

6-Part Webinar Series: Research Methodology

Part 2: Bias in Epidemiological Studies

Questions Asked: February 15, 2024

1. What is the difference between fraud or falsification and bias?

Research misconduct includes fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results.

Fabrication is making up data or results and recording or reporting such. Falsification is manipulating research materials, equipment, or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.

Bias results from any deviation from the truth, causing distorted results and wrong conclusions. Bias can occur at any phase of your research, including during data collection, data analysis, interpretation, or publication. Research bias can occur in both qualitative and quantitative research.

Research misconduct does not include honest error or differences of opinion.

2. Why is random error not biased?

Random error considers the variability, random variation, or 'noise in the system'. The level of heterogeneity in the population contributes to relatively large random variation in studies.

Systematic error is bias referring to deviations that are not due to chance alone.

3. Would a nested case-control study be a good approach to reduce selection bias?

In a nested case-control study, the cases and controls were selected from the database and therefore should be more representative of the population than those in a traditional case-control study. Hence, selection bias was minimized by using the nested case-control study design.

4. How might biases introduced by lost to follow-up impact the interpretation of study results?

After enrollment of subjects and collection of baseline data there is usually some loss to follow-up. Thus, when individuals leave the study before the end of follow-up. This biases the study when the association between a risk factor and a health outcome differs in dropouts compared with study participants.

5. Should we avoid informing the participants about the outcome of a cohort study?

Ethical guidelines indicate participants are often to be fully informed as to the purpose of the study and given an explanation of findings in the study.

6. How do you minimize bias due to loss of follow-up?

Some strategies to address Loss to Follow-up include:

- Enrolling motivated subjects.
- Using subjects who are easy to track.
- Making questionnaires as easy to complete as possible.
- Maintaining the interest of participants and making them feel that the study is important.
- Providing incentives.

7. In terms of non-response bias, how do we know how this will affect our research?

Assess the non-responders.

8. How do you care for missing data in cohort case studies, especially when you have a loss to follow up?

Missing data clearly leads to a loss of information and hence reduced statistical power. Another consequence is that this lack of data may introduce selection bias, which could potentially invalidate the entire study.

9. What is the best strategy for characterizing the controls in a case-control study to minimize selection bias?

A close match to cases.

10. If I'm doing a case-control study and my controls are hospitalized patients not suffering from the exposure under investigation, will I have data with Berkson's bias?

Berkson's bias is a form of selection bias that causes hospital cases and controls in a case control study to be systematically different from one another because the combination of exposure to risk and occurrence of disease increases the likelihood of being admitted to the hospital.

11. How can we minimize diagnostic bias when we undertake Hospital-based case-control studies?

The practice of reflection (the ability to reflect on one's actions so as to engage in a process of continuous learning) reinforces behaviors that reduce bias in complex situations. Simply increasing physicians' familiarity with the many types of cognitive biases—and how to avoid them—may be one of the best strategies to decrease bias-related errors.

12. How many controls should researchers take per case? What is the evidence regarding this?

if it is time-consuming or expensive to collect data on controls, the ratio of controls to cases should be no more than 4:1.

13. I have cases (children) who have brain damage due to TB; I have controls with minor injuries such as fractures/lacerations/amputations. We want a baseline MRI. They are matched by age, sex and socio-economic status. Are these good controls?

Age, race and SES seem good matches for controls.

14. Apart from matching, what are other key considerations to avoid selection bias in a case-control study?

Selection of controls in as similar manner as cases as possible.

15. How do we mitigate recall bias when studying longevity in the elderly group?

Several considerations to minimize recall bias:

- **Be mindful of questions when seeking information about past events because even the slightest alteration can trigger a biased response.**
- **Try to avoid leading questions altogether and stick to open-ended ones that allow for a more genuine recollection.**
- **Use prompts like photos or videos to trigger memories. These visual aids help keep things fresh and accurate.**
- **Create an environment that's conducive to honest recall. Reduce distractions and ensure privacy so people feel comfortable sharing their true memories without external influences.**

16. Can we have studies with no bias?

Randomized control trials at the top of the evidence pyramid do this best.

17. Is it even possible to address or anticipate all types of bias in research?

While an extra burden – addressing bias as a limitation in nearly all studies is very important.

18. Is information bias under systematic bias?

Information bias is any systematic difference from the truth that arises in the collection, recall, recording and handling of information in a study, including how missing data is dealt with. Major types of information bias are misclassification bias, observer bias, recall bias and reporting bias.

