CASE CONTROL STUDY

Learning Objectives

- Identify the principles of case control design
- State the advantages and limitations of case control study
- Calculate the sample size for a case control study (Epi program)
- Identify the characteristics of cases that will be selected
- Identify the characteristics of the controls that will be selected
- Discuss the issue of matching in case control study
- State the source of bias in a case control study
- Analyze data from a case control study
- Identify the repercussion of the limitations on the interpretation of the findings

Performance objectives

1. Design and implement a case control study and interpretation of the findings
2. Estimate the sample size required for a case control study

PRINCIPLE OF CASE CONTROL STUDY

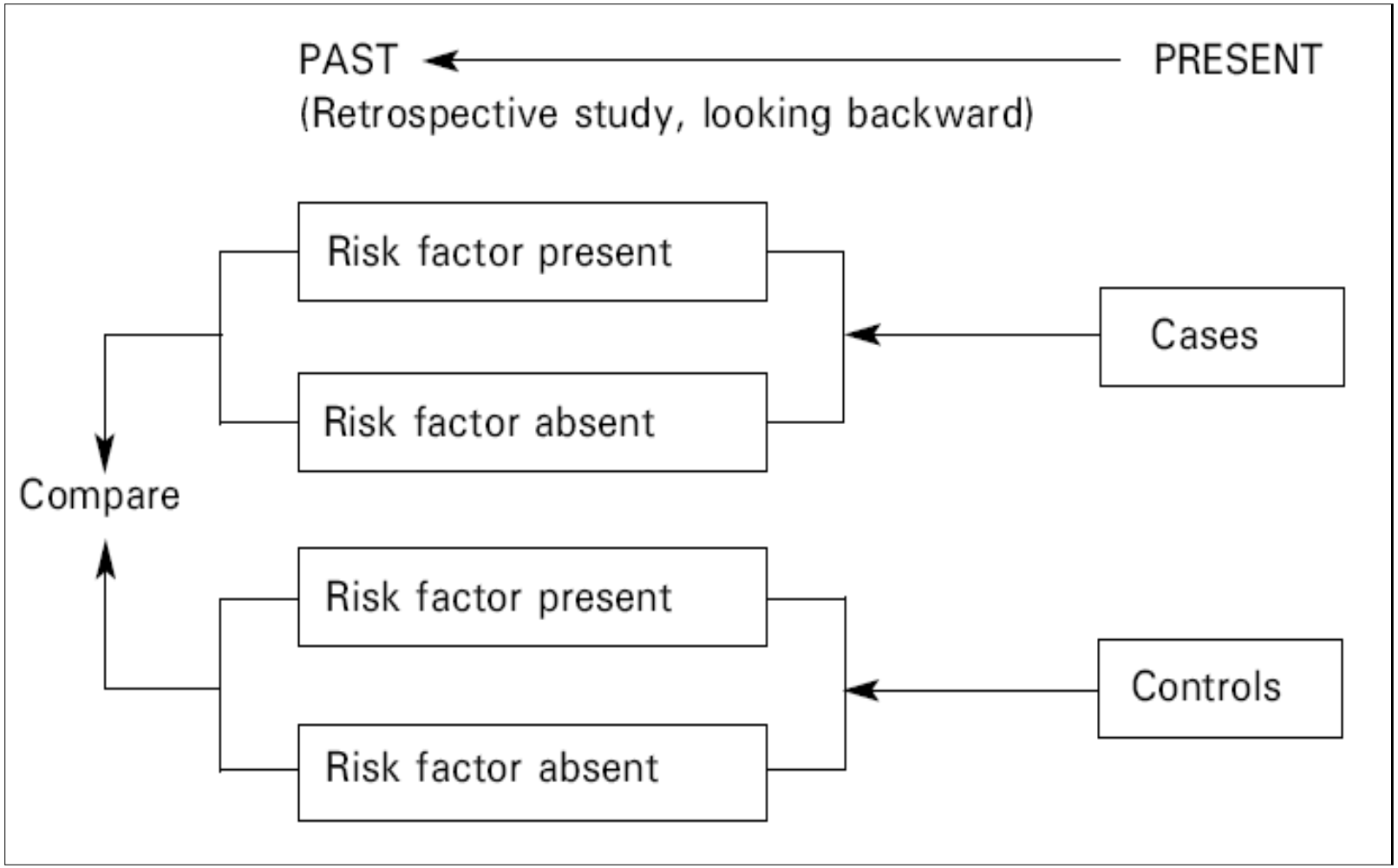
Selection of individuals on the basis of having the disease (cases) and not having the disease (controls)and comparing them in respect to the presence of risk factors.

