

13. Common Pitfalls

Many of the pitfalls discussed here have been referred to in earlier chapters in the context of the types of studies where they usually occur. It is the purpose to list and describe the pitfalls more fully in this chapter, as well as discuss different ways of dealing with them.

Bias (See also Appendix 13.1 and 13.2)

By bias is meant a systematic deviation from the truth resulting in distortion of the results. In the context of research the meaning of bias is not the same as in the literary sense. Bias means being "different". As we have noted in Chapter 2 there are three general categories of bias viz. selection bias, information bias and confounding. These are considered in detail below. But before going into the details of bias it is necessary to emphasize that once bias has been introduced into a study, however inadvertently, it is difficult to rectify it during analysis. For this reason it is important to pay careful attention to avoiding it in the design and conduct stages of the study. Different types of studies become prone to bias in different ways. For example, in case-control studies bias may arise from the prior knowledge of disease status influencing the determination of exposure status (due to more thorough probing of subjects by interviewers, or influencing recall by the subject). Alternatively, the knowledge of exposure status may lead to differential selection of subjects and controls resulting in selection bias. In cohort studies dropouts could cause a particular concern about bias. In intervention studies bias may be introduced by the nature of the comparison group, the placebo or alternative treatment used, and lax objectivity in measuring outcome.

Selection Bias

Persons being studied may differ in some significant ways from the larger population, which they are supposed to represent, or from the comparison group with whom they are supposedly comparable. Selection bias strikes at the heart of research since the essential feature of most studies is a comparison between two or more groups. The way in which individuals become study subjects is often intriguing. Certain people may be more likely than others to be selected for a study, or to continue to stay in a study. This often occurs for reasons unknown to the researcher. Whatever the underlying cause, when selective factors make the study group different in a way that can lead to mistaken conclusions, selection bias has occurred.

Selection bias takes many forms. The more common occurrences are listed below:

1. Persons who seek medical care, or get referred for specialist opinion, are likely to be those who are the sickest compared to those in the community. For example, the incidence of recurrent afebrile seizures following febrile convulsions in children is greater in studies

reported from hospitals compared to those in the general population. This is so because children with more prolonged or severe febrile seizures get referred for specialist care. For the same reasons, a hospital study of diarrhoea is very unlikely to include mild cases and would overestimate severity. In both these examples study subjects differ systematically from children with the same condition in the general population.

2. Migration bias. People with chronic lung disease tend to move out of the fumes and dust laden urban environment. A prevalence survey of chronic lung disease in a rural setting may produce the surprising finding of a high prevalence rate! On the other hand, people with emotional or psychiatric problems tend to seek the anonymity of crowded cities. A survey of inner city population may therefore show a higher prevalence rate of emotional and psychiatric problems.

3. High dropout rates. It is likely that those who continue to participate in a study are different in important ways from those who drop out. When persons lost to follow-up differ from those who remain in the study with respect to both the exposure and the outcome, the results are likely to be biased. A dropout rate in excess of 20% should be treated with caution. Similarly, people who refuse to participate in a study are different in the severity of illness, in personal characteristics, and in life-style compared to those who agree.

A high refusal rate is very likely to introduce bias. People who volunteer for a study are different from the population of cases in the general population. They tend to be better educated and have a different life-style. Studies based largely on volunteers are difficult to generalize.

4. Response bias. When subjects respond inaccurately to an investigator's questions, especially to the sensitive ones, bias occurs. How people answer a questionnaire is important. In every interview there are three factors involved - the interviewer, the respondent and the subject matter. Each of the three is a potential source of bias. In case-control studies any difference in non-response in either the cases or controls can become a source of bias.

5. Observer error. In a way this is similar to response bias. Even when identical objective criteria are used for clinical assessment discrepancies can occur between different observers. Each observer will interpret the criteria differently and apply own interpretation to the collection and evaluation of clinical information. It is useful to remember that even the most objective measurements carry some degree of subjectivity. This is as true for history taking and physical examination as it is for reading X-rays, E.C.G.s and viewing pathology slides. In case-control studies, as we have seen, prior knowledge of the subjects' disease status can easily result in differential probing about exposure. In cohort studies, information about exposure status is known at the time the disease status is being determined, and this can bias assessment of outcome.