19. What is Silver Blaze bias? Kindly explain.

In Arthur Conan Doyle's story on the disappearance of the champion racehorse Silver Blaze, the detective character Sherlock Homes homed in on his suspect by reasoning that whoever was involved in the disappearance of Silver Blaze that night must have been very well known at the stable because the stable dog never barked during the night. This key finding in solving the mystery exemplifies the general principle that a lack of response – and not just a response – may reveal its own inimitable truths. This principle is highly relevant to biomedical research. An example - investigators conducted a survey of patients with the first-ever code for heart failure and who resided in southeast Minnesota. They assessed the characteristics and the outcomes of those surveyed through the medical record database of the Rochester Epidemiology Project. In this survey involving approximately 8000 patients, the response rate was 43%, and, on follow-up of some 1.5 years after the survey was conducted, there was a remarkable divergence in outcomes between nonparticipants and participants in the survey: nonparticipants exhibited a two-fold increased risk of death and an increased rate of hospitalization after adjustment for a number of relevant factors. Nonparticipants differed from participants in several ways in that the nonparticipant group comprised a higher percentage of individuals who were female, single, nonwhite, less well educated, resident in rural areas, and with diagnoses of mental and psychiatric disorders; in contrast, nonparticipants were less likely to be diagnosed with cancer and assorted cardiovascular diseases. The findings have several major and far-ranging implications. First, they speak to the issue of participation bias as a limitation that may weaken the validity of epidemiologic studies; specifically, survey did not capture information from individuals with heart failure with significantly different health outcomes. Second, as pointed out by the authors, the current literature describes an association of a lack of participation in surveys with certain characteristics, several of which (lower levels of education, being single, among others) were also observed in their study. A lack of participation in surveys by patients may also reflect their lower sense of connectivity with their health care system or their greater sense of the burden of their disease and treatment. Third, health care delivery increasingly emphasizes patient-centric care, the latter enabled by patient reported outcomes; a lack of response in reporting.

20. Could you please differentiate between type 1 and type 2 errors with examples?

A type I error occurs when in research when reject the null hypothesis and erroneously state that the study found significant differences when there indeed was no difference. In other words, it is equivalent to saying that the groups or variables differ when, in fact, they do not or having false positives.

A type II error is a statistical term used within the context of hypothesis testing that describes the error that occurs when one fails to reject a null hypothesis that is actually false. A type II error produces a false negative, also known as an error of omission.

21. What is less harmful to the participants, a type 1 or a type 2 error, and what is the rationale behind that?

Both can be harmful – depends on the critical question. And the effects of false positives and false negatives.

22. Why is recall bias separated when that looks to be information bias?

A type of bias that occurs when participants in a research study or clinical trial do not accurately remember a past event or experience or leave out details when reporting about them. Recall bias is more likely to occur when the event happened a long time ago or when the study participant has a poor memory.

Information bias is a distortion in the measure of association caused by a lack of accurate measurements of key study variables. Information bias, also called measurement bias, arises when key study variables (exposure, health outcome, or confounders) are inaccurately measured or classified.

23. How do we get valid information and avoid recall bias? How can we reduce recall bias? How can we reduce recall bias?

Several considerations to minimize recall bias:

- Be mindful of questions when seeking information about past events because even the slightest alteration can trigger a biased response.**
- Try to avoid leading questions altogether and stick to open-ended ones that allow for a more genuine recollection.**
- Use prompts like photos or videos to trigger memories. These visual aids help keep things fresh and accurate.**

- **Create an environment that's conducive to honest recall. Reduce distractions and ensure privacy so people feel comfortable sharing their true memories without external influences.**

24. Are these biases having a poor impact on our analysis?

Biases, including Information bias, Interviewer bias, Publication bias, Researcher bias, Response bias, and Selection bias should all be considered in the analyses.

25. What is the difference between information bias and recall bias?

A type of bias that occurs when participants in a research study or clinical trial do not accurately remember a past event or experience or leave out details when reporting about them. Recall bias is more likely to occur when the event happened a long time ago or when the study participant has a poor memory.

Information bias is a distortion in the measure of association caused by a lack of accurate measurements of key study variables. Information bias, also called measurement bias, arises when key study variables (exposure, health outcome, or confounders) are inaccurately measured or classified.

26. Are recall and observer biases, not information biases?

Recall or reporting bias is another form of information bias due to differences in accuracy of recall between cases and non-cases or of differential reporting of a health outcome between exposed and unexposed.

Other types of information bias are misclassification bias, observer bias, recall bias and reporting bias.

27. Is there any difference between bias and confounding?

Bias refers to systematic error in how we measure or report data, while confounding refers to real but misleading associations. The ability to distinguish between biasing and confounding factors can be helpful in evaluating the true impact of a program or public health initiative on the desired outcome.

28. What gold standards were checked in the radiology study? (slide 50)

Consensus of radiologists on set of images.

29. How can we manage the refusal non-response?

Some strategies include:

- **Reconsider your survey's trigger and timing.**
- **Optimize your survey's design.**
- **Review your survey questions.**
- **Provide options for omission.**
- **Reward customers with incentives.**
- **Keep the survey length short.**
- **Make sure the information is confidential.**

30. Can you please explain social desirability bias?

Social-desirability bias is a type of response bias that is the tendency of survey respondents to answer questions in a manner that will be viewed favorably by others. For example, if a survey asks about sensitive topics like drug use or mental health, respondents might be reluctant to provide truthful answers for fear of judgment or repercussions.

