
MEDICAL RESEARCH
MODIFIED NO. 12
كتابة: Doctor 021
تدقيق: ميس مصطفى, محمد مجيد
الدكتور: منير أبو هلالة

Week 6
Analytical Studies Part 2
Case-control Studies

Dr Munir Abu-Helalah

Color code

- Slides
- Doctor
- Additional info
- Important

Note: This Modified was written by Doctor 2021, but it has been paraphrased for improved wording. Best of luck !!

Case-control studies

As discussed in the previous lecture, the classical approach to conducting an analytical study is through a cohort study. However, a major limitation of cohort studies is their ineffectiveness in studying rare diseases. For this reason, case-control studies are better suited for investigating rare diseases.

Are studies in which a group of people with a particular disease (the cases) are compared with a group of people without the disease (the controls). The purpose of the comparison is to determine whether, in the past, the cases have been exposed more (or less) often to a specific factor than the controls.

We begin by identifying cases and selecting matched control patients. For instance, cases could be newborns with a specific congenital heart disease, while controls could be healthy newborns born during the same time period and at the same hospital.

In case-control studies, we begin with the disease, whereas in cohort studies, we start with the risk factor.

The goal of the comparison is to assess whether the cases were more or less frequently exposed to a specific factor in the past compared to the controls.

The classical way to study several risk factors, especially for rare diseases, is through case-control studies.

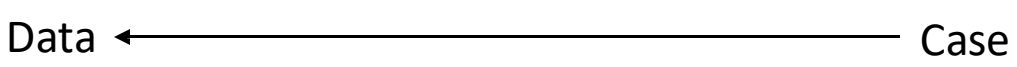
Reminder: If you are studying a rare exposure, a cohort study is required.

- This type of study is done to identify factors that could be responsible for the development of a disease or drug use problem.

CASE-CONTROL STUDIES

For example, if you have 200 lung cancer patients and 600 controls matched by age and gender, and you're investigating childhood exposure to radiation therapy, it's possible that none of the lung cancer patients had childhood malignancy. In cases of rare factors like this, it's better to start with a cohort study.

- The direction of time
- Cases identified now
  - Accordingly, in case-control studies, you can assess factors that no longer exist.
  - They are retrospective studies.
- Data on past events collected



CASE-CONTROL STUDY DESIGN

- Designed to assess association between disease occurrence and exposures (e.g., causative agents, risk factors) suspected of causing or preventing the disease.

Case-control studies

- **A group of people with a disease are compared to a group without the disease from the same population.**
- **Compare exposure to risk factors in both groups**
- **Able to look at many different possible risk factors**
- **Able to study diseases with a long latency period**

Some diseases, like most occupational cancers, have long latency periods. For example, exposure to certain environmental factors during the late teens or early 20s may lead to disease development in the 50s or 60s. When conducting a case-control study, we can gather a complete environmental and occupational history. For instance, we could ask lung cancer patients about their exposure starting at age 18, then calculate the odds ratio to evaluate these risk factors.

- **Most common analytic study design seen in the medical literature today**

Case-control studies

- **In general, the cases included in a case-control study include people with one specific disease only**
- **But, a case-control study can provide information on a wide range of possible exposures that could be associated with that particular disease**
- **Useful for the study of rare diseases**
- **Not suitable for the study of rare exposure**
- **Relatively small and inexpensive**
- **Takes a relatively short time to complete**
- **Can test current hypotheses**
- **Cannot measure disease incidence**

You might have 20 cases of congenital heart disease to study, but you still won't know how many additional cases exist in other hospitals.

CASE-CONTROL STUDIES

- Cases have the disease of interest
  Eg. Cerebral palsy

- Controls do not have the disease

Eg. Healthy babies born at the same time

Case-control study: challenges

- Selecting cases
  - Eligibility
- Selecting controls
  - Representativeness (The controls should match the characteristics of the patients, with the only difference being that they do not have the disease.)
- Exposure assessment
  - Accurate (If I'm investigating the effect of high aspirin intake during the first trimester of pregnancy as a risk factor for congenital heart disease, I need accurate information about the dosage of aspirin taken during that period.)

