TYPES OF EPIDEMIOLOGICAL STUDIES I
L 12

Associate Professor Dr. Eman A. Al-Kamil

Dep. Of Community Medicine

Collage of Medicine

Hashemite University

Types of Epidemiological studies

I. Observational studies :

The researcher observes and systematically collects information, but does not try to change the people

1. Descriptive studies

A. **Cross-sectional**

B. **Longitudinal**

2. Analytical studies

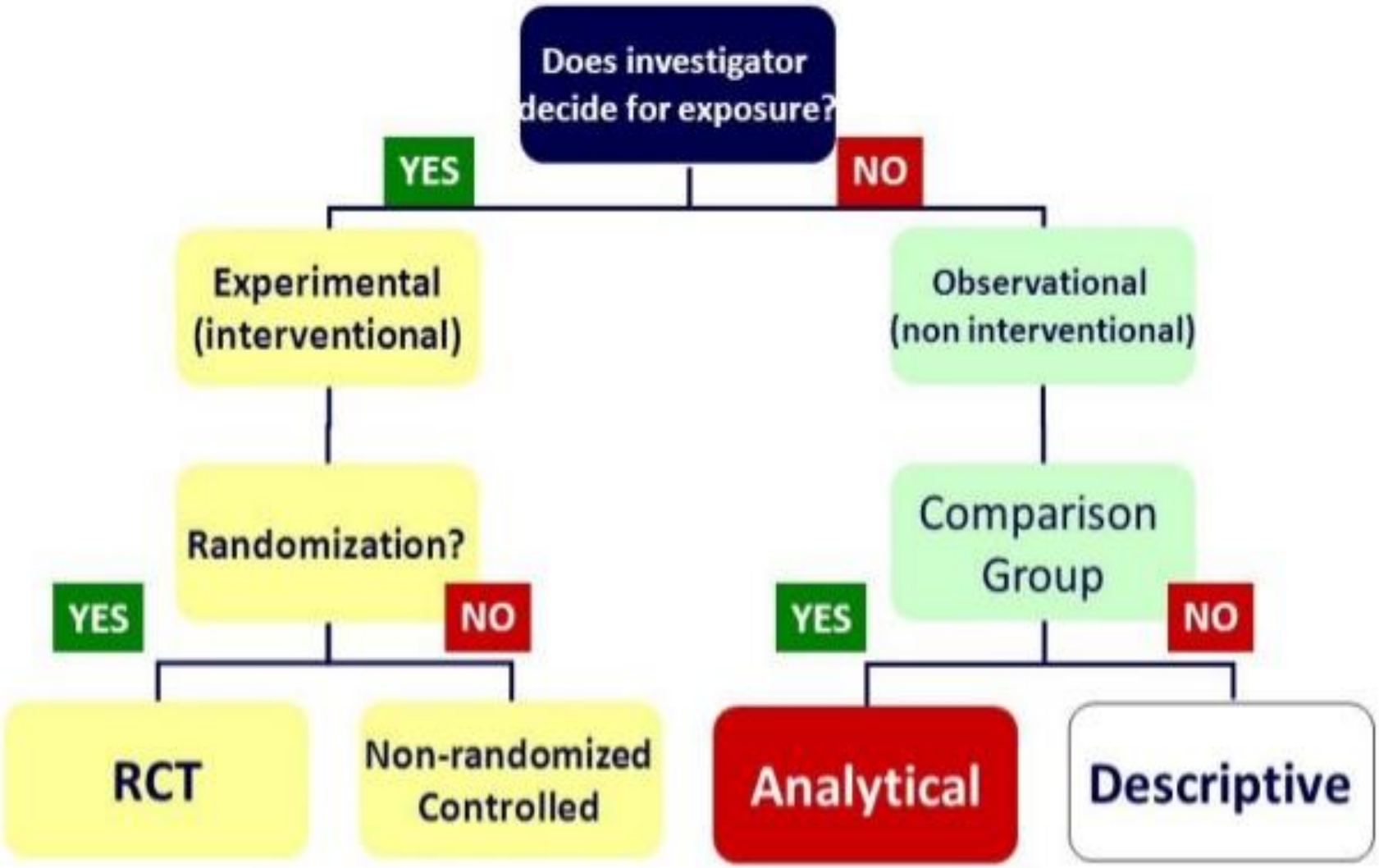
A. **Case - Control studies**

B. **Cohort studies**

II. Experimental or interventional studies:

- Involve an active trial to change disease determinant by the investigator who controls the exposure.
- Investigator allocates the exposure and follows the subjects.
- Participants are identified on the basis of their exposure status and followed to determine whether they develop the outcome or not.

