Analytical Epidemiological studies

Objectives

1. Understand the basic design features of case control study.
2. Draw flow chart for each of the studies.
3. Understand the rationale for applying case-control designs.
4. Calculate the odds ratio as a measure of association in case control studies.
5. Example applications applying case-control designs.
6. Recognize the advantages and disadvantages of case control studies.

Analytic Epidemiology

- The study of the **determinants** of the disease or health- related events. In the analytical studies, the investigator **observes** the natural course of events i.e. consider that temporal relationship.
- Second major type of epidemiological studies.
- In contrast to descriptive studies (look at entire population), subject of interest is the individual inside population.
- Objective is to test and not to formulate hypothesis.
- The key feature of analytic epidemiology is a comparison group

Types of analytical studies (analytic study designs):

✓ Case- control study.

✓ Cross-sectional study (prevalence survey).

✓ Cohort study.

✓ **Schematic design of case-control and cohort studies.**

Past ←Present→ future

- **Basic Question in Analytic Epidemiology:**

  Are exposure and disease linked?

<u>Direction of inquiry in case control study.</u>

D → E

Disease

E.g. Myocardial **Exposure**

Infarction (MI) ***Risks e.g. Smoking***

<u>Case control Study</u>

❖ **<u>Definition:</u>**

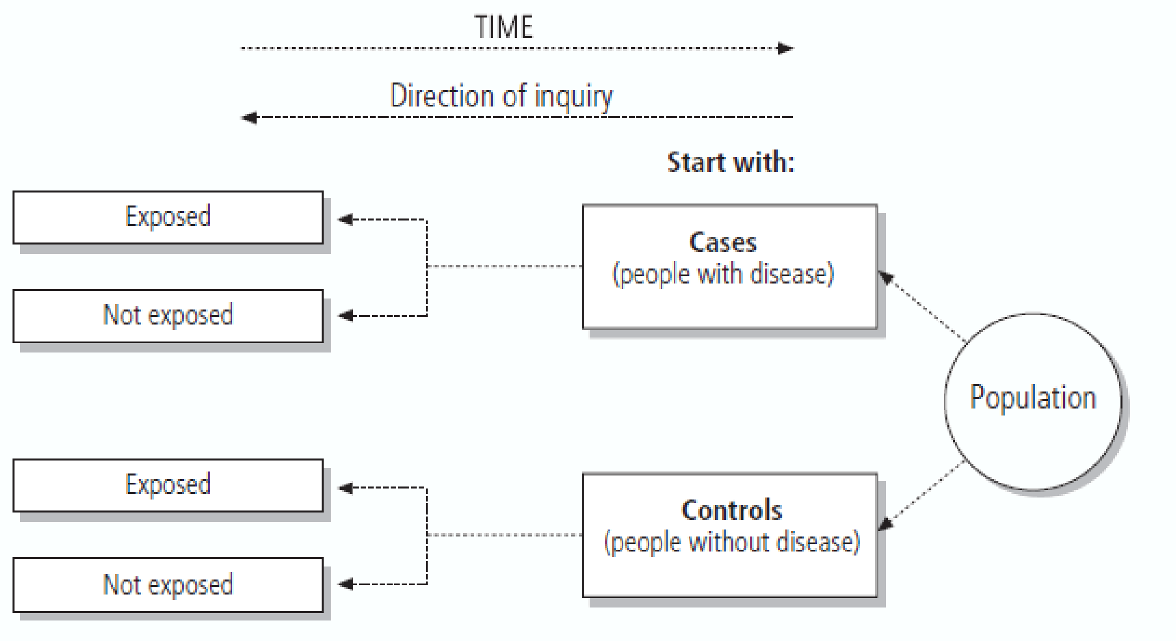
- Case control study is an *<u>observational epidemiological</u>* study in which subjects are selected on the basis of whether they have the condition (e.g. cases with disease or any health related events) or *<u>free</u>* from the condition (the control).
- The groups are then compared with respect to the proportion having a history of an exposure or certain characteristic
- Often called **"retrospective study".**

<u>*Characteristic features:</u>

- Both exposure and outcome have occurred before the start of study.
- Study proceeds backwards from effect to cause.