CASE-CONTROL STUDY DESIGN

- More efficient than a cohort study because a smaller sample size is required.
- One key feature of a case-control study, which distinguishes it from a cohort study, is the selection of subjects based on disease status.
- Controls are chosen from the same population yielding the cases

Design of case control studies

- **Comparability: Two groups must be as similar to each other as possible so selection of controls is very important. Controls must be as similar as possible to cases – except that they do not have the outcome (disease).**

- **Outcome (disease) must be very clearly defined. (Diagnostic criteria must be clear)**

- **Use objective data about exposure status wherever possible, to reduce the risk of bias** (You need to refer to the medical records to gather precise information about the aspirin dose the mother took during her first trimester—don't rely on the patient's memory.)

CASE-CONTROL STUDIES

Strengths

- Suited to study disease with long latency periods, but can be used in outbreaks investigations
- Optimal for rare diseases
- Efficient in terms of time and costs: relatively quick and inexpensive
- Allows for evaluation of a wide range of possible causative factors that might relate to the disease being studied
- Odds ratio estimated

CASE-CONTROL STUDIES

Limitations

- Very susceptible to bias (especially selection - how you select your cases, controls, and define your cases- and recall bias- sometimes, patients may not recall certain information accurately-) as both the disease and the exposure have already occurred when participants enter the study. Cases and controls might not be representative of the whole population
- We cannot calculate incidence or prevalence rate of disease
- We cannot be certain that exposure came before disease
- Choice of controls difficult
- Controls do not usually represent non-exposed populations (Sometimes, we prefer to have controls from the same hospital, while in other cases, we may select controls from the general population who match the cases in terms of age and gender.)
- Past records incomplete
- No absolute risk estimates

We also have what is known as interviewer bias. To explain this simply, imagine someone conducting interviews for a study on congenital heart disease. The interviewer has two groups: cases and controls. When interviewing the mothers of babies with congenital heart disease, they take time to conduct a detailed interview. However, when interviewing the mothers of healthy babies, they may rush through the questions, skip some, or not thoroughly review the records, simply because they know the babies are healthy.

- How can we avoid interviewer bias? By blinding the interviewer so they do not know which group each participant belongs to.

CASE-CONTROL STUDY DESIGN

- **Data Analysis**
- Data collection and analysis are based on whether the case-control study involves a matched or unmatched design. The measure used typically in case-control studies is the odds ratio.
- **Odds ratio (OR):** odds of a particular exposure among people with a specific condition divided by the corresponding odds of exposure among people without the condition under study