Study Designs



Observational studies

1. Descriptive studies

- A simple description of the health status of a community, based on routinely available data (health related data) or on data obtained in special surveys.

- Pure descriptive studies make no attempt to analyze the links between exposure and effect.

- They are usually based on mortality statistics and may examine patterns of death by persons, time, and place, during specified time periods or in various countries.

- Two types:
  - **A. Cross-sectional**
  - **B. Longitudinal**

Cross-sectional studies:

- It assess the presence of disease and risk factors at a point of time.

- These are based on a single observation usually carried out in a short time.

They are characterized by the following:

a. It is used to study conditions that are relatively frequent with long duration of expression (nonfatal, chronic conditions),

b. measure prevalence of disease or related outcome.

c. They are not useful for diseases of short duration, **rare or highly fatal** .

A single observation may miss cases.

It tends to identify prevalent cases of long duration, since people who die quickly or recover quickly or who are no longer employed in a particular occupation are less likely to be identified.

d. They suggest hypotheses.

e. Their results are difficult to interpret because of seasonal variation and cohort effect.

f. They are relatively quicker and cheaper to do.

Advantages and Disadvantages of Cross-Sectional Studies

Advantages

1. Gives general description or scope of problem.
2. Useful in health service evaluation and planning.
3. Baseline for prospective study
4. Identifies cases and controls for retrospective study
5. Low-cost
6. quicker

Disadvantages

1. No calculation of risk.
2. Not good for rare diseases.
3. Selective survival can lead to bias.
4. Selective recall can lead to bias.

Longitudinal or follow up studies:

- These are based on repeated observation of the study population over a defined period of time.

- They start with a base-line data provided by initial cross-sectional study.

a. They measure incidence of disease or related outcome.

b. They suggest hypotheses.

c. They are relatively more expensive and difficult to organize.

d. They are not useful for diseases of rare occurrence.

e. The results are easier to interpret.

f. They can be useful to determine seasonal variation of disease and other health related outcomes.

2. ANALYTICAL STUDIES

In analytical studies, the researcher attempts to explore how and why a disease process is initiated or maintained in a given population or place.

In this type of epidemiology, we always :

1. use 2 groups, study group and comparative or control groups.

2. we test hypotheses so that they are accepted or not.

Hypotheses are accepted when we have adequate evidence to support them. When the evidence is inadequate, hypotheses are not accepted and further studies may be required.

