not being exposed}}$.

- **Risk of a case being exposed**: $\frac{a}{a+c}$,

- **Risk of a case being not exposed**: $\frac{c}{a+c} = 1 - \text{risk of a case being exposed} = 1 - \left[\frac{a}{a+c}\right]$
## What is Odds? (2)

<table>
<thead>
<tr>
<th>Case</th>
<th>Control</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed (E+)</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>Unexposed (E-)</td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

- Odds = probability (P) of being exposed / probability of not being exposed.

- Odd that case being exposed = \( \frac{a}{a+c} : \frac{c}{a+c} = \frac{a}{a+c} : \{1 - [a:(a+c)]\} = \frac{a}{c} \)

- By the same procedure: Odd that control been exposed = \( \frac{b}{d} \)
What is Odds ratio (OR)?

<table>
<thead>
<tr>
<th></th>
<th>Case</th>
<th>Control</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed (E+)</td>
<td>a</td>
<td>b</td>
<td>a+b</td>
</tr>
<tr>
<td>Unexposed (E-)</td>
<td>c</td>
<td>d</td>
<td>c+d</td>
</tr>
</tbody>
</table>

- **Odds ratio** is the Ratio of the Odds of exposure = Case Odd / control Odd

  - Odds that case been exposed: \( \frac{a}{c} \)
  - By the same process: Odds that control been exposed: \( \frac{b}{d} \)

  - Then \( (OR) = \frac{(a/c)}{(b/d)} = \frac{ad}{bc} \)
Interpreting the odds ratio

- If OR = 1: the exposure is not related to the event

- If OR > 1: the exposure is associated with increase of the event (Probable causation)

- If OR > 1, then the exposure is associated with decrease of the event (probable protection)
Example of interpretation of Odds Ratio (OR) (Tebeu et al.)

<table>
<thead>
<tr>
<th>C/S</th>
<th>Risk of being exposed</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>23.29%</td>
<td>1</td>
</tr>
<tr>
<td>Yes</td>
<td>34.52%</td>
<td>1.7</td>
</tr>
</tbody>
</table>

-Women delivered by C/S had 70% risk to have being exposed (i.e.: aged 40 years and above)

-Aged 40 years or above is associated with increased risk of delivery by C/S
Stratification is a division of case and control by sub-groups according to one of one parameter for better analysis.

- Example: Parity as parameter of stratification
- Sub-group-1: Primiparous women
- Sub-group-2: Multiparous women (2-5)
- Sub-group-3: Grand multiparous women (more than five)
What are Confounders?

- Some baseline characteristics that can significantly influence the development of the event,
- There is no strong frontier from sub-group to confounders.
Examples of confounders

- Preterm delivery
- Previous caesarean section
- The attempt to induce the labor
- The intercurrent disease associated
- The history of uterine surgery,
- Qualification of the staff who followed the pregnancy, ....
Identification and use of Confounders

- Confounders are identified by their epidemiological impact (age, differentiation of the tumor, myometrial invasion......)
- They can also be identified at univariate analysis (by comparing the outcome in subset of patients presenting the characteristic or not)
- They are then using for more detailed analysis (multivariate analysis, but need software)
Limits of case-control study

- Missing to follow up
- Missing of confounders
- Does not provide the risk of developing the event if exposed to the risk
Conclusion

- Case-control study can be conducted in a rural health center.
- Analysis of data from case-control study can be easily performed in a setting with no existing calculator.
Useful links:

- [http://www.gfmer.ch/Medical_education_En/PGC_Yaounde_2004.htm](http://www.gfmer.ch/Medical_education_En/PGC_Yaounde_2004.htm)