Odds Ratio

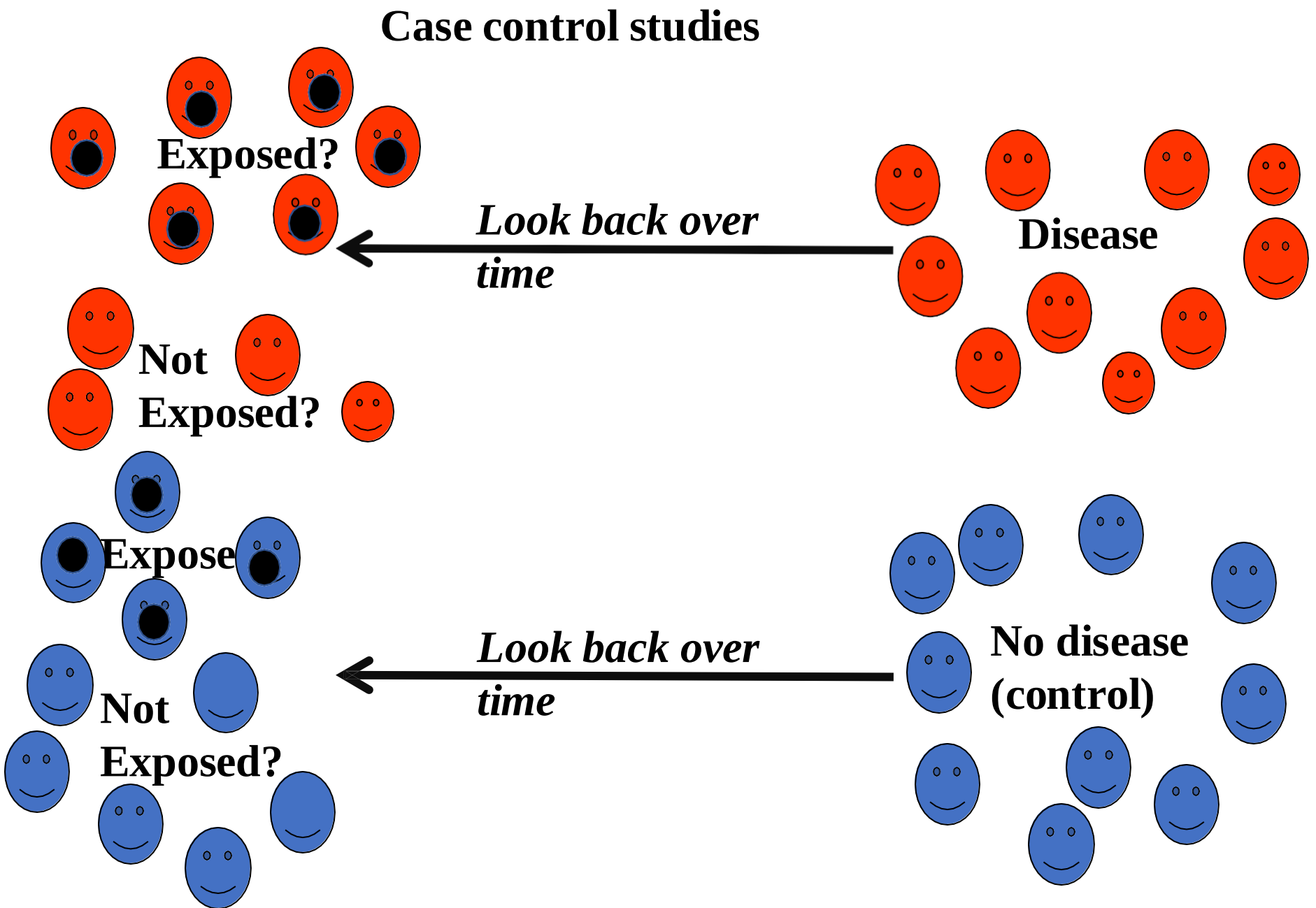
The word "odds" means the chances of an event to happen. The Odds of an event is the *ratio* of the event to happen over the event not to happen.

$$Odds(A) = \frac{probability(A\ happens)}{probability(A\ does\ not\ happen)} = \frac{prob(A)}{1 - prob(A)}$$

$$prob(A) = \frac{Odds(A)}{1 + Odds(A)}$$

Odds Ratio (OR)

$$OR = \frac{\text{Odds of exposure}_{\text{cases}}}{\text{Odds of exposure}_{\text{controls}}}$$


Case control studies
Exposed?
Look back over time
Disease
Not Exposed?
Expose
Look back over time
No disease (control)
Not Exposed?

Case-control study

	Disease Present	Disease absent	
Exposure Present	a	b	a+b
Exposure absent	c	d	c+d
Total	a+c	b+d	a+b+c+d

Odds of being ill in exposed=a/b

Odds of being ill in non exposed =c/d

Odds ratio (OR)=Odds in exposed/Odds in non exposed

= OR=(a/b)/(c/d)

$$Odds\ Ratio(OR) = \frac{ad}{cb}$$

Data from a case-control study of current oral contraceptive (OC) use and myocardial infarction in premenopausal female nurses

	Myocardial infarction		
	Yes	No	Total
Current OC use			
Yes	23	304	327
No	133	2816	2949
Total	156	3120	3276

Data from L. Rosenberg et al., Oral contraceptive use in relation to non-fatal myocardial infarction. *Am. J. Epidemiol.* 111:59, 1980.

$$OR = \frac{23 \times 2816}{304 \times 133} = 1.6$$

Women who were current OC users had a risk of MI 1.6 times that of nonusers

Two by two table

Exposure	**Outcome**		
	Yes	No	Total
Yes	a	b	a + b
No	c	d	c + d
Total	a + c	b + d	a + b + c + d