6. Measurement bias due to faulty instruments, bad technique, change in setting, and inter-observer error is well known. If data are inaccurate or incomplete spurious associations may be introduced. A form of measurement bias viz. recall bias, has already been referred to. In most studies, some degree of inaccuracy in reporting or recording information is inevitable. This leads to misclassification with regard to either exposure or disease. If such a misclassification is random affecting all comparison groups equally then there may be no difficulty. If misclassification affects different groups differently, then there is a potential for bias.
7. The Hawthorne effect. The act of observing a phenomenon changes the phenomenon itself. People's behaviour changes just because they know they are being observed. One can never be sure that what one observes in a group during a study is exactly the same as occurs in the absence of observation. In studies where groups are being compared the Hawthorne effect is likely if only one group know that they are part of a study and the others do not.
8. Unmasking bias. The often quoted example is that of oestrogen administered as hormone replacement therapy (HRT) in menopausal women and endometrial cancer. Use of oestrogen can cause bleeding in some women, and this would lead to investigation of the cause and possible detection of endometrial cancer. In women not receiving HRT the diagnosis of endometrial cancer is that much delayed.
9. A subtle form of selection bias can occur in case-control studies. It is axiomatic that controls should be drawn from the same population as the cases. In population-based studies there is unlikely to be any difficulty in following this principle. In hospital-based studies problems can arise with the selection of controls. Each control should have had the same probability of admission as the cases. This may not be so. For example, women known to be taking oral contraceptives are more likely to be admitted for investigation of pain in calf muscles or in the chest, compared to those not taking oral contraceptives. Hence admission criteria should be carefully considered in planning case-control studies which are hospital based. In all instances where selection bias has occurred any observed relation between exposure and disease among the study subjects would be different from those who would have been eligible as subjects but were not chosen.
10. Intervention studies are open to a number of biases. One or the other group may receive preferential treatment with regard to thoroughness of investigation, treatment or ancillary care. Compliance is always a problem. Dropouts, self-medication or changes in life-styles are other difficulties to be constantly aware of.
11. Bias can creep in during analysis because of wrong assignment of cases who have missing values, or of those developing unexpected outcomes.

Strategies for Dealing with Bias

The best approach is to be aware and watch out for possible sources of bias whilst designing the study. The importance of **clear and precise definitions** for populations, disease, exposure, cases and controls, inclusion and exclusion criteria, methods of recruiting the subjects into the study, units of measurements, and so on, is obvious. **Blinding** of the subjects, researchers, interviewers, and others regarding the status of the subjects concerning exposure, intervention or group membership will reduce bias stemming from such prior knowledge. **Measurement bias** can be reduced by good quality control and prior training of technicians, interviewers, and other staff. **Careful follow up** of all those entering the study including dropouts or those who had incomplete data is essential.

Control of bias requires carefully checking the study design, method of choosing study subjects, and methods of data collection. Careful attention to selection of subjects, ascertainment of disease and exposure status, willingness of subjects to continue with the study, methods of data collection and designs of questionnaires is needed in the planning of any study. Study personnel who administer questionnaires, examine subjects or abstract data from records should be unaware of the study hypothesis as well as the disease and exposure status. Clearly written protocols and standardized training of study personnel would also help in eliminating bias. Verifying all information by checking and cross checking with different sources of data (for example about exposure status), and checking the accuracy of all data obtained also helps.

