MedicalBiostatistics.com

VARIETIES OF BIAS TO GUARD AGAINST

Medical research results often become clouded because some bias is detected after the results are available. Therefore, it is important that all sources of bias are considered at the time of planning a study, and all efforts are made to control them. Various sources of bias are as follows. These are not mutually exclusive sources. In fact the overlap is substantial. Some of the biases in this list in fact are collection of many biases of similar type. If we state all these separately, the list may become unmanageable.

1. **Bias in concepts** — Lack of clarity about the concepts that are to be used in the proposed research. This gives an opportunity to the investigators to use subjective interpretation that can vary from person to person. Sometimes the logic used can be faulty and sometimes the premise itself of the logic can be incorrect. For example, it is generally believed for body mass index and blood pressure that the lower the better. In fact their very low values are also associated with increased morbidity and mortality.

2. **Definition bias** — The study subjects should be sharply defined so that there is no room for ambiguity. For example, if the cases are of tuberculosis, specify that these would be sputum positive, Montoux positive, radiologically established, or some combination. Blurred definition gives room to the assessor to use subjective interpretation that can affect the validity of the study.

3. **Bias in design** — This bias occurs when the case group and control group are not properly matched, and the confounding factors are not properly accounted for at the time of analysis. Concato et al. (2001) reports another type of bias in designs for prostate cancer detection when groups were asymptomatic men who received digital rectal examination, screening by prostate specific antigen and transrectal ultrasound, but there was no 'control' group with 'no screening'. Thus the effectiveness of screening could not be evaluated.

4. **Bias in selection of subjects** — This occurs when the subjects included in the study are not truly representative of the target population. This can happen either because the sampling was not random, or because sample size is too small to represent the entire spectrum of subjects in the target population. Studies on volunteers always have this kind of bias. Selection bias can also occur because the serious cases have already died and are not available with the same frequency as the mild cases (survival bias). See also Length bias.

5. **Bias due to concomitant medication or concurrent disease —** Selected patients may suffer from other apparently unrelated condition but their response might differ either because of this condition itself or because of medication given concurrently for that condition.

6. **Instruction bias —** When unclear or no instructions are prepared, the investigators use discretion and this can vary from person to person, and from time to time.

7. **Length bias —** A case-control study is generally based on prevalent cases rather than incident cases. Prevalence is dominated by those who survive for a longer duration. And these patients are qualitatively different from those who die early. Thus the sample may include disproportionately more of those who are healthier and survive longer. The conclusions can not be generalised to those who have less survival time.

8. **Bias in detection of cases —** Error in the diagnostic or screening criteria, e.g., being able to use a laboratory investigation properly in the hospital setting but not in the field setting where the study is to be actually done. In a prostate cancer detection study if prostate biopsies were not performed in men with normal test results, true sensitivity and specificity of the test can not be determined.

9. **'Lead-time' bias —** All cases are not detected at the same stage of the disease. In cancers, some may be detected at the time of screening such as by pap smear, and some may be detected when the disease has started clinical manifestation. But the follow-up is generally from the time of detection. This difference in 'lead-time' can cause systematic error in the results.

10. **Bias due to confounder —** Failure to take proper care of the confounders so that any difference or association can not be fully ascribed to the antecedent factors under study.

11. **Contamination in controls —** Control subjects are generally those that receive placebo or the usual therapy. If these subjects are in their homes, it is difficult to know if they have received some therapy that can affect their status as a control. In the prostate cancer detection project reported by Concato et al. (2001) and discussed in preceding paragraphs, the controls subjects are those who are under usual care. But some of these may be screened outside the study and treated. Thus their survival rate would not be sufficiently 'pure' to be compared with the survival of those who were screened by the test procedures. In a field situation, contamination in control group occurs if it is in close proximity of the unblinded test group and learns from their experience. The neighbouring area may not be the test area of the research but some other program may be going on there that has spill-over effect on the control area.

12. **Berkson's bias —** Hospital cases when compared to hospital controls can have bias if the exposure increases the chance of admission. Thus cases in a hospital will have disproportionately higher number of subjects with that exposure. Cases of injury in motor vehicle accidents have this kind of bias.

13. **Bias in ascertainment or assessment —** Once the subjects are identified, it is possible that more care is exercised by the investigators for cases than for controls. This can also occur when subjects belonging to a particular social group have records but others have to depend on recall. Sometimes this is also called information bias.