31. How do we determine and decrease the sources of biases in a study?

It is usually very difficult or impossible to measure whether or not the results of a particular study has been affected by bias; so risk of bias is frequently assessed by looking for features of the study design and conduct of the study that have been shown by empirical evidence to minimize the risk.

So first, consider potential sources and types and bias and address in the study design.

32. What are your thoughts about "noise" and its relation to bias? The work of Daniel Kahneman.

Returning to systematic error vs random error discussions - To improve judgments, we need to reduce error as much as possible by correcting for noise and bias. The targets are the judgements where inconsistency = noise and inaccuracy = bias.

Daniel Kahneman is a Nobel prize-winning psychologist working in this area.

33. Can we consider long intervals between data collection and "onset of illness" as recall bias? For example, a person got sick, say, January last year, and I am interviewing the patient today about events surrounding their illness that started last year.

Indirectly – Recall bias is a type of bias that occurs when participants in a research study or clinical trial do not accurately remember a past event or experience or leave out details when reporting about them. Recall bias is more likely to occur when the event happened a long time ago or when the study participant has a poor memory. The impact is particularly great when study participants recall differently,

34. How are the P-value and confidence intervals related to bias?

P-value and confidence intervals can be considered in identifying potential impact of bias.

35. Does Neyman's Bias affect survival study, and how?

Neyman Bias is a selection bias where the very sick or very well (or both) are erroneously excluded from a study. The bias (“error”) in your results can be skewed in two directions: Excluding patients who have died will make conditions look less severe.

36. Which studies are more likely to present publicity bias?

Studies with high appeal to participate

37. If bias in the design/methodology of the study comes to the notice of the investigator during conduct - and the investigators mention it in the manuscript - how would that affect the chances of publication?

Describing potential bias in a manuscript is very important and is the responsibility of an author. Such information is critical for future studies.

38. Is a telephone survey different from using an online application like WhatsApp to conduct a study?

Similar if the online addresses are selected randomly

39. If the study's purpose is not to generalize the results and our selection method is biased (e.g., not everyone has the same probability of being involved in the study), will it be acceptable? Or, in this case, will the study be judged for biased random sampling?

While the population would be ‘select’ – if the approach is sound and rigorous, the results could provide value to this select population.

40. Does enlarging the studied groups reduce the impact of selection bias?

Bias skews the results, whereas random errors increase the variance but do not skew the results. In a random sample, larger sample size can help reduce the influence of random noise. But larger sample size usually does nothing to minimize the effect of bias.

41. When selecting control villages for a case-control study on measles outbreaks, should I directly pick the nearest ones to the cases, or do I need to choose them randomly? This means that one village might be selected more than once if it's close to multiple cases or villages with measles outbreaks.

The control village would wish to replicate the same village as closely as possible. If that is the closest – that would be the best.

42. In study design one, why weren't other admitted patients, other than those with arthritis, included as control? (slides 41-42)

This was a hypothetical example to demonstrate bias due to the selection of controls.

43. How much will cluster sampling affect the study as a proxy of selection bias? Or is cluster sampling a type of selection bias?

There is a higher risk of bias in cluster sampling compared to other methods of sampling such as simple random sampling or probability sampling methods. Individuals within each cluster may be more similar to each other than to individuals in other clusters.

44. Regarding selection bias, can I trade-off for a less serious bias if there is little option? Assuming trading of loss to follow up for conducting a study among health workers who could be easily followed up. Also, if yes, how could you make sure the trade-off is the best of decisions?

The proper approach would be to design the study to minimize bias. But then describe any potential bias. This is important as the follow-up studies can address in future designs.

45. How can one ensure that in trying to prevent volunteer bias, the volunteer group is also represented in the selection process? In other words, how can one ensure that the study population also represents the "volunteer group"?

Important question. The major driving determinant is the question being addressed. In general – your sampling scheme should identify a sample representative of all the

population including those who tend to volunteer. The big question is what to do with volunteers NOT selected in the sampling. In those situations in which desperate volunteers are recruited, it would be important to try to minimize the negative effects—for example, by using unequal randomization in favor of the experimental arm in the trial design.

Instead of randomly selecting participants, researchers can also invite individuals who willingly volunteer to participate in the study. It ensures that only those who are interested in the topic come to the forefront. It provides you with valuable data through methods like online surveys, interviews, or questionnaires.

46. Please share some key aspects of the validation of the study questionnaire.

Generally it involves running a study that is designed to determine a specific kind of validity, although it is sometimes possible to add a validation arm onto a trial with other primary objectives. One way to check the validity of a questionnaire is to compare its results with results from more objective measures.

General

- Thank you very much for the excellent webinar. Please introduce key references.
Thank you – we enjoy the Program and appreciate all participants.
- Please share the flier or a link to the website for the World Hypertension Congress so we can participate.
Will do with details.
- Is there a scholarship for attending the Congress?
Not at this time but we are working on such a scholarship.