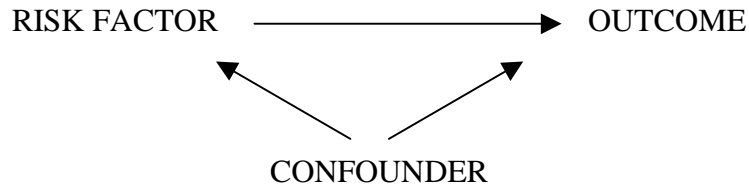
In the analysis stage, when results are in hand it is best to pause a while and ask whether the results could have been biased in any way. Whereas all analytical studies are prone to bias, there are pitfalls which are specific to different types of studies, and these have been mentioned above.

Information bias

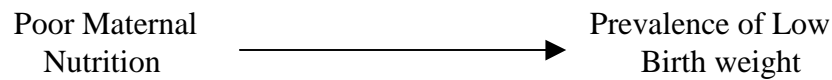
A distortion in the estimate of the outcome can occur if there are errors in measurement, or misclassification of subjects on one or more variables. Important sources include invalid measurements, incorrect diagnosis, omissions, and imprecision or other mistakes in data entry. Unequal diagnostic surveillance of exposure groups in follow-up studies is another significant cause of bias in longitudinal studies.

Confounders

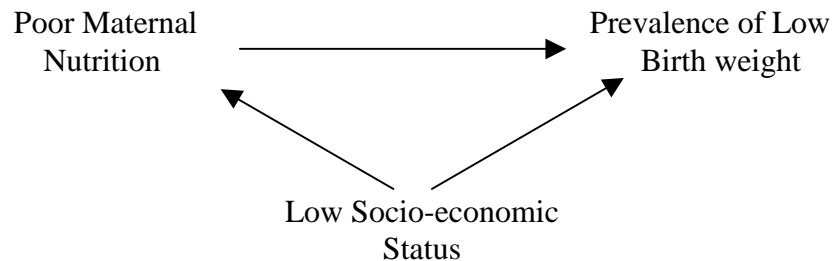
What looks like a causal relationship between a presumed risk factor and a disease may in fact be due to another factor, which has not been taken into account. Unlike bias, which is usually introduced by the investigator or study participants, confounding is caused by the complex interrelationships between various kinds of exposure and disease. The confounding factor acts by being associated with both the risk factor and the disease in a way that makes the risk factor and the disease seem to be related.



As an example, let us consider low socio-economic status and prevalence of low birth weight. Wherever studies have been carried out the prevalence of low birth weight is higher in women from the lower social class, and one may write this association as follows:



But this relationship may be due to poor maternal nutrition which is more common in women from the lower social class, and the relationship may be written as:



Such a situation where the true association between a risk factor (low socio-economic class) and outcome (low birth weight) gets blurred because of another factor (poor maternal nutrition) which is distributed differently between study and comparison groups is called *confounding*.

A confounding variable must be a risk factor for a disease. It need not necessarily be a cause of the disease, but it could be a marker of a cause or for the diagnosis of the disease.

A confounding variable must satisfy the following conditions:

- (i). It should be associated with both the outcome (disease) and the risk factor under study.
- (ii). It should be extraneous to both the disease and exposure but distort their relationship.
- (iii). It must not be an intermediate step on the causal pathway between exposure and disease.

If the extraneous factor lies on the causal pathway between exposure and outcome it cannot be considered a confounder. In such a situation it acts as *effect modifier*.



Similarly, by definition, if the extraneous variable influences the risk factor but not the outcome it cannot be considered a confounder.

At the time of planning a study it is not always clear which factors are confounders. Mistakes can be avoided by carefully scrutinizing published reports of similar studies, discussing with other researchers working on same or similar topics, and collecting adequate information on the would-be confounders so that one can take them into account during analysis.

It is a good practice always to ask "Is there some plausible explanation for this finding besides direct cause and effect?" In particular this question should be asked whenever an association has been found during analysis of the data which was intended for answering another research question. Such unplanned observations about associations are highly susceptible to spurious conclusions.

