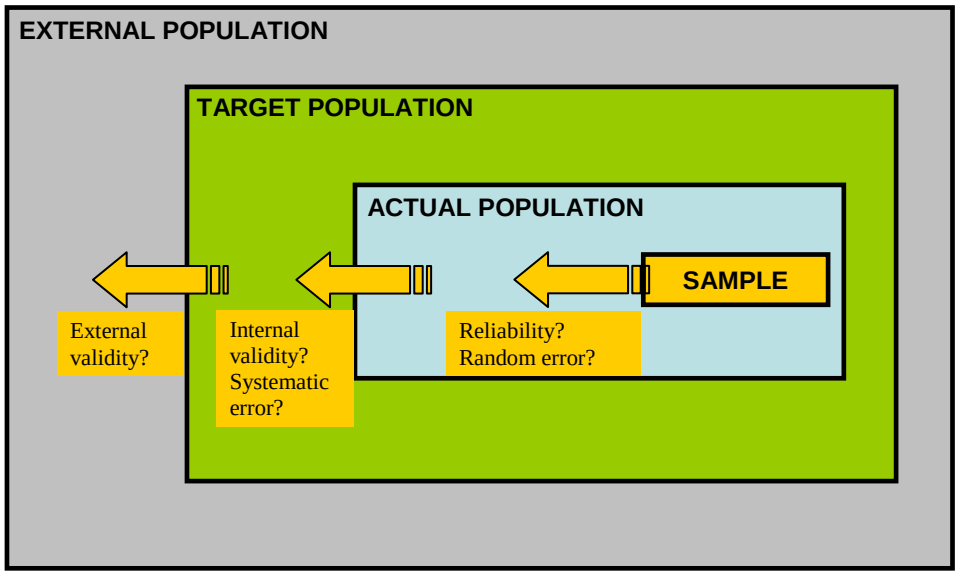
Dissertation Module: Research Skills Program
Topic 8: SOURCES OF BIAS

THE CONCEPT OF BIAS

Epidemiological research is conducted to increase knowledge about diseases and related issues with the general aim of identifying risk factors, preventing future disease, improving treatment, and improving health service delivery to a community. We discussed the epidemiological research cycle at length in Chapter 1 and we have worked through numerous examples. We know that epidemiological research studies people. However, usually the people who participate in the research are not all the people to whom we would like to make inferences. We usually work with a sample of people and hope that the results gained from the study can be **generalized** to a much larger group of people. In order to distinguish formally between the people who are actually being studied and the people to whom the results might be generalized, we use the words **sample** and **population**, respectively. We differentiate between several populations as depicted in Figure 1.

Figure 1: The hierarchy of populations



Box 8.1: Sample and Population

Example: We want to conduct a study to determine the occurrence of stuttering in pre-school children in Australia. We approach day-care centres in Sydney, Melbourne and Adelaide to recruit children into our study. The parents of 265 children aged 3 to 5 years from 15 day-care centres consent to participate in the study. The 265 children are examined

for stuttering and their parents are interviewed to collect information about possible factors which influence the occurrence of stuttering such as traumatic experiences (for example, sexual abuse, severe neglect), maternal behaviour during pregnancy (for example, smoking, alcohol abuse) and home environment (for example, extent of social contact with the child).

In this example, the 265 children and their parents are the sample.

Definition of a sample:
In general, a sample is a selected subset of a population. The sample is the group of individuals from whom the study data are collected.

Definition of a target population:
The target population is the group of people from whom the sample has been taken and about which you want to draw conclusions. The sample may or may not be representative of the target population. The target population is partly defined by the exclusion and inclusion criteria of a study. Please note that when we plan a study it is most important to define the target population first to ensure appropriate sampling! Unfortunately, many publications do not specifically provide information about the target population of their studies so we often have to guess to whom the researchers wanted to generalise their results.

In our example above the target population was all Australian pre-school children.

Definition of actual population:
The actual population is the group of people from whom the sample was drawn and that the sample represents. Thus, the actual population is an expansion of the sample and includes all individuals who fulfil the study inclusion and exclusion criteria. Please note that the actual population is often hard to define, and who belongs to the actual population can be a matter for substantial debate.

If we want to identify the actual population of our example, we need to think about the sample and the sampling procedure.
The study involved children aged 3 to 5 years who attended day-care centres in Sydney, Melbourne and Adelaide.
Who do these children represent?

- Australian children aged 3 to 5 years;
- Who attend day-care
- Resident in large southern Australian cities

In addition, day-care centres as well as parents and children had to consent to participate in our study……

If the sample represents the target population, then the actual population and the target population are identical rendering the description of the actual population superfluous. However, if the sample does not represent the target population, then the actual population

has to be constructed as the actual population is the group of people to whom findings of an epidemiological study can be generalized if the sample does *not* represent the target population.