DEFINITION OF CASE CONTROL STUDY

Analytic research design used for testing a hypothesis regarding a particular health problem and single or multiple exposures.

Starts with the outcome (health problem) then look backward for the exposure to identify potential risk or protective factors.

Provide answer or an explanation about “why” one group of people is affected and not the other.


PAST
PRESENT
(Retrospective study, looking backward)
Risk factor present
Risk factor absent
Cases
Compare
Risk factor present
Risk factor absent
Controls

PROCEDURE OF IMPLEMENTATION

1. Define objectives and research questions
2. Determine the “cases” for a case control study
3. Determine the “controls” for a case control study
4. Decide on the ratio of cases to control
5. Decide on matching cases and controls
6. Estimation of the sample size
7. Select cases and controls
8. Obtain data on exposures
9. Analyze and interpret the findings

OBJECTIVES AND RESEARCH QUESTIONS

The overall aim is to identify and quantify the risk factor(s) associated with the occurrence of a health problem

- Research question
  - what are the risk factors associated with the occurrence of a health problem?

  or
  - is the rate of exposure to a particular factor differ between cases and control?”

- Aim : identifying predictors of a health problem

DETERMINE WHO IS THE CASE

- Establish a “standard case definition”
- Adopt a “standard diagnostic criteria”
- Set inclusion and exclusion criteria
- Decide on the type of cases (incident or prevalent cases)
  - Incident cases reflect the determinant of the disease
  - Prevalent cases reflect the determinant and disease duration

DETERMINE WHO IS THE CONTROL

- The ideal “controls” are the healthy however
- Enroll controls who are by definition “sick” yet with specifying conditions of inclusion
- Free from the health problem under investigation
- Free from health problems known to be associated with the exposure
- Being comparable to cases in terms of susceptibility

DECIDE ON CASE CONTROL RATIO

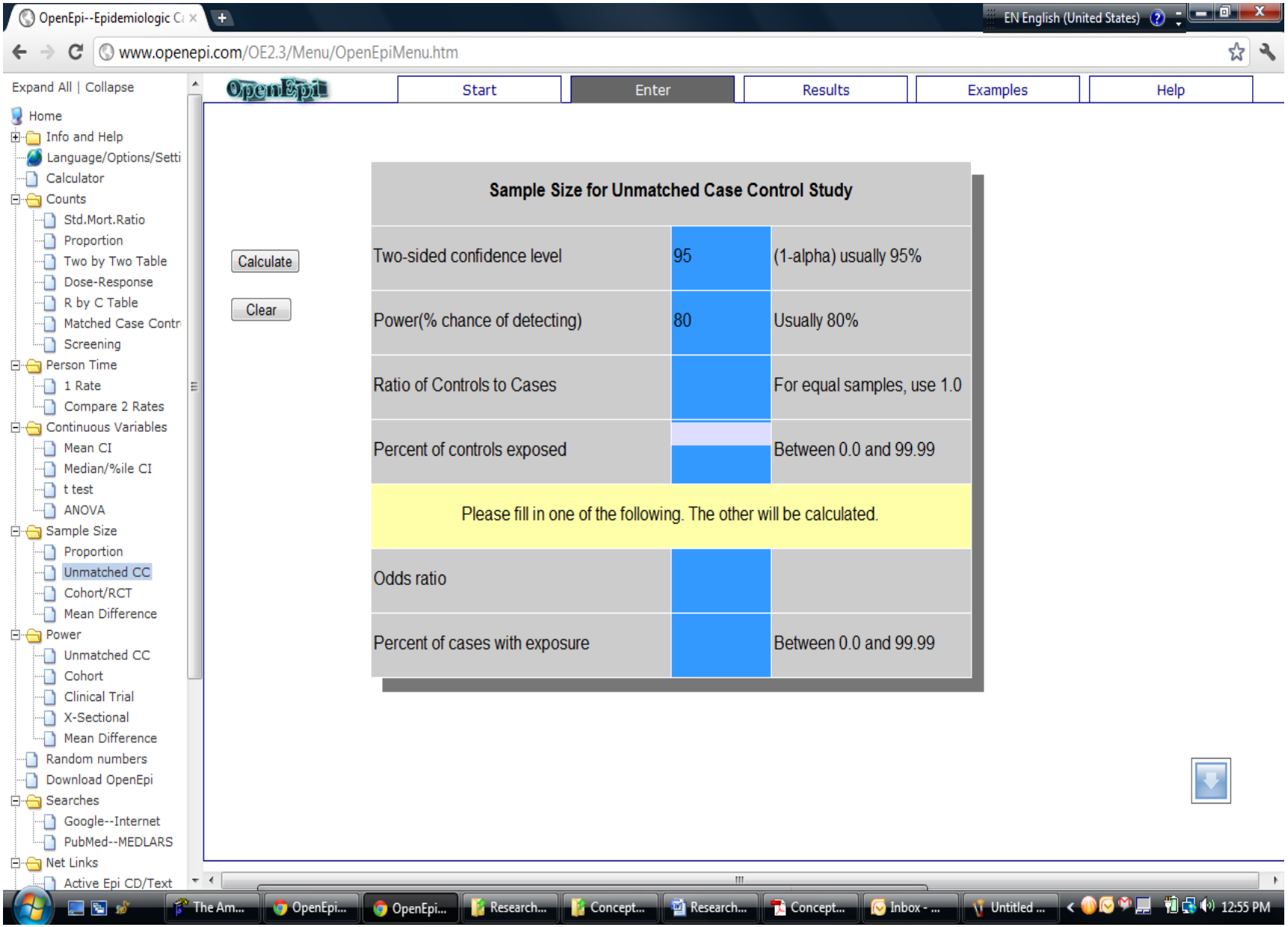
- The ratio of cases to control should be at least 1:1
- Increase in the ratio lead to increase in “study precision”
  1:2
  1:3
  1:4
- Further increase is associated with little increase in study precision relative to the cost involved