During analysis confounding may be suspected if the odds ratio gets altered after adjustment for another factor. This is called change-in-estimate criterion. In other words, one first obtains the odds ratio for the exposure variable of interest. Next one calculates the same for different values (strata) of the suspected confounder. If the odds ratios for different values of the suspected confounder are not materially different then it is not a confounder. Another quick test is to see if the would-be confounder is unevenly distributed between the study and comparison groups. If it is then it should be taken as a warning signal. However, in order to be able to perform these checks it is necessary to include all possible confounders during the design stage of the study so that adequate data can be collected about them. Uncontrolled confounding is a major threat to the validity of results. To avoid this, it is necessary that the design of the study permits adequate data collection to address the issue during analysis.

Strategies for dealing with Confounders

In intervention studies randomization of subjects and enrolling a sample of adequate number of controls on whom information has been collected for known and unknown confounders is helpful.

In the case of observational studies an important step is to make the groups as similar as possible. Some of the basic precautions to be taken in the early planning stage like, for example, careful scrutiny of all published reports on the topic and discussions with colleagues as well as those known to be working on similar topics have been mentioned earlier. At the time of designing the study one should

first list all those factors which are likely to be confounders, and then decide how to deal with each in the design and/or analysis stages as follows:

Design stage strategies

One strategy is to use very strict inclusion criteria so that the would be confounder is kept the same in the groups under study. Rigid inclusion criteria usually ensure that major confounders get excluded. The disadvantage is that the groups will become very small (restricted), and the generalizability is affected. To take an example, in a study about availability of clean water within the household and prevalence of diarrhoea maternal education level can act as a confounder. To overcome this problem one may recruit only households with illiterate parents. This would put a limit on the number of subjects eligible for recruitment into the study. Such limitations can become additive when several confounders are to be dealt with simultaneously.

A second strategy is *matching*, so that for each case one or more controls with the same value for the confounding variable are selected. The decision about matching should be done *before* starting the study, and is irreversible during the course of the study. If the variable for which matching has been done lies on the causal pathway it is not possible to study its effects on the outcome, and bias may result. Matching is most effective when the nature of a major confounder is known, or when confounders are better matched than measured (e.g. life-style, neighbourhood, sibship, and so on).

The purpose of matching between study and comparison subjects is to ensure that "cases" and "controls" have equal susceptibility to a given disease in terms of all risk factors *except the one of interest*. It is often difficult to know which factors to match for. In general, cases and controls are matched for all those risk factors which are known to be related to both the exposure and the disease and therefore may act as confounders.

Undermatching occurs when important confounding factors are not accounted for in the matching process. When under-matched, a case-control study can come up with strong associations which are in fact spurious because of the underlying effect of an uncontrolled risk factor. Various statistical and analytical methods may need to be used at the conclusion of the study to control for the confounding effects of unmatched variables.

Over-matching occurs when controls excessively resemble the cases. In an over-matched case-control study one fails to discover an association that is present and is real. For example, in a case-control study of AIDS, homosexual men attending a clinic for sexually transmitted diseases who were HIV negative were selected as controls. When the exposure factor 'anal intercourse' was analyzed no difference was seen between cases and controls.

Over-matching makes differences in exposure rate less apparent even when difference is present. The real difference in exposure gets "diluted out". Matching is a tricky job, and there is no simple answer. Significant associations may change with strategic changes in the character of the match.

Another strategy is **randomization**. It is by far the best method in intervention studies. It can ensure that all potentially confounding factors, both currently known and unknown, are evenly distributed between the study groups.

Analysis stage strategies

Stratified analysis and multivariate modeling are the two commonly employed strategies for identifying independent effects of variables.