In our example, the sample did not represent the target population of all Australian pre-schoolers. Children who do not attend day-care and children from regional, rural or remote locations were not represented. These omissions might have an effect on the results of the study!

Definition of external population:
The external population is the wider group of people who do not fulfil the selection criteria of the study but to whom one also wishes to infer at least some findings of the study.

In our example, the external population of interest could be all pre-school children from developed countries.

Figure 1 illustrates the situation. The sample is sampled from and therefore part of the actual population which in turn is part of the target population which in turn is part of the external population.

Please note that the actual population, the target population, and the external population might not be clearly defined in scientific publications. Therefore, when we read articles we need to critically assess the sample and the sampling procedure in order to decide for ourselves to whom the results of a study can be generalized. This implies that when reading published work we might not be entirely sure who is included in the actual population and/or the target population.

In order to allow generalization of the study results from the sample to the wider target population the sample has to be representative of the target population. However, representativeness, although necessary, is not sufficient for generalizability.

Box 8.2: Generalizibility of a study

Generalization of results from an epidemiological study to the target population is possible if:

(1) The study results are reliable, that is, if the study was repeated under the identical conditions, the results would be similar.
Reliability (also called: consistency, repeatability) refers to random error. The smaller the random error the more reliable the results.
Please note: Every sampling involves a certain amount of random error. Thus random error cannot be completely avoided in epidemiological studies but it can and should be kept reasonably small by enrolling a sufficiently large sample size. Random error is assessed during the statistical analysis of the data collected in a study.

(2) The study results are valid, that is, if the study was repeated in another setting but still within the same target population, the results would be similar.
Validity (also called: conformity) refers to systematic error. The smaller the systematic error the more valid the results.
Please note: Systematic error can, at least theoretically, be avoided all together or at least minimized by meticulous attention to the design and conduct of the study.

In our example in Box 8.1 Reliability (1) refers to a repeat of the study recruiting different children but in an identical manner, that is, from the same day-care centres located in Sydney, Melbourne, and Adelaide (same actual population), collecting data with the same instruments. High reliability, that is, small random error, implies findings similar to the original study. Indeed in this example, we would expect similar results.

Validity (2) refers to a repeat of the study in another setting – but still within the realm of the target population which were all Australian pre-school children (see Box 8.1). Thus we could think about enrolling children from regional, rural and remote areas; or children who do not attend day-care (same target population but different actual populations).

Since it is quite obvious that in our example the originally sampled children do not represent the target population, internal validity is not given and the study results probably can not be generalised to all Australian pre-school children.

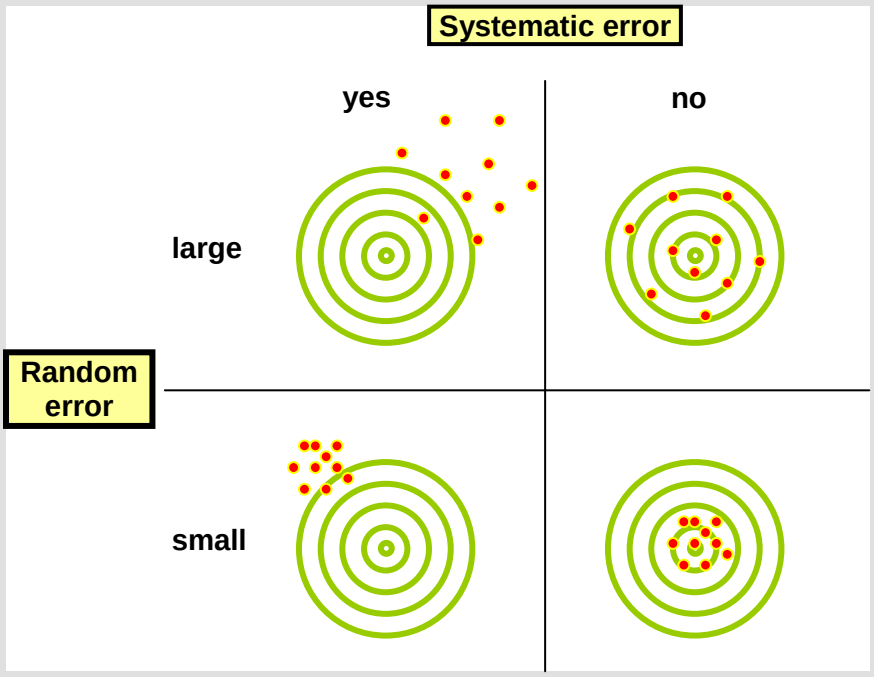
Condition (1) of generalizibility deals with the question of generalising sample findings to the actual population. This condition addresses the question of whether a study result could be repeatedly gained by further studies conducted within the same actual population or whether the initial results occurred by chance alone. Of course, in reality repetitions of a study are not performed nor are they necessary, since the question of **reliability** - referring to random error - is dealt with by **statistical inference**. Statistical inference is a statement about the likelihood that the observed findings of a study have occurred by chance alone, that is, by random error.

