



Bias and confounding

Prof. Dr. Rabaa Mahmoud

Faculty of Dentistry, Mutah University

Credit to: Prof. Dr. Waqar Al-Kubaisy & Dr. Israa Al-Rawashdeh



Learning objectives

► **By the end of this lecture, the students will be able to:**

1. Discriminate between different types of bias and confounders.
2. Identify the drawbacks of introducing bias and confounders in the research.
3. Understand methods to overcome bias and confounders.

What does bias mean?



systemic error

Bias occurs when there is a systematic difference between the results from a study and the true situation.



Bias may be introduced at all stages of the research process, from study design, to analysis and publication.



Between exposure and outcome.

Bias can create a fake association (i.e., overestimation or underestimation of an effect).



Types of bias

➤ **There are many forms of bias which may be categorized as:**

A. Selection bias.

B. Information bias.

C. Funding bias.

D. Publication bias.

E. Confounding bias.

A. Selection bias

➤ Selection bias occurs when patients included in the study are not representative ^{بغير تمثيل} of the population to which the results will be applied,

➤ *For example:*

Patients who agree to participate in a study may differ from those who do not agree to participate (this form of bias is a particular problem in retrospective studies, where patients who have died are excluded from the study).

→ 4 types

Selection bias includes the following:

(ما أخترته بطريقة عشوائية)

➤ **Ascertainment bias** may occur when the sample included in a study is not randomly selected from the population.

➤ **For example:**

When doctors are interested in the genetics of a particular medical condition, they collect information on the patients in their clinic, rather than using a random sample from the population.

loss of follow up

➤ **Attrition bias** arises when those who are lost to follow-up in a longitudinal study differ from those who are not lost to follow-up.

ينسحب بعض المشاركين من الدراسة بشكل غير متساوي

أو الانسحاب يؤثر على النتيجة

➤ **Response bias** is caused by differences in characteristics between those who choose or volunteer to participate in a study and those who do not.

➤ **Survivorship bias** occurs when a study includes only those subjects who have "survived" long enough to receive **certain intervention**. → as drug

➤ This can lead to systematically distorted results, because those who died were unable to receive the intervention.

توضيح الفكرة
بالمثال تحت

لازم أبدأ أخذ
النتائج من 2015

For example,

- Imagine a long-term cohort of cancer patients being studied from 2010–2020.
- A new targeted drug becomes available in **2015**.
- Researchers compare survival between patients who received the new drug and those who did not.
- Patients who lived until 2015 could receive the drug, but patients who died before 2015 obviously could not.
- If survival is measured from 2010, it will look like the drug-treated group lives longer, but much of that “extra survival” occurred **before** they even received the drug.

* *The patients who died Before 2015 must be not included*

B. Information Bias

- Arises when there are systematic errors in measuring exposure or outcome.

For example

يعني ما يذكروا انهم يرضعوا مثلاً

- Participants may underreport behaviors such as smoking or alcohol use (recall bias), or medical records may contain incorrect diagnoses.

تشويه

- Information bias can distort the association between exposure and outcome, leading to either **overestimation** or **underestimation** of true effects.

Information bias includes the following:

I. Central Tendency Bias

لا مثلاً صاحب مطعم به يقمين رضا الزبائن عن طعامه
فعادة الزبائن تميل لا اختيار الخيار الوسط

- Arises when using a Likert scale (comprising a small number of graded alternative responses such as very poor, poor, no opinion, good, excellent), where responders tend to move towards the mid-point of the scale (usually ‘no opinion’ or ‘just right’).
- As a result, data become clustered around the center of the scale.
- This bias can distort study findings, especially in assessments of satisfaction, performance, or attitudes.

II. Measurement Bias →

الخطأ في قياس الـ Data
مثل أنه يكون الجهاز مش مشغول

- Measurement bias occurs when the instruments or procedures used to collect data distort the results in one direction.

For example

- Using a ^{معايرة خاطئة} miscalibrated blood pressure cuff that consistently reads higher values can lead to an overestimation of hypertension prevalence.

III. Misclassification Bias

- Arises when individuals are **incorrectly categorized** with respect to their exposure or outcome variables.