In stratification, the relation between exposure and outcome is separately assessed for different strata or levels of the confounding variable. In the above example of clean water in the home and the prevalence of diarrhoea the study and comparison households may be grouped by years of mothers' schooling, and the odds ratio for each group is then calculated. If the difference in the prevalence of diarrhoea is demonstrable across the groups then one may safely conclude that a strong association exists between availability of clean water and the prevalence of diarrhoea. The main disadvantage of stratification is that the number of confounders which can be controlled for simultaneously is small. Another problem is one of over aggregation during grouping. For example in all analyses the variable "Age" is aggregated into "Age groups" for convenience. If the band is too wide (e.g. 10 years) there is a strong likelihood of residual confounding even after stratification. A narrower band (e.g. 5 years) is preferable. Similarly, with smoking residual confounding is more likely when only two categories - "smoker" / "non-smoker" are used as compared to when the groups are "non-smoker", "ex-smoker", "light moderate and heavy smoker". Another type of difficulty is when a variable is used as a proxy measure, and is in reality a very poor proxy. For example, socio-economic class is usually a proxy measure for life-style and cultural practices. Several attributes may have been used for assessing socio-economic status. It may be more informative to analyze the data using these individual attributes as well as their composite score for socio-economic class.

Multiple regression analysis is another strategy which helps overcome the problem of controlling for several confounders simultaneously. The danger here is that of inappropriate use, which is very likely if the attitude is leave-it-to-the-computer-and-stand-back. The advice of a statistician can be invaluable in deciding about regression analysis.

Defining the denominator

Disease rates constitute an important tool in all medical research. Numerator alone does not tell much. For example, case reports is a good source of numerator data, but at best they are only a beginning of an investigation.

In cross-sectional studies subjects are examined or interviewed once. The number of "cases" may then be divided by the total number of subjects in the study to give the **prevalence rate**. In this example the denominator is made up of the total number of persons in the study at the time. Another measure of disease frequency in such a study is the number of cases divided by the number of people without the disease of interest. This measure gives the **prevalence odds**.

In studies where a defined population is followed over time to record cases of a disease as they occur, the denominator is the population initially at risk. The number of incident cases (the numerator) divided by the population at risk (the denominator) gives the cumulative incidence of the disease. A common problem is that cohorts change over time because of migration, being lost to follow-up, or death due to some other cause. To overcome this difficulty another type of denominator is used viz. person-time unit. Each individual followed over one unit of time represents 1 person-time unit. For example, 1 subject followed for ten years is 10 person-years equivalent to 2 subjects followed for five years, or 10 subjects for one year.

During the course of a study the researcher's attention is focused on the subjects being studied, or the numerator. It is equally important not to lose sight of the denominator.

Validity

In all studies the most important issue is "Are the conclusions true?" The best way to ensure validity is to recognize threats to it, and carefully watch out for them. The common threats to validity are:

1. Confounding factors.
2. Unexpected causes beyond the investigator's control might have produced the same effects as the interventions under study.
3. Differential loss of subjects in various groups. Those who dropped out of a study or control group may have done so because of factors related to the characteristics one is studying.
4. Selectivity or bias in assigning subjects to various groups.
5. Instrument errors causing distorted response.
6. Hawthorne effect - What is observed tends to show improvement.

Strategies to deal with threats to validity.

1. Having a control group helps against confounding, various unexpected events occurring, and the Hawthorne effect.
2. Random assignment of subjects to different groups.
3. Before and after measurements.
4. Unobtrusive methods of observation, especially when the subject is likely to change behaviour when being observed.
5. Careful research design. If instruments are to be used then pre-testing of instruments as well as quality control are needed.
6. Knowledge of various environmental events, especially when studies are over a prolonged period.

Statistical Pitfalls.

The conclusions derived from a study always relate to the target population, which is defined by the sampling unit. Most statistical tests assume that the sampling units are selected at random from a well-defined population. Also that each observation is independent of any other. Such assumptions are necessary because of the wide variations between individuals. It is for this reason that wherever possible researchers seek to do randomization. The purpose is to distribute subjects among groups in an unbiased fashion. However, when subjects differ groups will also differ no matter how they are formed. It is because of the variation between individuals that researchers describe their subjects not only with regard to the average value but also to the amount of variation present. The standard deviation is a measure of the variation of individual values around the mean. The standard error of the mean is the variation of sample means around the true mean, which is always unknown. Another point to bear in mind is that the main purpose of significance testing is to account for the potential for distortion in the results due to variations amongst subjects.