Condition (2) of generalizibility deals with the question of generalising sample findings to the target population. This condition addresses the question of whether a study result can be repeated by further studies conducted within a different actual population but still within the same target population. This question of **interval validity** refers to systematic error and is one of the central issues of epidemiological research. Validity is dealt with by a thorough and detailed study plan and by conducting the research in a flawless and meticulous manner.

Sometimes we want to generalize even beyond the target population to an external population which includes the target population but is not restricted by the exclusion and inclusion criteria of the current study. This generalization then hinges on what is called **external validity**. Whether or not a study has external validity is a matter of our judgement and requires our skills in critical appraisal as well as an in-depth understanding of the subject matter at hand.

Box 8.3: Random and systematic error

Figure 2: Schematic representation of random and systematic errors.



Random error occurs in any sampling but it can be large or small. A small random error implies that if a study was repeated under identical conditions the results would be very similar; that is the dots in Figure 2 are close together. However, a small random error does not automatically mean that we are close to the target! We are only close to the target if there is no systematic error and we have only a small random error.

Comments:

(1) Random error can be controlled rather elegantly by conducting a sample size calculation during the design phase of a study. Random error can be limited if we use a sample large enough for the specific research hypothesis being studied and by applying the correct statistical analysis.

(2) On the other hand, and as discussed in detail throughout the rest of this chapter, it is indeed very difficult to control systematic error completely. Most studies involve at least some potential for systematic error. It is up to us to identify potential systematic error when designing and conducting a study and when reading the publication of a study.

Interval validity is one of the central issues of epidemiological research. Consequently, a main part of every epidemiological study is devoted to the discussion of potential systematic errors, how they can be avoided, and whether they might influence the results. The central question of internal validity is whether systematic error occurred during the conduct of a study which led to distorted results or whether the results are unbiased and therefore valid

for the target population. Indeed, systematic errors often lead to distortion of the results. If this happens, then we talk about **biased results**. The avoidance and control of bias is an extremely important issue in any epidemiological research because only results from unbiased studies are valid for the target population.

Internal validity in epidemiological research hinges on **three pre-requisites** which are listed in Box 8.4.

Box 8.4: Pre-requisites for internal validity

(1) Structural uniformity
The structural characteristics and potentially influencing factors of the sample(s) being studied have to be as alike as possible.
For instance, in the example in Box 8.1, we aimed to compare children who stutter with children who do not stutter. Structural uniformity means that these two groups should be of similar age, have the same or a similar gender ratio, and so on. All potentially influencing factors, which we do not wish to examine as study factors, should be similar in the groups that are being compared.

(2) Observational Uniformity
The sample(s) must be observed in the same way, with the same intensity, making the same measurements and observations in the same way, and with equal documentation. The information collected must be valid.
For the example in Box 8.1 this means, that it is not advisable to use information which is available in speech pathology files for the children who stutter and only send questionnaires out to the families of the unaffected children.

(3) Representative Uniformity
The sample(s) must represent the target population.
The sample of children attending day-care centres in Sydney, Melbourne and Adelaide in the example in Box 8.1, is probably not representative of all pre-school children in Australia, which in this example is the target population. Therefore, representative uniformity is not provided in this example.

These three essential pre-requisites will also lead us to the definition of the three main types of bias in epidemiological research. But let's begin by defining bias in general.

Box 8.5: Definition of bias

If systematic error occurs when conducting an epidemiological study, which leads to misinterpretation of the effect measure (for example prevalence, relative risk, or attributable rate) this misinterpretation is called **bias**. If there is no misinterpretation the effect measure is

called **valid** (or unbiased).

Source: Kleinbaum, D.G., et al., 1982.

Comments:
(1) A bias renders the results of a study invalid, that is, incorrect – incorrect for the target population. If no bias exists then our results are valid – valid for the target population.
(2) The above definition means that bias actually affects our results! A study might have a systematic error, for instance, in the sampling regimen, however as long as this error does not affect the results of the study it does not constitute a bias.
(3) "Effect measure" in the above definition also refers to studies where the aim is to estimate a single characteristic. For example, a study which sets out to estimate the prevalence or incidence of a disease.

In the example in Box 8.1 we identified problems with the sampling approach, that is the children from the large southern Australian cities may not represent all Australian children and the children attending day-care may be different from children who do not. These differences could have an effect on the results of the study and hence the results of the study are potentially biased.