Odds of outcome in exposed = a / b

Odds of outcome in non- exposed = c / d

Outcome odds ratio = (a/b) / (c/d) = ad / bc

Case-control study

Early life exposure to diagnostic radiation and ultrasound scans and risk of childhood cancer: case-control study

BMJ 2011;342:d472

Objective To examine childhood cancer risks associated with exposure to diagnostic radiation and ultrasound scans in utero and in early infancy (age 0-100 days).
Design Case-control study.
Setting England and Wales.
Participants 2690 childhood cancer cases and 4858 age, sex, and region matched controls from the United Kingdom Childhood Cancer Study (UKCCS), born 1976-96.
Main outcome measures Risk of all childhood cancer, leukaemia, lymphoma, and central nervous system tumours, measured by odds ratios.

Case-control study: example

The table below presents the number of participants in a case-control study examining the link between early life exposure to diagnostic radiation and ultrasound scans, and the risk of childhood cancer

Radiation	**Case**	**Control**	Total
Yes	140	165	305
No	1550	5693	7243
Total	1690	5858	7548

Odds of outcome in exposed = 140 / 165 = 0.85

Odds of outcome in non-exposed = 1550 / 5693 = 0.27

Outcome odds ratio = (a/b) / (c/d) = 0.85/0.27=3.1

This means that children exposed to radiation have three times the risk of developing cancer compared to those who were not exposed.

CASE-CONTROL STUDIES

Methods of data collection

Case-note review: Completeness

Postal questionnaire: response rate

Interview: Detailed information

How many controls?

- **control-to-case ratio is 1 : 1**
  is the optimal when the number of available cases and controls is large and the cost of obtaining information from both groups is comparable
- **control-to-case ratio is 1 : n**
  When the number of cases is limited or when the cost of obtaining information is greater for cases or controls

- **As the number of controls per case increases, the power of the study also increase**

- **It is not recommended that this ratio increase beyond 4 : 1**

CASE-CONTROL STUDY DESIGN

- **Selecting Cases and Controls**
- Identification and collection of cases involves specifying the criteria for defining a person as a case—in other words, as having the disease (also called *case definition*).
- This definition consists of a set of criteria, also called *eligibility criteria,* for inclusion in the study. There also are criteria for exclusion from the study.

CASE-CONTROL STUDY DESIGN

- The next step is selection of the controls.
- Controls are chosen from the source population.
- The source population is usually defined by geographic area. It is important to select controls so that participation does not depend on exposure.

CASE-CONTROL STUDY DESIGN

Source of controls

- The ideal situation is a random sample from the same source population as the cases.
- Investigators may use more than one control group.
- Controls can be selected by sampling:

  The general population in the same community; the hospital community (patients in the same hospital); individuals who reside in the same block or neighborhood; and spouses, siblings, or associates (schoolmates, co-workers) of the cases.

Obtaining cases and controls for case control studies

Study	Source of cases	Source of controls
PROM (premature rupture of membrane)	Hospital patients	Hospital patients
Rheumatoid arthritis	Outpatient clinic	Other outpatient clinic
Cervical screening	GP register	GP register

CASE-CONTROL STUDY DESIGN

Matching Cases and Controls

- **Matching is a popular approach to control for confounding and selection bias in case-control studies.**
- **Matching cases and controls helps to ensure that these groups are similar with respect to important risk factors, thereby making case-control comparisons less subject to confounding or selection bias.**

CASE-CONTROL STUDY DESIGN

Prior exposure to the risk factor(s) of interest

- Once cases and controls are selected, information must be collected on prior exposure to the risk factor(s) of interest.

- Interviews and questionnaires are the most common means of determining a subject's exposure history and medical records review is another source
- The most objective means for characterizing exposure is the use of a biological marker.