In carrying out research the investigator is arguing from the evidence to the truth. Random variation and bias can lead to the emergence of many potential explanations, many of which are spurious. The problem is made worse by the availability of powerful computers and software that make data dredging easy. One should be aware that hypotheses resulting from such fishing expeditions are coincidences stemming from subject variation. When the conclusions of a study are suspected of arising from the mere churning of data, one should ask the following:

1. Was the analysis intended at the commencement of the study?

2. Is the analysis likely to be prone to bias?
3. What came first - Did the idea originate the observation or was it the other way round?
4. Is the conclusion biologically feasible?

Appendix 13.1

Biases in specifying and selecting the study sample**Admission rate bias (also called Berkson bias)**

If hospitalization rates differ for different exposure/disease groups, the relation between exposure and disease will become distorted in hospital-based studies.

Specialist hospital bias

The reputation of certain hospitals (and physicians) makes subjects with specific disorders to gravitate towards them.

Unmasking bias

An innocent factor may become suspect if it causes a sign or symptom which sets off an intensified search for the disease.

Diagnostic access bias

Subjects differ in their geographic, social, and economic access to the diagnostic facilities which label them as having a disease.

Diagnostic vogue bias

The same illness may receive different diagnostic labels at different times, and places.

Conceptual bias.

Faulty conceptualisation of the research question can lead to faulty interpretation and conclusions.

Design bias

Studies with faulty designs, methods, sampling procedures, and/or group assignment procedures, or when inappropriate techniques of analysis have been used can lead to wrong conclusions.

Bias in handling outliers.

This can arise from failure to discard an unusual value occurring in a small sample; or the exclusion of unusual values which should have been included.

Membership bias

Membership of a group (unemployed, slum dwellers, obese, heavy drinkers or smokers) may imply a proneness to illness which differs from the general population.

Non-contemporaneous control bias

Changes over time in definitions, diagnoses, exposures, treatments may render non-contemporaneous control not comparable.

Non-respondent bias

Non-respondents or late recruits in a sample may exhibit exposures or outcomes which differ from those of responders or early recruits. No response is a major source of potential bias, as it reduces the effective sample size resulting in loss of precision of the outcome estimates.

Popularity bias

The admission of patients for certain treatments and procedures is influenced by the interest stirred up by the presenting condition.

Prevalence-incidence bias

Short lived and fatal episodes are missed in cross-sectional studies. Similarly mild or 'silent' cases and cases in which evidence of exposure disappears with disease onset get missed out.

Previous opinion bias

Procedures and results of a previous diagnosis, if known, may influence subsequent diagnostic processes on the same patient.

Treatment selection bias

Certain clinical procedures may be preferentially offered to those who are poor risks. (For example, selection of patients for medical versus surgical treatment).

Referral filter bias

As patients get referred from primary to secondary and tertiary care facilities the proportion of severe, or rare conditions, or with multiple diagnoses may increase.

Unacceptable disease bias

When disorders are socially unacceptable (insanity, venereal disease, leprosy and so on) they tend to be under reported.

Volunteer bias

Volunteers in a sample may exhibit exposure or outcome characteristics which differ from those of non-volunteers.

Appendix 13.2
Type of study and main possible sources of errors

Study Type	Selection Bias	Information Bias	Diagnostic Bias	Confounding
Cross-sectional	Yes	Yes	Yes	Yes
Case-control (exposure status based on recall)	Yes	Yes	No (if blinding done)	Yes
Case-control (exposure status based on records)	Yes	No (exposure data obtained prior to diagnosis)	No (if blinding done)	Yes
Prospective Cohort	No (all at risk interviewed)	No	Yes	No
Retrospective Cohort (exposure status based on records)	No	No	Yes	Yes
Retrospective Cohort (exposure status based on recall)	No	Yes	No (if blinding done)	Yes
Nested case-control (exposure status based on records)	No	No	No (if blinding done)	Yes
Clinical trial (double blind)	No	No	No	No