Books and articles often refer to a confusing multitude of apparently different biases, using a disturbing number of names (see Box 8.6 for some examples). Fortunately most of these different biases can be categorized into three main types: **selection bias, information bias** and **confounding bias.**

Box 8.6: The spinning world of BIAS

Bias of rhetoric; All's well literature bias; One-sided reference bias; Positive result bias; Hot stuff bias; Popularity bias; Centripetal bias; Referral filter bias; Diagnostic access bias; Diagnostic suspicion bias; Unmasking bias; Mimicry bias; Previous opinion bias; Incorrect sample size bias; Admission bias; Prevalence-incidence bias; Diagnostic vogue bias; Diagnostic purity bias; Procedure selection bias; Missing clinical data bias; Non-contemporaneous control bias; Starting time bias; Unacceptable disease bias; Migratory bias; Membership bias; No respondent bias; Volunteer bias; Contamination bias; Withdrawal bias; Compliance bias; Therapeutic personality bias; Bogus control bias; Insensitive measure bias; Underlying cause bias; End-digit preference bias; Apprehension bias; Unacceptability bias; Obsequiousness bias; Expectation bias; Substitution game; Family information bias; Exposure suspicion bias; Recall bias; Attention bias; Instrument bias; Post-hoc significance bias; Data dredging bias; Scale degradation bias; Tidying-up bias; Repeated peeks bias; Mistaken identity bias; Cognitive dissonance bias; Magnitude bias; Significance bias; Correlation bias; Under-exhaustion bias; Clinical practice bias; Do something bias; Favoured design bias; Resource allocation bias; Prestigious journal bias; Prominent author bias; Famous institution bias; Flashy title bias; Substituted question bias; Credential bias; Esteemed professor bias; Rivalry bias; Personal habits bias; Moral bias; Friendship bias; Tradition bias; Territory bias; I am an epidemiologist bias; …..

Source: Spilker, B.L. (1991)

Let us introduce the three main types of bias by an example. We are investigating sexual violence in Australian prisons (study idea based on Richters, J. et al. 2009) using a cross-sectional survey. A random sample of 2,626 male prisoners from N.S.W and Queensland were invited to participate in the study; 2018 prisoners consented. Amongst other things we are interested in finding out whether sexual coercion in prison is more frequently among homosexual male prisoners than it is among heterosexual male prisoners. In this example, the study factor is sexual orientation and the outcome is sexual coercion, which was defined as being "forced or frightened into doing something sexually that [they] did not want".

To introduce the three main biases by using this example let us now assume:

(1) During the analysis of the data we find out that the homosexual men were on average 5 years younger in comparison to the heterosexual men and that a younger age in general increases the prevalence of sexual coercion (the outcome).
CONFOUNDING: This finding shows that the age of the two groups to be compared is different (structurally dissimilar). Therefore age could introduce a confounding bias to our study: the homosexual men are more likely to be sexually coerced because of their younger age, irrespective of their sexual orientation; the resulting prevalence odds ratio (POR) for assessing the relationship between sexual orientation and coercion is likely to be overestimated because of the difference in age structure. Please note that if structural uniformity is violated, this can result in confounding bias.

(2) Sexual orientation is not a topic that is easily talked about and thus some participating prisoners may not have disclosed their homosexuality when interviewed.
INFORMATION BIAS: It is likely that homosexuality was systematically under-reported by the participants systematically not disclosing the "study factor". Therefore, if there was an association between sexual orientation and sexual coercion, this association would probably be underestimated by the observed POR because of systematically introduced incorrect information about sexual orientation. Please note information bias occurs if the observational uniformity or correctness of collected information is violated.

(3) For a correct study design we needed to invite all of the 2,626 male prisoners who were part of our initial random sample to participate in the study. Overall, 2,018 (76.8%) consented to participate.
SELECTION BIAS: Let us assume that men who decided not to participate in our study were also less likely to have encountered sexual violence. As a consequence, our sample is not representative of the target population, which is all male prisoners in N.S.W. and Queensland. In this case, the prevalence of the outcome would be systematically too high. Please note there can be selection bias if the sample does not represent the target population.

In reality, no study is 100% perfect! There is always random error as well as systematic error in any epidemiological study. Random error can be assessed by statistical methods and is controlled for as long as we follow the rules of statistical planning and analysis. The central problem for most epidemiological studies is to identify potential sources of systematic error. Ideally we try to avoid potential systematic error during the design phase of the study. But even then complete avoidance of bias might not be possible.