For example:

- A person **with a disease** is mistakenly recorded as **disease-free**.
- A person **exposed** to a risk factor is recorded as **unexposed**.

Causes:

1. **Inaccurate measurement tools** (faulty test, poor questionnaire design)
2. **Recall errors by participants** → المریض ما يتذكر History
3. **Data entry errors**
4. **Observer bias during assessment** → صحت الـ data بطريقة متحيزة

IV. Observer Bias

الباحث متحيز لفكرة وطول ما هو شغال في الدراسة يحاول يثبت فكرته

- Occurs when one observer tends to under-report (or over-report) a particular variable; also called **assessment bias**.

V. Reporting Bias

→ Patient يمشي بنفس الاتجاه الي حسه من observer به اياه

- Occurs when participants give answers in the direction they perceive is of interest to the researcher or under-report socially unacceptable or embarrassing behaviors (e.g., alcohol consumption or sexually transmitted diseases).

C. Funding Bias

المؤسسة الساعمة

- Funding bias occurs when the *source of financial support* for a study influences its design, methods, analysis, or reporting in a way that favors the sponsor's interests.

What does it mean?

When a company, organization, or group funds research, there is a risk, intentional or unintentional, that:

- The study is designed to produce favorable outcomes.
- Unfavorable results are suppressed or not published.
- Data analysis is interpreted in a sponsor-friendly way.
- Only certain comparisons are chosen to increase the chance of positive findings.

D. Publication Bias

- ▶ **Publication bias** refers to the tendency for studies with favorable results to be published more often and more quickly than those with unfavorable results.

Why does it happen?

Publication bias occurs because:

- Journals prefer positive findings.
- Researchers may not submit negative results (“file-drawer effect”).
- Sponsors may discourage publication of unfavorable data.
- Peer reviewers may view null results as less valuable.

النتائج السلبية

E. Confounding: What is it?

- ➔ **Confounding** occurs when we find a FAKE association between a potential risk factor and a disease outcome or miss a real association between them because we have failed to adjust for any confounding variables.

- ➔ **A confounding variable** or confounder is associated with both the exposure and the outcome.

الذي يسبب *fake association*

الطرف الثالث

For Example,

- ▶ We may be interested in studying the effect of **smoking** status on the incidence of coronary heart disease (CHD) in men.
- ▶ However, we know that **alcohol** consumption is associated with the development of **CHD**, and that alcohol consumption and smoking are also related to each other (i.e., men who consume alcohol are more likely to smoke than men who do not drink alcohol).
- ▶ Thus, in this study, unless we adjust for it, the effect of alcohol consumption may confound an apparent relationship between smoking and the incidence of CHD.



Methods to Control Bias and Confounders

1. During Study Design
2. During Data Collection
3. During Data Analysis

1. During Study Design

➤ **Randomization:** → *avoided selection Bias*

- Randomly assign participants to exposure or treatment groups.

➤ **Restriction:** → *avoided confounding Bias.*

- Limit study participants to a specific category of a confounding variable.

* يعني مثلاً نختار بس المدخنين بس ما يشربوا كحول

- Example: include only non-smokers or only females.

➤ **Matching:**

- Select control groups with similar characteristics (e.g., age, sex) as the case groups.

- Commonly used in case–control studies.

* يعني لما نختار مجموعة controls ومجموعة cases الاثنيتين يكونوا

يحملون نفس الصفات لالاختلاف الوحيد مجموعة معها المرض ومجموعة لا



■ **Blinding (Masking):** → محاضرة د. فانت

- Keep participants, investigators, biostatisticians, and sponsors unaware of exposure or treatment status.

■ **Standardization of Procedures:**

- Use uniform instruments, questionnaires, and protocols for all participants.
- This reduces measurement variability and interviewer bias.

2. During Data Collection

► Use of validated tools:

* يعني مثلًا الأسئلة التي بدنا نسألها للمرضى
نعرضها على خبراء ليقبوا صحة الأسئلة ومدى مناسبة

► Employ standardized, pretested instruments to ensure accuracy.

► Training and calibration:

► Train and calibrate examiners to maintain consistency.

► Reducing recall bias:

► Use objective data sources (e.g., medical records) or shorten recall periods.