MATCHING IN CASE CONTROL STUDY

- Matching reduce the possible confounding effect
- Matching on several characteristics is not advisable as it
  - Creates difficulties in finding controls
  - Requires more complex statistical analysis
  - May result in overmatching
- Matching design should be followed by matched analysis
- Few statistical programs provide opportunities for matched analysis as STATA and SAS.

SAMPLE SIZE ESTIMATION

www.openepi.com



Two side confidence level

Power of the test

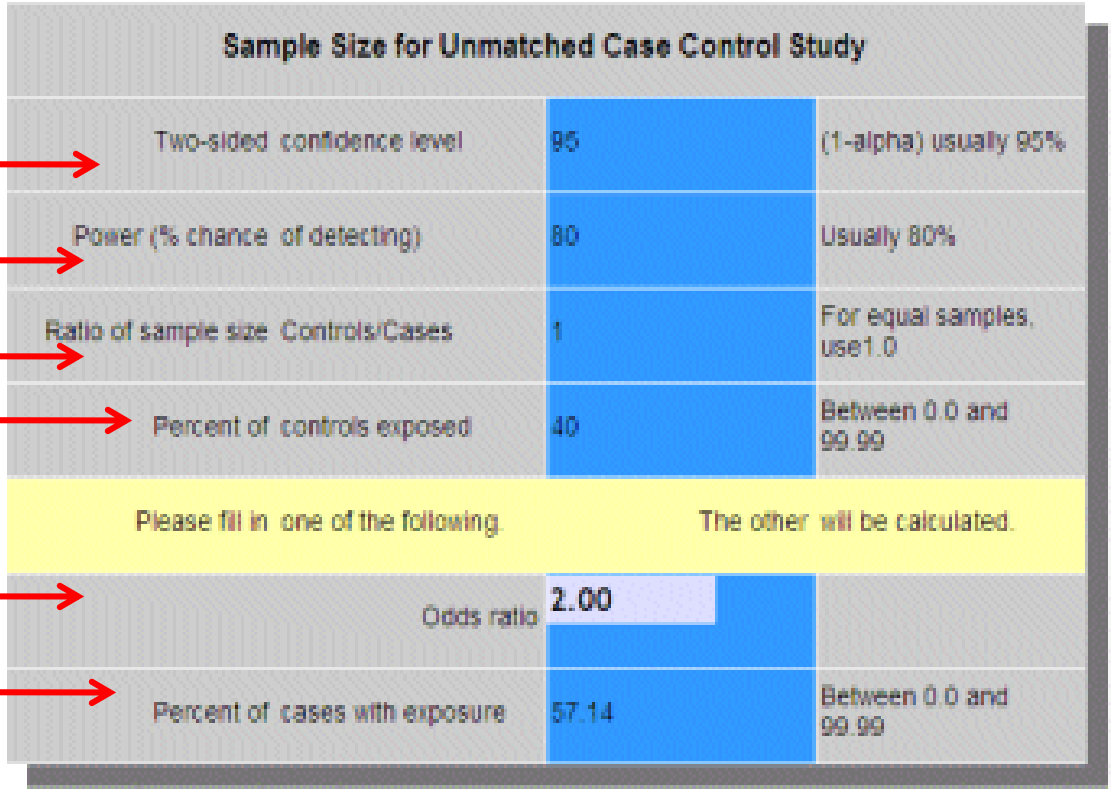
Ratio of cases to control

Percent of controls exposed

Then one of the followings

Odds ratio

Percent of cases exposed



SELECTION OF CASES AND CONTROL

Ensure that cases are representative

