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

	Case (Even)	Controls (No even)	Total
Exposed	a	b	a+b
Not exposed	c	d	a+d
Total	a+c	b+d	-
Risk-of exposure	R1=a/a+c	R2=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.)

.	C/S =Cases	No C/S =controls	Total
Exposed to the risk (40-46)	29	153	182
Non exposed to the risk (20-29)	55	494	549
Total	84	647	-
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)

	case	Control	Total
Exposed (E+) =	a	b	a+b
Unexposed(E-) =	c	d	c+d

- Odds of exposure among cases= probability (P) of being exposed/ probability of not being exposed.
- Risk of a case being exposed: = a/a+c,
- Risk of a case being not exposed: = c/a+c= 1-risk of a case being exposed =1-[a:(a+c)]

What is Odds? (2)

	case	Control	Total
Exposed (E+) =	a	b	a+b
Unexposed(E-) =	c	d	c+d

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

- Odd that case being exposed = (a/a+c):(c/a+c)=(a/a+c):{1-[a:(a+c)]}=a/c
- By the same procedure: Odd that control been exposed=b/d

What is Odds ratio(OR)?(3)

	Case	Control	Total
Exposed (E+) =	a	b	a+b
Unexposed(E-) =	c	d	c+d

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

  Odds that case been exposed: =a/c

- By the same procss: Odds that control been exposed=b/d

- Then (OR)= (a/c) : (b/d)= ad / bc

Interpreting the odds ratio

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

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

-If OR>1, then the exposure is associated with decrease of the even (probable protection)

Example of interpretation of Odds Ratio (OR) (Tebeu et al.)

C/S	Risk of being exposed	OR
No	23.29%	1
Yes	34.52%	1.7

-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

- 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 provided the risk of developing the event if exposed to the risk

Conclusion

- Case-control study can be conduct 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.ccnmtl.columbia.edu/projects/episim/study2f.html
- http://www.gfmer.ch/Medical_education_En/PGC_Yaounde_2004.htm
- http://bmj.bmjjournals.com/collections/statsbk/8.shtml