► Blinded data collection:

► Ensure data collectors do not know participants' exposure or outcome status.

3. During Data Analysis

► Stratification:

- Analyze results within subgroups of a confounding variable (e.g., by age or sex).

Female alone , male alone * بي اقيس الضغط نجيب ال

انه عوامل الضغط تختلف Between male and female

► Multivariable analysis:

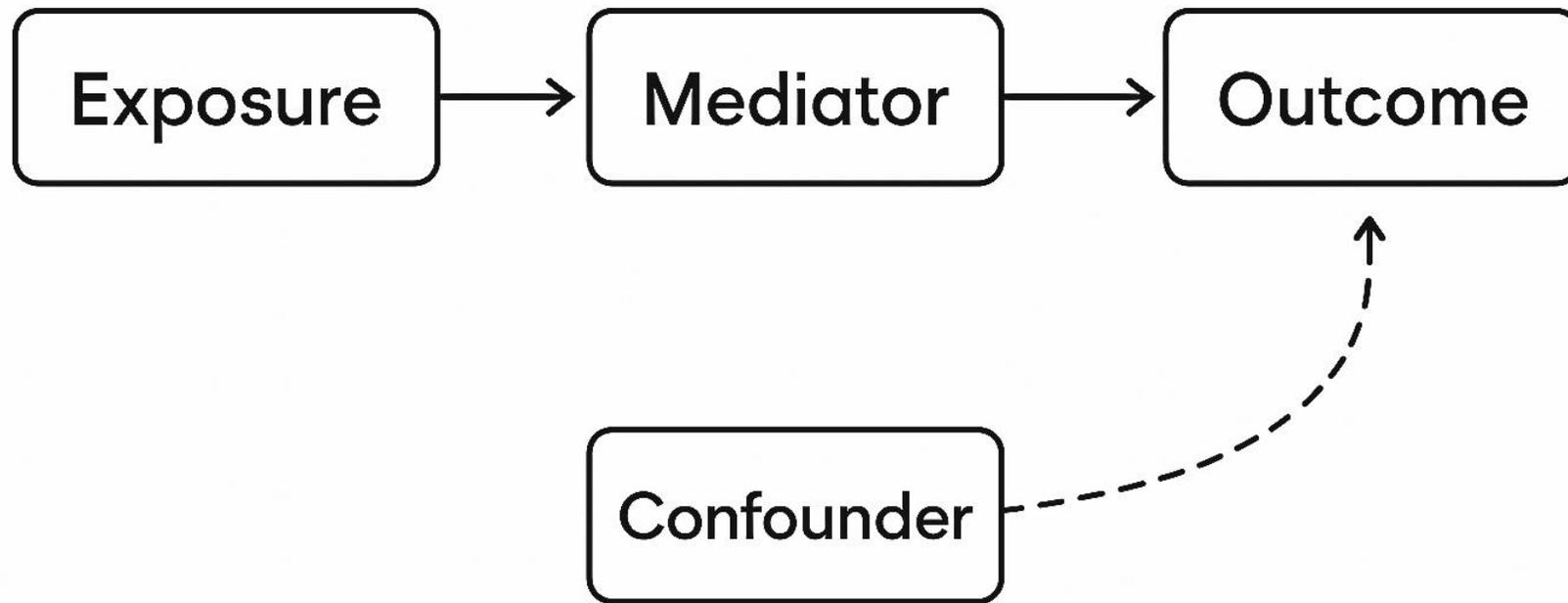
- Use statistical models such as logistic or linear regression to adjust for several confounders simultaneously.

* عوامل كثيرة بي تا نشوفه مينه أكثر
ارتباط في disease هذا

The causal pathway and confounding

- The **causal pathway** is the chain of events that, in sequence, lead to an outcome.
- The causal pathway helps us **consider opportunities for disease prevention**.
- **We should not confuse mediators**, which are part of the causal chain and confounders, which are not part of it.
- **For Example: multiple births → preterm delivery → neonatal cerebral damage in cerebral palsy**
another Example → exposure to sucrose → Bacteria fermentation sucrose to acid → acid damage tooth.
- We should **not adjust** for variables that lie **on the causal pathway** between exposure and outcome.

The causal pathway and confounding



Thanks ➔