Bias

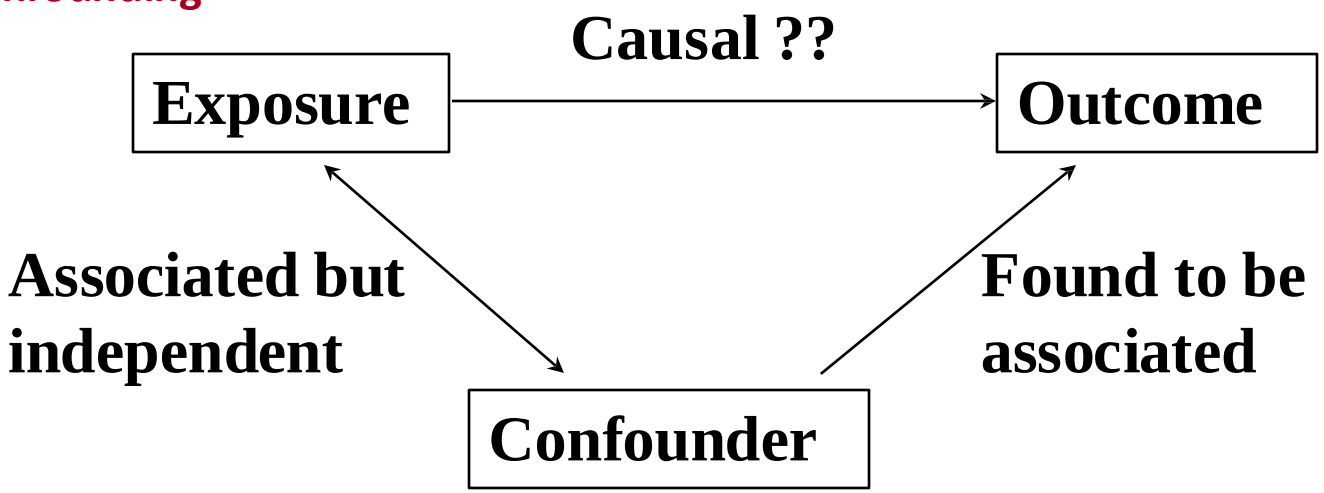
Bias is any systematic error in an epidemiological study that results in an incorrect estimate of the association between exposure and risk of the outcome

- **Selection bias:** inappropriate controls
- **Observation bias**
  - **Subject and recall bias:** eg recall bias of mothers with cerebral palsy babies
  - **Interviewer bias: blind if possible**
  - **Misclassification** (You should consider case definition and eligibility criteria to avoid misclassification).

Confounding

A confounding factor is one that is associated with the exposure and that independently affects the risk of developing the outcome, but that is not an intermediate link in the causal chain between the exposure and the outcome under study

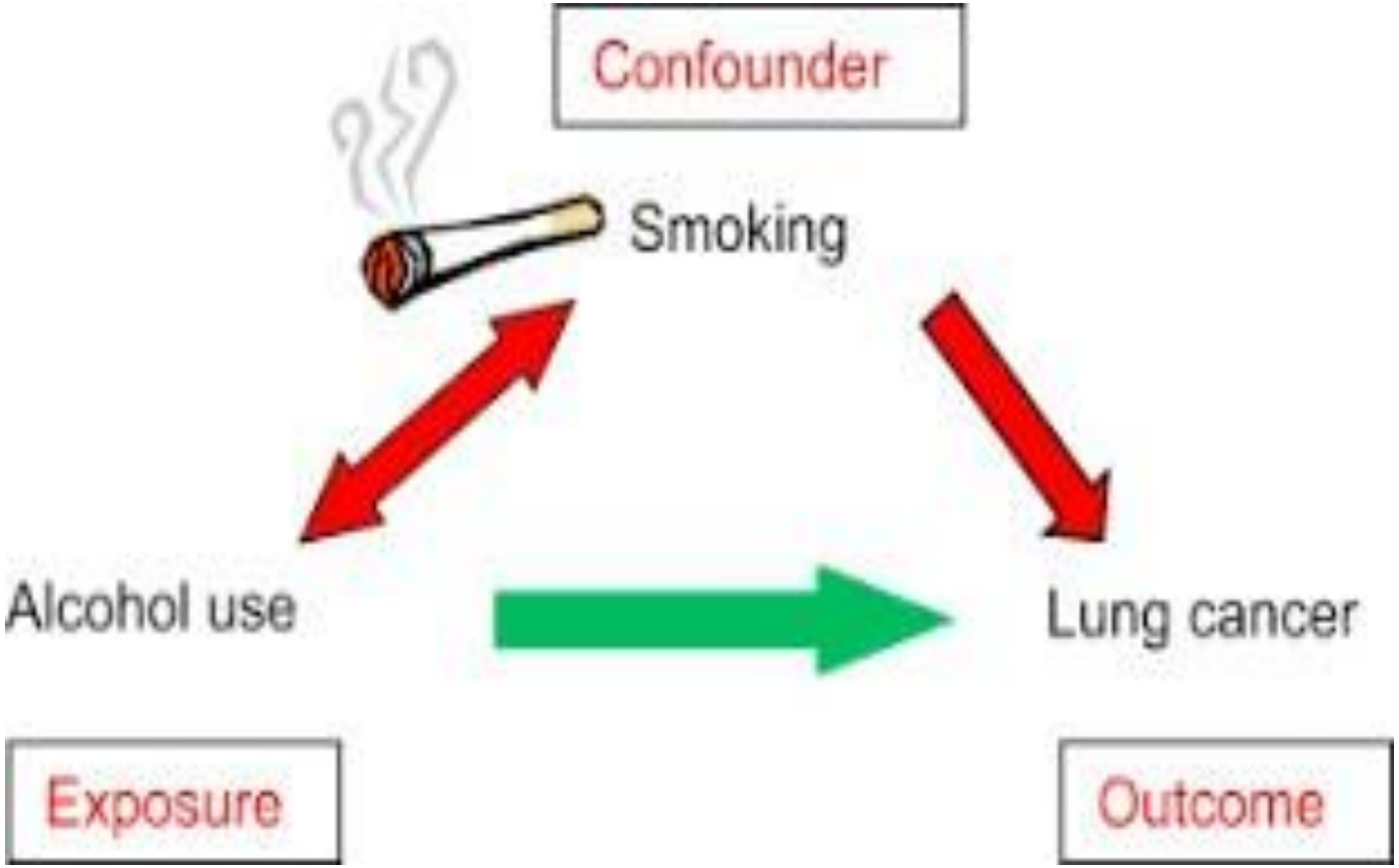
Matching - often used in case-control studies to decrease confounding