14. **Interviewer bias or observer bias —** Interviewer bias occurs when one is able to elicit better response from one kind of patients (say, those who are educated) relative to the other kind (such as illiterates). Observer bias occurs when the observer unwittingly (or even intentionally) exercises more care about one type of responses or measurements such as those supporting a particular hypothesis than those opposing this hypothesis.

Observer bias can also occur if he is, for example, not fully alert in hearing Korotkoff sounds while measuring blood pressure or not being able to properly rotate endoscope to get an all round view of, say, duodenum in a suspected case of peptic ulcer.

15. **Instrument bias** — This occurs when the measuring instrument is not properly calibrated. A scale may be biased to give a higher reading than actual, or lower than actual. The other possibility is inadequacy of an instrument to provide complete picture such as endoscope not reaching to the site of interest and giving false information from a distance.

16. **Hawthorne effect** — If a subject knows that he is being observed or being investigated, his behaviour and response can change. In fact, this is the basis of including a placebo group in a trial. Usual responses of subjects are not the same as when under a scanner.

17. **Recall bias** — There are two types of recall bias. One, arising from better recall of recent events than those occurring long time ago. Also, serious episodes are easy to recall than the mild episodes. Two, cases suffering from disease are able to recall events much more easily than the controls if they are currently healthy subjects.

18. **Response bias** — Cases with serious illness are likely to give more correct responses regarding history and current ailments compared to the controls. Some patients such as those of STDs may intentionally suppress sexual history and other information because of stigma attached to these diseases. Injury history may be distorted to avoid legal consequences. If the subjects are able to exchange notes, the response to questions might alter, in some cases might even be uniform. An unsuspecting illness, death in the family, or any such drastic event may produce an extreme response. Response bias also comes under information bias.

19. **Repeat testing bias** — In a pretest-posttest situation, the subjects tend to remember some of the previous questions and they may remove previous errors in posttest—thus do better without the effect of the intervention. Observer may acquire expertise second or third time to elicit correct response. Conversely fatigue may set in repeat testing that could alter the response. It is widely believed that most biological measurements have strong tendency towards mean. Extremely high scorers tend to score lower in subsequent testing, and extremely low scorers tend to do better in a subsequent test.

20. **Mid-course bias** — Sometimes the subjects after enrolment have to be excluded if they develop an unrelated condition such as injury, or become so serious that their continuation in the trial is no longer in the interest of the patient. If a new facility such as a health centre is started or closed for the population being observed for a study, the response may alter. If two independent trials are going on in the same population, one may contaminate the other. An unexpected intervention such as an outbreak can alter the response of those who are not affected.

21. **Self-improvement effect** — Many diseases are self-limiting. Improvement over time occurs irrespective of the intervention, and it may be partially or fully unnecessarily ascribed to the intervention. Diseases such as arthritis and asthma have natural periods of remission that may look like the effect of therapy.

22. **Digit preference** — It is well known that almost all of us have special love for digits 0 and 5. Measurements are more frequently recorded ending with these digits. A person of age 69 or 71 is very likely to report his age 70 years. Another manifestation of digit preference is in forming intervals for quantitative data. Blood glucose level categories would be 70-79, 80-89, 90-99, etc., and not like 64-71, 72-79, etc. If digit zero is preferred, 88, 89, 90, 91 and 92 can be recorded as 90. Thus intervals such as 88-92, 93-97 and 98-

102, are better to ameliorate the effect of digit preference, and not the conventional 85-89, 90-94, 95-99, etc.

23. **Bias due to nonresponse** — Some subjects refuse to cooperate, injure, die, or become untraceable. In a prospective study, there might be some dropouts for various reasons. Nonrespondents make two types of effects on the responses. First, they are generally different from those who respond, and their exclusion can lead to biased result. Second, nonresponse reduces the sample size that can decrease the power of the study to detect differences or associations.

24. **Attrition bias** — Differential nonresponse in various groups. The pattern of nonresponse can differ from one group to the other in the sense that in one group more severe cases drop out whereas in another group mostly mild cases drop out.

25. **Bias in handling outliers** — No objective rule is available to label a value as outlier except a guideline that the value must be far away from the mainstream values. If the duration from HIV infection to development of AIDS is mostly between 6 and 10 years, some researchers would call 16 years as outlier and exclude it on the suspicion of being wrong reporting, and some would include in their calculation. Some would not exclude any outlier, however different it might be. Thus the results would vary.