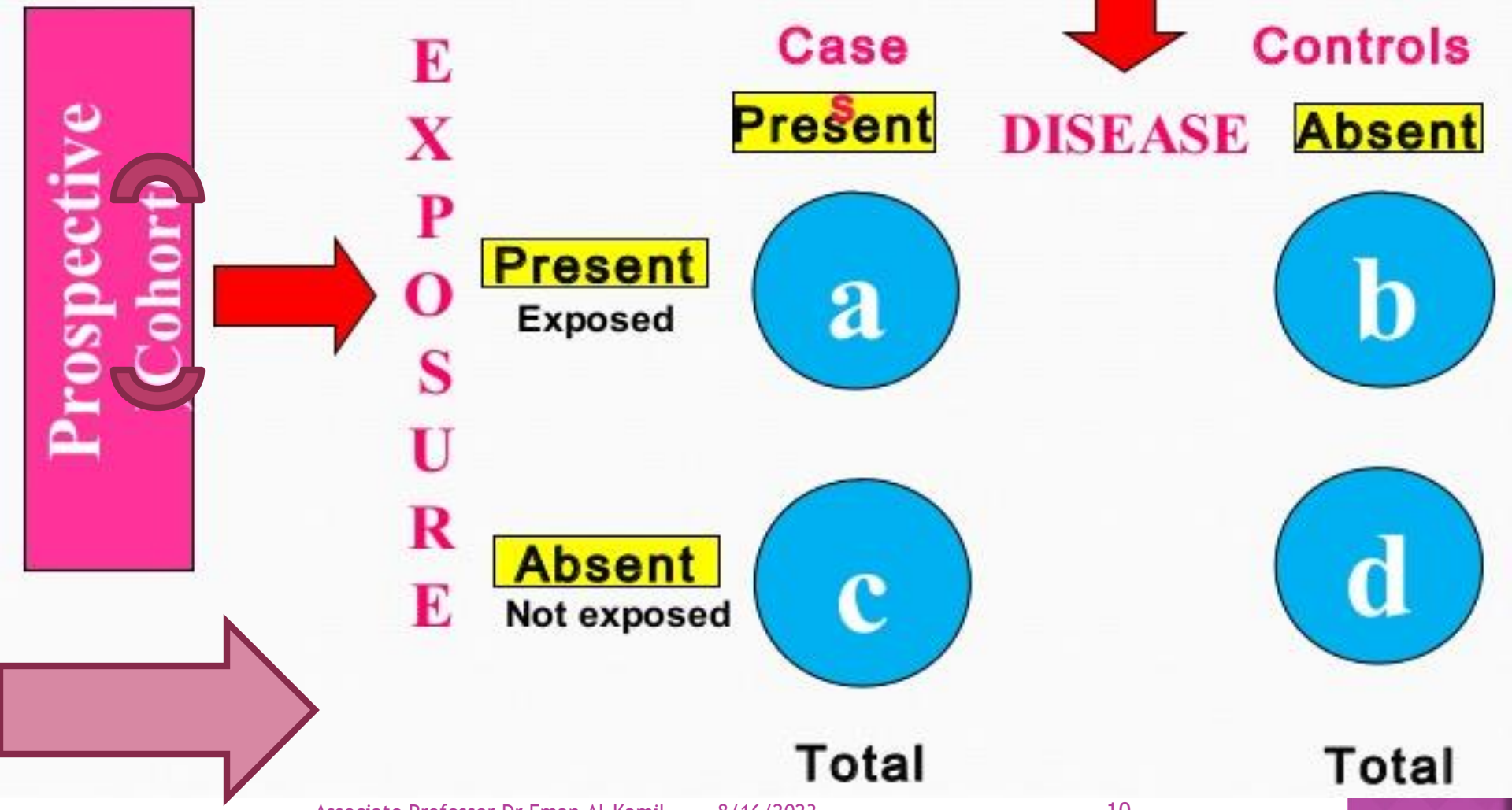
To use an example, it might be suggested (hypothesized) that parental smoking increases the risk of acute respiratory infection among children aged under five years. To test this hypothesis, two types of analytical epidemiological studies may be used:

A. **Case - Control studies**

B. **Cohort studies**

Retrospective Case-Control

A fourfold table



Case - Control studies

In case-control studies:

1. Both exposure and outcome (disease) have occurred before the start of the study.
2. The study proceeds backwards from outcome to cause (retrospective).
3. Epidemiologists survey a group of people with disease (cases) and a group without disease (controls) about their histories.

4. Controls are used to support or disprove any inference.

The survey may involve direct questioning or examination of medical or other records.

The basic question: *What differs in the histories of these two groups that could explain why one is diseased and the other is not?*