For example, selection bias in our study about stuttering in children could have been minimised if we had sampled appropriately. On the other hand, in the study of sexual violence in prison it would be difficult to completely avoid bias as there is not much that can be done if people do not want to participate or do not want to disclose their sexual orientation or unpleasant sexual experiences. In specific situations it might be possible to address bias during the analytical phase of the study. However, in many situations the additional information necessary to do so is unavailable, and would have to be collected in further studies. Thus often the only way to deal with bias after the study has been completed is critical review of its possible effect.

In most studies the size of bias cannot be directly measured as additional independent studies would be required to measure its extent. Then again, it is sometimes possible to judge and discuss the suspected **direction of a bias**. However, before we introduce the challenging concept of direction of bias, we should discuss the three main biases in some detail.

Selection Bias

Selection bias is one of the three main types of bias and constitutes a fundamental challenge for epidemiological research since hardly any study is completely free of it.

Box 8.7: Definition of selection bias

Selection bias refers to a distortion in the effect measure, resulting from the manner in which the people are selected for the sample. Thus, selection bias may be introduced if sampling techniques are inappropriate. If selection bias occurs, the sample(s) do not **represent** the target population.

Source: Kleinbaum, D.G., et al., 1982.

Comment:
Bias occurs when the effect measure, for example the odds ratio from a case-control study, is distorted and does not give a reasonable estimate of the effect measure in the target population. Selection bias occurs when this distortion is due to the way the sample was selected.

Please note that producing a random sample or another form of probability sample should reduce problems with selection bias.

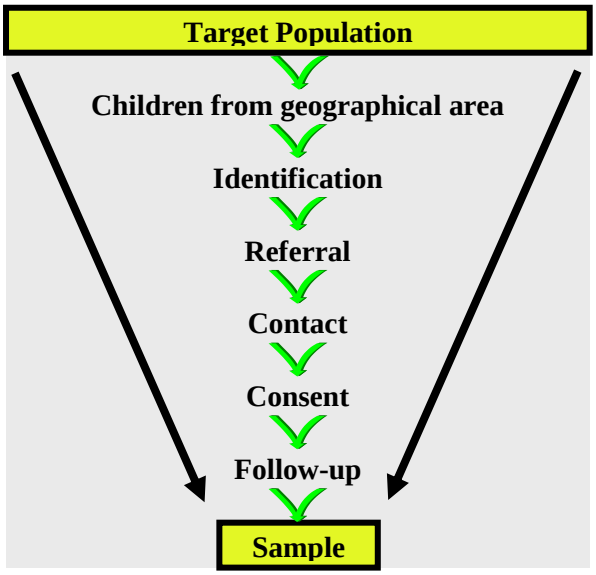
There is an high potential for selection bias in every study, because the procedures necessary to obtain a sample are inevitably bound to feasibility. It is rather difficult, often quite complex and a demand on resources to achieve a true random sample.

For example, assume we are planning a randomised controlled trial to investigate whether stuttering in pre-school children can be improved by using an intervention program

developed by speech-language therapists (Jones, M. et al., 2005). Our target population is pre-school children aged 3 to 6 years with a frequency of stuttering in at least 2% of syllables. The ideal sampling would be randomly from all Australian pre-school children who stutter at this level. However, no complete list of these children is available and hence we will be unable to sample randomly. In addition, even if there were such a list, the practicality of including children from different locations all over Australia would lead to a prohibitively resource intensive study. Thus we will instead include pre-school children who stutter at least 2% of syllables from one or two speech pathology clinics, preferable close to where we live. We will be working closely with the staff of these clinics, so that they support the study by identifying and referring the children and their parents to us. We then will contact these families and obtain consent for them to participate in the study. Within our trial we plan to follow-up the children for one year so we can show any long-term benefits of the intervention. The results of our study will be based on the children who have a final outcome assessment (= sample). Figure 3 shows a step-wise representation of the selection process in this hypothetical example. Please note that selection bias could be introduced at each step in the selection process.

Selection bias can occur at every step from the target population to the sample. At every single step, a subgroup is selected which may not be representative of the whole group and consequently could introduce selection bias. In the example where we wanted to represent all Australian pre-school children aged 3 to 6 years with a frequency of stuttering of at least 2% of syllables, one of the selection processes was identifying children via a speech pathology clinic. Children who seek treatment for stuttering at a clinic might be different from those who do not. In another selection step, parents and children had to consent to participate in the study. Families who consent might be different from those who do not, and so on.

Figure 3: From the target population to the sample



Unfortunately it is difficult to identify potential sources of selection bias in general as the selection procedures are specific to each study. However, some common potential sources for selection bias are worth mentioning: (1) almost all studies require people to consent to participate introducing a potential "volunteer bias"; (2) often people do not respond during data collection; and (3) frequently a substantial proportion of people are lost during follow-up. These potential selection problems are often helpful considerations when reading a publication or planning a study.