26. **Recording bias** — Two types of errors can occur in recording. One arising due to inability to properly decipher the writing on case sheets. Physicians are notorious for illegible writing. This can happen particularly with similar looking digits such as 1 and 7, and 3 and 5. Thus the data entry may be in error. Second is due to carelessness of the investigator. A diastolic level of 87 can be wrongly recorded as 78, or a code 4 entered as 5 when the dependence is on memory that can fail to recall the correct code.

27. **Bias in analysis** — This again can be of two types. First, gearing the analysis to support a particular hypothesis. For example, while comparing pre- and post- values such as Hb level before and after weekly supplementation of iron, the increase may be small that will not be detected by comparison of means. But it may be detected when evaluated as proportion of subjects with level <10 mg/dl before and after supplementation. Second can arise due to differential P-values. When $P = 0.055$, one researcher can straight refuse to say that it is significant at 0.05 level and the other can say that it is marginally significant. Some researchers may change the level of significance from 5 percent to 10 percent if the result is to their liking.

28. **Bias due to lack of power** — You will soon notice that statistical tests are almost invariably used to check the significance of differences or associations. The power of these tests to detect difference or association depends to a large extent on the number of subjects included in the study—the sample size. If the study is conducted on small sample, even a big difference can not be detected, leading to a false negative conclusion. When conducted on an appropriate number of subjects, the conclusion can change.

29. **Interpretation bias** — The tendency among some research workers to interpret the results in favour of a particular hypothesis ignoring the opposite evidence. This can be intentional or unintentional.

30. **Reporting bias** — Researchers are human beings. Some can create a report such that it gives a premonitioned result yet based on evidence. It is easy to suppress the contradictory evidence by not talking about it.

31. **Bias in presentation of results** — Scale for a graph can be chosen to depict a small change look like a big change, or vice-versa. The second is that the researcher may

merely state the inconvenient findings that contradict the main conclusion but does not highlight them in the same way as the favourable findings.

32. **Publication bias —** Many journals are much too keen to publish reports that give a positive result regarding efficacy of a new regimen, compared to the negative trials that did not find any difference. If a 'vote count' is done on the basis of the published reports, positive results would hugely outscore the negative results, although the fact may be just the reverse.

The purpose of describing various types of bias in so much detail is to create awareness to avoid or at least minimise them. Some of these steps are listed in Box 4.10.

Box: Steps for minimising bias

Following steps can be suggested to minimise bias in the results. All steps do not apply to all the situations. Adopt the ones that apply to your setup.

- Develop an unbiased scientific temperament by realising that you are in the occupation of relentless search for truth.
- Specify the problem to the minutest detail.
- Assess the validity of the identified target population, and the groups to be included in the study in the context of objectives and the methodology.
- Assess the validity of antecedents and outcomes for providing correct answer to your questions. Beware of epistemic uncertainties arising from limitation of knowledge.
- Evaluate the reliability and validity of the measurements required to assess the antecedents and outcomes, as also of the other tools you plan to deploy.
- Carry out a pilot study and pretest the tools. Make changes as needed.
- Identify all possible confounding factors and other sources of bias, and develop an appropriate design that can take care of most of these biases if not all.
- Choose a representative sample, preferably by random method.
- Choose an adequate size of sample in each group.
- Train yourself and coworkers in making correct assessments.
- Use matching, blinding, masking, and random allocation as needed.
- Monitor each stage of research, including periodic check of the data.
- Minimise nonresponse and partial response.
- Double check the data and cleanse it of errors in recording, entries, etc.
- Analyse the data with proper statistical methods. Use standardised or adjusted rates where needed, do the stratified analysis, or use mathematical models such as regression to take care of biases that could not be ruled out by design.
- Interpret the results in an objective manner based on evidence.
- Report only the evidence based the results – enthusiastically but dispassionately.
- Exercise extreme care in drafting the report and keep comments or opinions separate from the results.

Bias and other aspects of design can be very adequately taken care of if you could imagine yourself presenting the results a couple of years hence to a critical but friendly audience (Elwood 2002). Consider what your colleagues could question or advise at that time, consider their reaction when you conclude that the results are significant and also if you conclude that the results are not significant. Can there be noncausal explanations of the results?

Are there any confounding factors that have been missed? Whether chance or sampling error could be an explanation? Such consideration will help you to develop proper design, and to conduct the study in an upright manner.