- Identifying cases through registries and records of health facilities to represent cases in the community

Or/

- Enroll those meeting the diagnostic criteria from those attending a specific health facility

SELECTION OF CASES AND CONTROL

- Selection of "healthy controls implies the selection of controls from healthy subjects in the general population

Or/

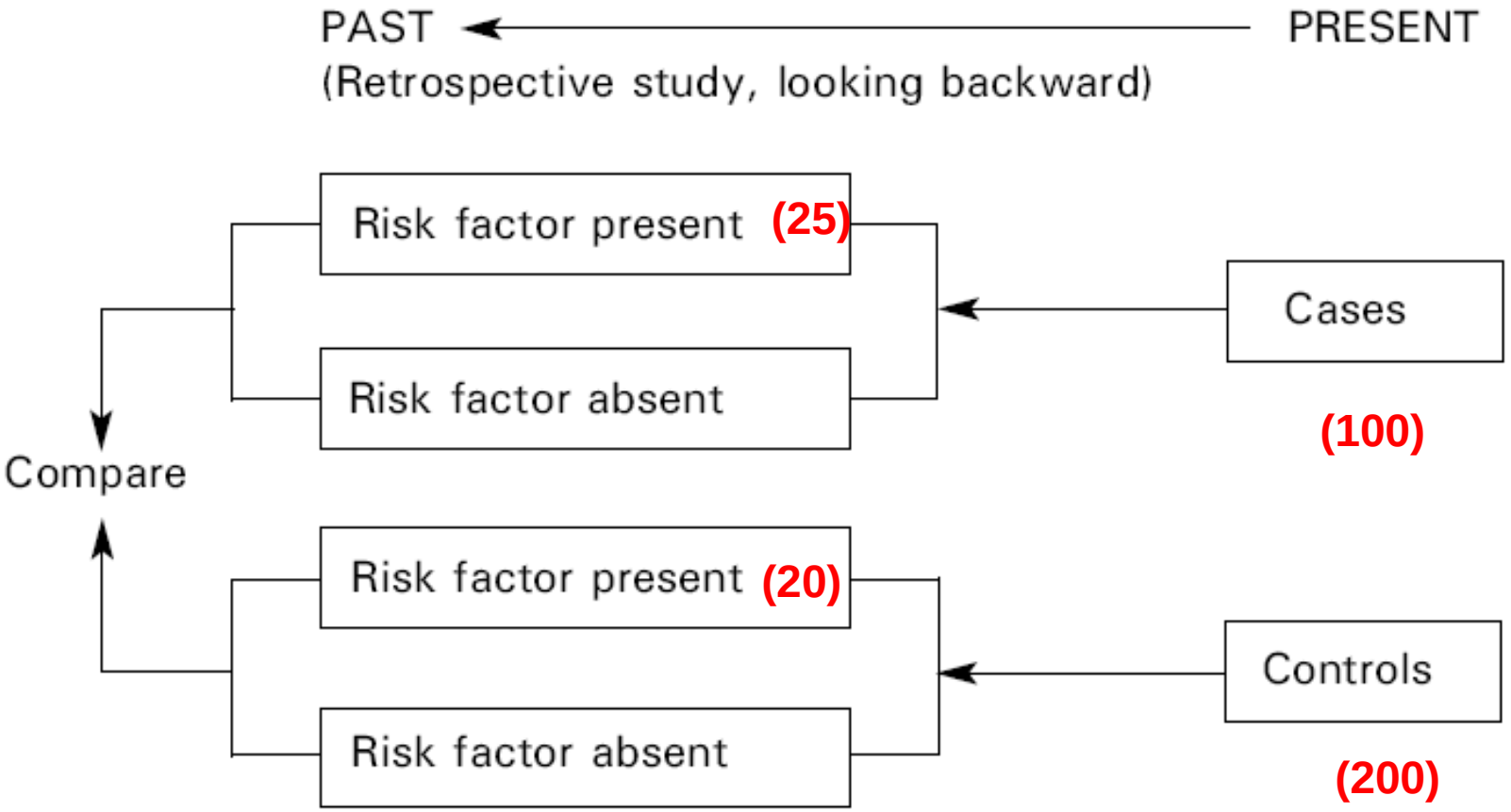
- Select controls meeting the inclusion criteria from those attending the same health facility attended by the cases.
  - Lower cost and time
  - Higher participation rate (minimize non-response)
  - Tend to recall past events better (minimize recall bias)
  - Tend to have over presentation of risk factors
  - Difficult in determining which appropriate illness for inclusion