Selection bias – Example 1: Volunteer bias

Volunteer bias is a potential source of distortion of the effect measure in nearly all epidemiological studies as participants usually have to provide their consent to participate (written informed consent). It is quite likely that people who agree to participate in a study (the "*volunteers"*) differ from people who decline. Depending on the topic being studied and the country where the study is conducted, volunteers may differ from the people who decline to be involved by being more or less informed, educated, wealthy, desperate, et cetera. In most cases it is extremely complex to quantify volunteer bias.

Selection bias – Example 2: Detection Bias

Detection bias is a special form of selection bias in case-control studies resulting from differing intensities of observation between cases and controls. A classic example of detection bias was first described by Ralph Horwitz and Alvan Feinstein (1978) who claimed that previous findings from case-control studies assessing the association between the use of oestrogen (study factor) and endometrial cancer (outcome) were spurious because of a form of selection bias they called detection bias. This bias resulted from a more intense

surveillance of women who used oestrogen than of those who did not use oestrogen. Horwitz and Feinstein argued that users frequently suffer from oestrogen induced vaginal bleeding, and are therefore more frequently examined. As a result endometrial cancer is more likely to be detected in users than in non-users of oestrogen artificially inflating the exposure odds ratio.

Horwitz and Feinstein argued that detection bias can be avoided if a "proper" control group is used. They suggested that in the investigation of oestrogen and endometrial cancer the controls for a case-control study should be recruited from gynaecological patients who had undergone procedures such as a dilatation and curettage or hysterectomy because of uterine bleeding (Horwitz, R.I. and Feinstein, A.R., 1978). However, their approach was heavily criticised by George Hutchison and Kenneth Rothman (1978) who argued that invasive endometrial cancer will ultimately be detected and, therefore, endometrial cancer will only be slightly over-represented in oestrogen users. Hutchison and Rothman suggested that Horwitz and Feinstein had actually introduced their own selection bias by using control women who were too similar to the case women, thereby leading to a considerable underestimation of the relationship between oestrogen use and endometrial cancer (Hutchison G. B. and Rothman K.J., 1978; Greenland, S. and Neutra, R., 1981). The estimates for the odds-ratios did in fact decrease from almost 12 when comparing women who did and did not use oestrogen to 1.7 when controls were used as suggested by Horwitz and Feinstein (1978).

Despite the controversial discussion by these eminent epidemiologists about how to best avoid the described bias, they agreed that **detection bias potentially occurs whenever the procedure to identify disease status varies with the exposure status**. In his book "Clinical Epidemiology" Feinstein (1985a) comments that "detection bias arises from natural phenomena that lead to differences in the medical surveillance, ordering of tests, or interpretation of outcome evidence for the recipients of different manoeuvres." Detection bias can be produced by concentrating attention on a new intervention or exposure.

Please note the mechanism behind detection bias is not restricted to case-control studies, however, it is only an example of selection bias in case-control studies. Detection bias in cohort studies is classified as information bias! If we consider the above example of oestrogen use and endometrial cancer in a cohort setting then the same mechanism would still occur (the procedure to identify endometrial cancer still varies with exposure status), however, now the bias is about the differences in observing exposed and unexposed women and not about the way in which cases are identified and thus selected.

Selection bias – Example 3: Self-selection bias and Selective survival

Self-selection bias is similar to volunteer bias which was described previously. In his book "Modern Epidemiology" Kenneth Rothman and co-workers (2008) presented a classic example of self-selection bias. The Centers for Disease Control (CDC) investigated the incidence of leukaemia in troops who had been at the Smoky Atomic Test site in Nevada. Overall 76% of the troops could be traced; of those 82% were identified by the investigators and 18% self-referred. The results of this study showed why self-selection might cause issues for validity; 4 cases of leukaemia were diagnosed in the 82% of participants who were

traced by the investigators, while another 4 cases were diagnosed amongst the 18% of self-referred participants (Caldwell, G.G., et al., 1980). Self-referral can potentially create a large bias in a study because the reason(s) why people self-refer are often linked to the outcome. Imagine the headlines if there had been only self referred soldiers in the study!

Selective survival is a form of selection bias that can occur in cross-sectional studies or in case-control studies that use prevalent cases. The problem arises because people who developed the disease and "died" before the study began obviously can't participate in the study. However, the "surviving" people may no longer be representative of the entire target population. If exposure status is either over- or under-represented among the survivors, the use of prevalence data can lead to biased results.