❖ **<u>Case control Study Design:</u>**

Figure 3.5. Design of a case-control study



***Steps to conduct case control study**:

1-*Selection of cases*:

These cases should represent all the cases in a specified population group. Cases are selected on the basis of disease, not exposure. **This includes:**

a. Establishment of **diagnostic criteria** (standard case definition) and must be specified before study e.g. the documented criteria of diagnosis of myocardial infarction are electrocardiograph abnormalities, enzyme changes and characteristic chest pain and so on.

b. **Eligibility criteria:** e.g requirement that only incident (newly diagnosed) cases within a specified period, or cancer of a specified stage of disease.

c. **Sources of cases:** This includes:

i) Hospitals or any health care facility.

All cases admitted during a specific period of time or a ***random sample*** is selected for the study.

ii) **General population:** a survey is carried out to identify the cases in certain area within a specific period of time.

2-*Selection of the control*: the control should be: ***Free from the studied disease***.

*****The steps for selection of control are:***

a. **Matching:** is required to ensure comparability between cases and control. *Matching:* is the process in which we select the controls in a way that they are similar to cases with regards to certain confounding factors (e.g. age) which are known to influence the outcome of the disease.

- Age could be a confounding factor, if we are investigating the relationship between contraceptives and breast cancer. If the control group is younger than the cases (women with breast cancer), they would be protected from breast cancer because of their lower age as breast cancer increases with age. So we have to select the control to be of similar age to the cases to eliminate the effect of this confounding factor (age).

- In the same way other confounding factors should be matched E.G. sex, social class,
- Matching can be expensive and time-consuming.

b. <u>Sources of the control</u>: The sources of the control are hospitals, relatives, neighbors and general population.

i) **Hospitals**: The controls are selected from the same hospital as the cases but with different illness other than the study disease. For example the control group of cancer breast women could be non-cancerous patient attending the same hospital.

ii) **Relatives or friend**: They are co-operative however they are unsuitable control when genetic conditions are under study.

iii) Neighborhoods: The controls are selected from the same locality of the cases, persons in the same factory, children attending the same school.

iiii) **General population:** ***random sample*** of individuals in ***same geographic area*** of cases and free from the disease under stud should be selected. (Expensive, time consuming, difficult)

d. **<u>Size of the control:</u>** Use one control for each case if the cases are available and large in number.

- ✓ Ratio cases to control -1 : 1 up to 1 : 4

If the number of cases is less than fifty, 2,3 or even 4 controls can be selected for each study subject (case).

3. <u>Assessment of the exposure:</u>

Information about the exposure may be obtained by interviews, by questionnaires, or by studying past records of cases such as hospital records, school or occupational records

4. <u>Analysis and interpretation of the results:</u>

a. Tabulation of data

b. Calculation of exposure rates among cases and controls

c. Estimation of the disease risk associated with exposure (odds ratio)

Tabulation of data

Framework of case control Study

Exposure (E)	Cases	Control
Exposed	**a**	**b**
Not exposed	**c**	**d**
Total	**a+ c**	**b+ d**

b. Calculation of exposure rate:

The rate of exposure among the cases=

$$\frac{\text{The number of those exposed among the cases X100}}{\text{The total number of cases}} = \frac{a}{a+c} \text{ x } 100$$

The rate of exposure among the controls=

$$\frac{\text{The number of those exposed among the control X100}}{\text{The total number of control}} = \frac{b}{b+d} \text{ x}100$$

c. Estimation of risk associated with exposure: (Odds Ratio)

- Odds ratio is the measure of the strength of the association between the risk factor and the disease.
- The odds ratio is sometimes called the **cross-product ratio**

Odds Ratio= $\frac{ad}{bc}$

How to calculate the odds ratio?

- What is the odd that a case is being exposed?

$$\frac{a}{a+c} \div \frac{c}{a+c} = \frac{a}{c}$$

- What is the odd that a control is being exposed?

$$\frac{b}{b+d} \div \frac{d}{b+d} = \frac{b}{d}$$

-What is the estimated risk (odds ratio)?

$$\frac{a}{c} \div \frac{b}{d} = \frac{a\,d}{b\,c}$$

<u>If the value of odds is:</u>

- More than one, there is a risk.
- Less than One the exposure factor is protective
- Equal one there is no relation between the exposure factor and the disease.