The basic design

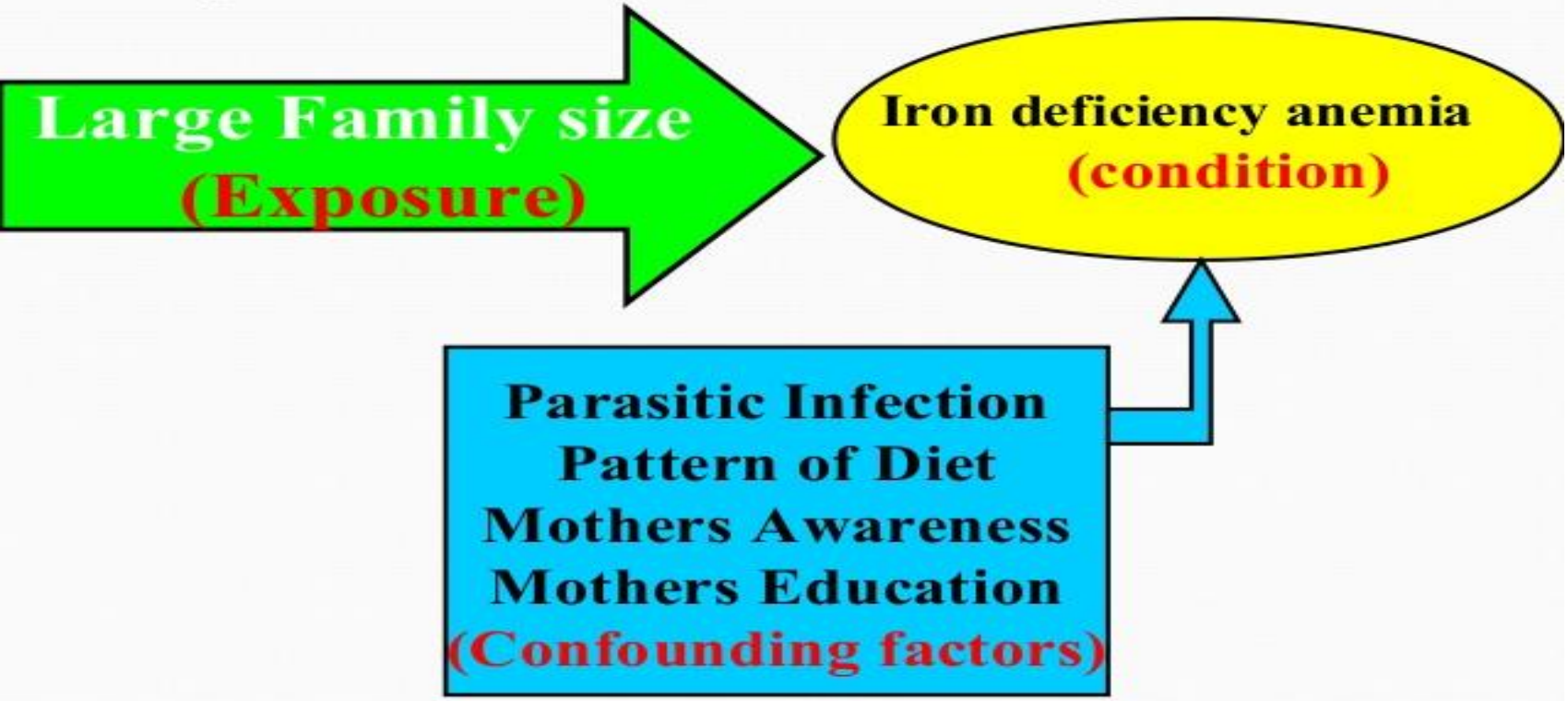
1. Two groups of persons are studied.

2. The first consists of subjects who have the disease under study at the time of the beginning of the study (cases).

3. The second group consists of subjects who are free from the disease under study (controls).

4. Both cases and controls are preferably matched for age and sex or other factors which may affect the results (confounding factors).
5. Some times they are matched for other variables but overmatching is to be avoided.

Impacts of the Confounding Factors



Confounding factors:

- Confounding occurs when the effects of two exposures (risk factors) have not been separated and the analysis concludes that the effect is due to one variable rather than the other.

Steps to conduct the case control study:

1-Selection of cases:

a. **Establishment of diagnostic criteria (standard case def.).**

b. **Sources of cases:**

i) Hospitals or any health care facility ii) General population:

2-Selection of the control:

a. **Matching.**

b. **Sources of the control.**

i) Hospitals ii) Relatives. iii) Neighborhoods.

c. **Size of the control**

3. Assessment of the exposure:

4. Analysis and interpretation of the results.

a. **Tabulation** of data

b. **Flow chart**

c. Calculation & interpretation of the estimated risk (**odds ratio**)

2-*Selection of the control*:

a. Matching: It is the process in which we select the control in a way that they have the same confounding factors affecting the cases (e.g. age) which are known to influence the outcome of the disease.

b. Sources of the control:.

i) Hospitals or any health care facilities.

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

iii) Neighborhoods

vi) *General population*: it is expensive, time consuming, difficult and the individuals may be uncooperative.

c. Size of the control:

If the number of the cases is >50 cases ,use one control for each case.