We have written "died" and "surviving" in inverted commas, as death is just one possibility – and is responsible for the name of the bias. In Chapter 8 we introduced the "healthy worker effect" (McMichael A.J., 1976). The healthy worker effect is an excellent example of selective survival. Sick workers will seek a less hazardous job, while healthy workers will be more likely to stay in the same job. The healthy worker effect results in lower than expected morbidity and mortality rates among workers, so that even if a problem is present, it may not be detected by studies which are based on prevalence measures. In the situation of an existing selective survival, prevalence is not sufficient and cross-sectional studies will typically underestimate associations between study factor and outcome. Please note that cross-sectional studies are especially prone to the problem of selective survival. In case-control studies this type of selection bias can be avoided by only considering incident cases.

Information Bias

The second of the three main biases which we want to introduce is called information bias.

Box 8.8: Definition of information bias

Information bias refers to a distortion in the effect measure, due to measurement error or misclassification of participants for one or more variables. Information bias occurs when the measurement of either the study factor or the outcome is **systematically inaccurate**.
Because one or more variables are wrong, therefore the participants have been misclassified.
Source: Kleinbaum, D.G., et al., 1982.

Comment:
(1) Bias occurs when the effect measure, for example the relative risk in a cohort study, is distorted and does not give a reasonable estimate of the effect measure in the target population. Information bias occurs when this distortion is due to the way the study factor and/or outcome are assessed.

(2) Information collected for a participant might be incorrect. For example, when a person is asked about the number of cigarettes smoked on average during the past week (s)he might reply about 20 per day, while in reality the average per day was 35 cigarettes. This one person reporting incorrectly is not information bias – as yet! Information bias occurs only if there is a general tendency of all or most participants to underestimate their smoking habit, that is, a **systematic** underestimation.

There are many potential sources of information bias including a defective measurement device, questions in a questionnaire which do not measure what they are supposed to measure, an inaccurate diagnostic test, or an erroneous data source such as when participants recall exposure retrospectively. Think again about our example of recalling the number of cigarettes smoked during the past week. We all know that smoking is unhealthy and recently it has been made more and more socially unacceptable to smoke in public. Within this context it might be challenging for smokers to admit just how much they smoke, and when asked, smokers might be tempted to underestimate their cigarette consumption. Such a systematic underestimation can lead to information bias.

When we measure the strength of the association, for example between smoking cigarettes (= study factor) and coronary artery disease (= outcome), we often categorise the study factor and the outcome so we can calculate odds-ratio or relative risk. Any incorrect information will potentially lead to a *misclassification* of the participants. In the above example, for instance, if people underestimate cigarette consumption, some might say they are non-smokers while in actual fact they smoke (misclassified to be “unexposed”), others might say they smoke about 5 to 10 cigarettes per day while in actual fact they smoke 15 to 20, and so on. Thus, systematic error in the collected information will most likely lead to misclassification of the study factor and/or outcome; and this is the reason why information bias is also often called **misclassification bias**.

As with selection bias, in the following we will describe some common sources of information bias in epidemiological research using some classic and not so classic examples.

Information bias – Example 1: Misclassification of disease status

In epidemiological studies where we are following up people to detect new cases of a disease (= outcome), a diagnostic test is often used to assess the disease status. As discussed in Chapter 3, sensitivity and specificity of a diagnostic test define how well the test can identify whether someone has the disease or not. Since diagnostic tests are rarely 100% sensitive and 100% specific, the use of diagnostic tests inevitably produces false positive and false negative cases, that is, it introduces some misclassification of the disease status.

Box 8.9: Non-differential and differential misclassification

We distinguish in general between two types of misclassification: **non-differential and differential misclassification.**

In **non-differential** misclassification, the misclassification acts identically on exposed and unexposed or diseased and disease free.

In **differential** misclassification, the misclassification acts differently on diseased and disease free or on exposed and unexposed, or the misclassification occurs only with certain combinations of exposure and outcome.

Examples:

(1) The discussed misclassification caused by diagnostic testing is a typical example of a non-differential misclassification, as the misclassification acts identically on the exposed and unexposed people. Sensitivity and specificity are identical when applied to the exposed or to the unexposed.

(2) An example of differential misclassification often occurs in case-control studies and is called **recall bias**. People diagnosed with a serious disease (= cases) tend to reflect more carefully about potential causes of their illness. For example, melanoma cases are probably more likely to carefully consider how frequent and severe their life-time sunburns were and may thus recall severe sunburns more accurately than controls. This differential recall is an example of differential misclassification and may lead to an overestimation of the relationship between sunburn and melanoma. In the context of direction of bias we will call this a bias “away from the null value”. Please note that case-control studies are prone to recall bias.