Let’s say a case-control study finds that the odds ratio for lung cancer and heavy alcohol consumption is 10. This means that heavy alcohol drinkers are 10 times more likely to develop lung cancer compared to those who don’t drink. While alcohol appears to be a risk factor here, it may actually be a confounding factor.

If we consider smoking, we can determine whether alcohol or smoking is the true risk factor. We would divide the heavy alcohol drinkers into two groups: heavy alcohol drinkers who smoke and heavy alcohol drinkers who don’t smoke.

In the group of heavy alcohol drinkers who don’t smoke, the odds ratio hypothetically drops to one, which is not significant. However, in the group of heavy alcohol drinkers who do smoke, the odds ratio increases. This suggests that smoking, not alcohol, is the actual risk factor for lung cancer.



P.S.: this is not the real situation. The doctor just wants you to understand the concept. If we evaluate the association between alcohol and lung cancer and discover that smoking also plays a role, we might conclude that smoking is the confounder.

Confounding

Matching Cases and Controls

- For example, if age and sex are the matching variables, then a 35 year old male case is matched to a 35 year old male control
  - Pair matching (one to one individual matching)
- The use of matching usually requires special analysis techniques (e.g. matched pair analyses and conditional logistic regression)

اللَّهُمَّ عَامًا كَعَامِ الْفِيل اللَّهُمَّ
طَيرًا أَبَابِيلَ تَرمِيهِم بِحِجَارَةٍ مِن سِجِّيل
اللَّهُمَّ تَدَخُّلًا رَبَّانِيًّا وَمُعجِزَةً مِن عِنْدِكَ
تَجعَلهُم كَعَصفٍ مَأْكُول
فَقَط ضَاقَت بِأَهلِ غَزَّةَ السُّبُلَ
وَقَلت بِأَيدِيهِم الْحِيَلَ اللَّهُمَّ نَصرًا
مُؤَزَّرًا وَرَحمَةً تَتغَمد بِهَا
شُهَدَائهُم الْأَبرَار

VERSIONS	SLIDE #	BEFORE CORRECTION	AFTER CORRECTION
V1→ V2			
V2→V3			

امسح الرمز و شاركنا بأفكارك لتحسين أدائنا !!