Example

	Cases	Controls
Smokers	110	150
Never smokers	90	250
Total	200	400
% smokers	*55%*	*37.5%*

If disease is related to smoking then more cases than controls would be smokers

15/12/2010 Case control studies ١٥

***Measures of association**

- Relative risk
- Odds ratio
- We often need to know the relationship between an outcome and certain factors (e.g., age, sex, race, smoking status, etc.)

- Used to guide planning and intervention strategies

Measures of association

- Odds ratio ~ cross product ratio

	Cases	Controls
Exposed	110	150
Not exposed	90	250

OR= = ad/bc = [(110*250)/(150*90)] = 27500/ 13500 = 2

Odds of exposure in cases : two times compared to controls

15/12/2010 Case control studies 16

Interpretation of odds ratio (OR)

- OR > 1: means the exposure is a risk factor
- OR < 1: Decreased odds, or protective effects, among those with outcome
- OR = 1: No association between exposure and outcome(disease)

Note: "Exposure" is a broad term that represents any factor that may be related to an outcome.

- Remember that the relative risk can only be calculated in prospective studies
- Odds ratio can be calculated for any design
  - Cohort / prospective

- Case-control
- Cross-sectional

- **<u>APPLICATIONS</u>**

1. Evaluating Vaccine Effectiveness
2. Evaluations of Treatment & Program Efficacy
3. Evaluation of Screening
4. Outbreak Investigations
5. Indirect Estimation in Demography
6. Genetic Epidemiology
7. Occupational Health Research
8. Predictive Modeling

❖ **<u>Advantages of case control study:</u>**

1. Efficient, quick and cheap.
2. It is suitable for diseases with long latency. E.g. chronic diseases.
3. Suitable to study rare diseases.
4. Does not require large samples. .
5. It can examine multiple etiologic factors for a single disease (smoking, physical exercise, stress, blood lipids in CHD).
6. Requires no follow up of subjects, no attrition problem.
7. It faces minimal ethical problems.
8. No risk to subjects

❖ **<u>Limitations of case control study:</u>**

1. It cannot compute the incidence rates in exposed and non-exposed.
2. Errors in Case definitions.
3. Difficulties in selection of controls group
4. Other possible effects of exposure cannot be studied

5. The problem of bias is more common in case control study compared to other designs.

Bias is any systematic error in the determination of the association between the exposure and the disease.

- Recall bias: when cases and controls are asked about the exposure, it may be more likely for the cases to recall the existence of exposure than the control who are healthy persons.
- Interviewer's bias: This bias can be removed by matching.
- Selection bias: the cases and the controls may not be representative of cases and controls in the general population.
- Interviewer's bias: Bias may also occur when the interviewer knows the hypothesis and also knows who the cases are.

Example:

An investigator selected 100 patients with CHD admitted to hospital (X) during the year 2004, and 100 subjects free from the disease as a control from general population. Both groups were interviewed to obtain information on history of smoking. The smokers were 60 among cases and 20 among the control.

a. **Tabulate the previous data**
b. **Calculate exposure rate among case and control**
c. **Calculate Odds ratio**
d. **Analyze and interpret the results.**

A=

Exposure (E)	Cases	Control
Smoker	**60**	**20**
Not smoker	**40**	**80**
Total	**100**	**100**

B. Exposure rate in case = 60\ 100 X100 = 60%
Exposure rate in control: 20\100 X100 = 20%

C. Odds ratio= $\frac{ad}{bc}$

60x80
hufy-------- = 6

40x20

D: Smoker has more effects on CHD.

Example 2:

Use the data in table 3.15 to calculate the odds ratio.

Table 3.15 Exposure and Disease in a Hypothetical Population of 10,000 Persons

	Disease	No Disease	Total	Risk
Exposed	a=100	b=1,900	2,000	5.0%
Not Exposed	c=80	d=7,920	8,000	1.0%
Total	180	9,820	10,000	

Answer:

Odds ratio = (100 x 7,920) / (1,900 x 80) = 5.2

Example 3

Table 3.2. Association between meat consumption and enteritis necroticans in Papua New Guinea[11]

		Exposure (recent meat ingestion)		
		Yes	No	Total
Disease (enteritis necroticans)	Yes	50	11	61
	No	16	41	57
	Total	66	52	118

Use the data in the previous table to calculate the odds ratio.

Odds ratio = (50 x 41) / (11x16) = 11.6