If the number of cases is < 50, use 2,3 or even 4 controls.

3. ***Assessment of the exposure:*** By interview, by questionnaires, or by studying past records of cases "hospital records, school or occupational records"

4. Analysis & interpretation of the results:

Tabulation of data:

Framework of case control Study

Exposure	Cases	Control
Exposed	a	b
Not Exposed	c	d
Total	a+c	b+d

b. Exposure rate:

Exposure	Cases	Control
Exposed	a	b
Not Exposed	c	d
Total	a+c	b+d

The rate of exposure among the cases =

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

The rate of exposure among the controls =

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

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

Measure of the strength of the association between the risk factor & the disease.

Odds ratio (OR) is synonymous to relative risk (RR)

The odds ratio = $\frac{ad}{bc}$

What is the odds that a case is being exposed?

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

- What is the odds 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**)?

$$\frac{a}{c} \div \frac{b}{d} = \frac{ad}{bc}$$

Exposure	Cases	Control
Exposed	a	b
Not Exposed	c	d
Total	a+c	b+d

Odds ratio

To illustrate the study design,

1. we identify 240 children who are suffering from acute respiratory infection (say pneumonia) cases
2. An equal or more number (380) of children matched for age and sex but are free from acute respiratory infection at the time of the study is also selected (controls).

3. Now, for children in both groups, the smoking habits of their parents are ascertained through careful interviewing of these parents. We try to know whether parent(s) do smoke or not and if they do, what is the number of cigarettes smoked per day. Suppose we found that the parents of 170 cases and 200 controls were smokers.

The analysis and interpretation:

The first step is to present the data in a 2x2 table

History of maternal smoking (Risk factor)	Cases of pneumonia (cases)	Children without pneumonia (controls)
Positive	170 a	200 b
Negative	70 c	180 d
Total	240 a + c	380 b+ d

a = no. of individuals with the disease have exposure to the studied risk factors = 170

c = no. of individuals with the disease but have no exposure to the studied risk factors =70

b= no. of individuals without disease but have exposure to the studied risk factors =200

d= no. of individuals without the disease and have no exposure to the studied risk factors = 180

The second step is to calculate the percentage of smokers (exposed) among parents of cases and controls.

$$\text{Percentage of smokers among parents of cases} = \frac{170}{240} \times 100 = 70.8\%$$

$$\text{Percentage of smokers among parents of controls} = \frac{200}{380} \times 100 = 52.6\%$$

It is clear that the habit of smoking was more frequent among parents of cases as compared to parents of controls.

Interpretation

Cases were more likely to be children of smoking parents.

The third step is to measure the strength of association between parental smoking and acute respiratory infection.

This is achieved by calculating a proxy measure to the relative risk. This measure is called the Odds ratio (OR).

$$OR = \frac{\text{Cases exposed (a)} \times \text{Controls not exposed (d)}}{\text{Cases not exposed (c)} \times \text{Controls exposed (b)}}$$

$$= \frac{170 \times 180}{70 \times 200} = 2.2$$

Interpretation

This means that the risk of acute respiratory infection among children of smoking parents is nearly two times the risk among children of nonsmoking parents.

Benefits of case control study:

1- Suitable :

- to test the hypothesis that the disease of interest is caused by an exposure.
- for diseases with long latency period.
- to study rare diseases

2- Easy, rapid, & cheap (compared withy prospective cohort)

3- Requires few subjects.

4-Can examine multiple exposure factors for a single disease.

5-Estimation of the risk (odds Ratio)

6-Minimal ethical problems.

7- No attrition problem. Not costly

Limitations of case control study:

1- Incidence & Prevalence rates can not be calculated.

2- Not suitable for studying rare exposures.

3-The problem of bias.

Sources of controls in case control studies

In **case control** studies, the main sources are:

1. The total population in a given area, on the assumption that we know the extent of exposure in the general population.

2. Relatives and neighbors. This is useful to control for genetics and immediate environment.

3. Hospital patients other than those with the disease under study. Cases with a disease which may be related to the risk factor under study must not be used as controls.

4. Associates of cases in place of residence, schools, place of work.


Thank You!