- Control could be relatives and neighbors as they are comparable to cases in respect to a large number of exposures

OBTAINING DATA ON EXPOSURE

- Asking questions of relevance to the exposure using
  - Self administered questionnaire
  - Interviewing participants either face to face or by telephone
- Review of medical or employment records
- More than one source of information can be use to complement each others
- In deciding on the source one should consider availability, accuracy as well as the logistics and cost of data collection.

ANALYSIS OF CASE CONTROL STUDY



QUANTIFICATION OF RISK

Exposure	Disease status	
	Case	Control
Exposed	25 (a)	20 (b)
Not exposed	75 (c)	180 (d)
Total	100 (a+c)	200 (b+d)

Rate of exposure among cases = (a/a+c) x constant

(25/100) X 100 = 25%

Rate of exposure among controls= (b/b+d) x constant

(20/200) X 100 = 10%

Odds ratio (OR) = (axd)/(cxb)

(25x180)/(20x75)= 3.0

QUANTIFICATION OF RISK

Exposure	Disease status	
	Case	Control
Exposed	25 (a)	20 (b)
Not exposed	75 (c)	180 (d)
Total	100 (a+c)	200 (b+d)

Odds ratio (OR) = (axd)/(cxb)

(25x180)/(20x75)= 3.0

Odds of exposure among cases= $\frac{(a/a+c)}{(c/a+c)} = \frac{a}{c}$

Odds of exposure among controls= $\frac{(b/b+d)}{(d/b+d)} = \frac{b}{d}$

Odds ratio= $\frac{(a/c)}{(b/d)} = \frac{ad}{cb}$

INTERPRETATION OF ODDS RATIO

Exposure	Disease status	
	Case	Control
Exposed	25 (a)	20 (b)
Not exposed	75 (c)	180 (d)
Total	100 (a+c)	200 (b+d)

Odds ratio (OR) = (axd)/(cxb)

(25x180)/(20x75)= 3.0

Interpretation of OR

< 1	Protective
=1	Not related
>1	Risk

ELABORATION ON ODDS RATIO

OR > 1

Percentage increase of disease as a result of exposure

$$[(OR - 1) * 100]$$

OR < 1

Percentage decrease of disease as a result of exposure

$$[(1 - OR) * 100]$$

Utilities

- Study of rare diseases
- Study of diseases with long latency period
- Evaluate all possible factors associated with the disease
- Quantification of the risk associated with exposure (s)
- Save cost and time
- No value in the study of rare exposure
- Not for study of several diseases associated with a single exposure

Limitations

- Cases don't represent cases in the general population
- Selection bias
- Recall bias
- Inability to define the temporal sequence between the disease and the exposure