Information bias – Example 2: Detection bias

We will now return to detection bias which was an example of selection bias in case-control studies. Here, however, we discuss detection bias as an example of information bias in cohort studies. Please do not get confused! The same mechanism is at work in both cohort and case-control study, but the type of bias differs because of the design. Above we discussed the classic example of oestrogen use and occurrence of endometrial cancer (Horwitz, R.I. and Feinstein, A.R., 1978) as an example of detection bias in a case-control study leading to selection bias, because the cases were identified more readily due to their exposure status. If we now assume we want to investigate the association between oestrogen use and endometrial cancer using a cohort design, then we are potentially faced with the same mechanism. Women who use oestrogen will probably keep in regular contact with their GP to renew their prescriptions. These "exposed" women are therefore more likely to be regularly examined by their GP and therefore any disease, including endometrial cancer, is more likely to be detected earlier in these women. Hence in cohort designs, detection bias is a form of differential misclassification bias.

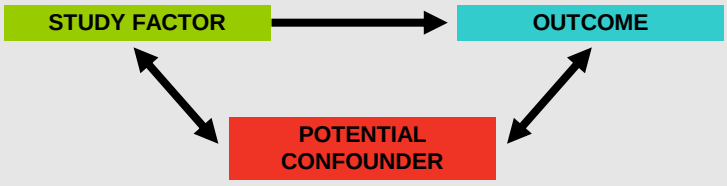
Confounding Bias

The last of the three main types of bias we will introduce here is confounding. Confounding is a type of bias that may occur when the effect of the study factor on the outcome is mixed in the data with the effect of another or third variable (= confounder). Confounding is a very common bias and can occur in all study designs.

Box 8.10: Potential confounder

A potential confounder is a variable that is:
(1) related to the study factor;
(2) related to the outcome (independently from it's relationship to the study factor); and
(3) not an intermediate variable between the study factor and the outcome.

Figure 6: Schematic representation of potential confounding



Example:
Earlier we introduced a study which investigated sexual violence in Australian prisons (study idea based on Richters, J. et al. 2010). We were interested in finding out whether sexual coercion in prisons occurs more frequently among homosexual rather than heterosexual male prisoners. In this context, the study factor is sexual orientation (homosexual or heterosexual) and the outcome is sexual coercion. Which

variables could be potential confounders?

Taking illegal drugs would be a potential confounder if (1) there was a relationship between drug taking and sexual orientation, that is, if say heterosexual men were more likely to use illegal drugs than homosexual men in this study group; and if (2) drug taking was related to sexual coercion, that is, if for example people taking illegal drugs in this study were more likely to be victims of sexual coercion than those not taking illegal drugs; and if (3) drug taking was not an intermediate variable, that is drug taking is not logically related to sexual orientation or to sexual coercion.

Comments:

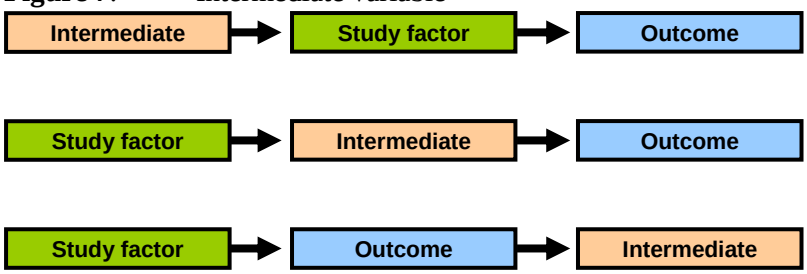
(1) Please note that all three conditions need to be fulfilled for a variable to be a potential confounder.
(2) A potential confounder *may* induce confounding bias. Whether it actually does is another question and can only be checked when the data are available. We might need to stratify the analysis by the potential confounder or conduct multivariable analysis to see whether there was confounding bias (see example in Box 9.11).
(3) From the example of sexual orientation and sexual coercion amongst an Australian prison population, it should be obvious that we do not always immediately know whether a characteristic could act as a potential confounder or not. Insider knowledge and careful analysis are often required. Always consider all known risk factors for the outcome as potential confounders to be on the safe side.
(4) The most important issue about confounding is to identify potential confounders prior to data collection so that information about these variables is at least recorded for subsequent analysis.

The question remains: **"What is an intermediate variable?"**

Intermediate variable

There are three situations when a variable is called an **intermediate variable**: (1) the variable is a logical pre-stage of the study factor; (2) the variable is a logical consequence of the study factor or a logical pre-stage of the outcome; or (3) the variable is a consequence of the outcome (Figure 7).

Figure 7: Intermediate variable



It is difficult to argue the concept of intermediate variable in general, because the approach taken to analyse the data is a function of the hypothesis being evaluated. In other words, whether or not a variable is considered a study factor or an intermediate variable is dependent on the conceptual model that is translated into study questions and research hypotheses. For example, if we want to investigate the association between sun exposure and skin cancer, we need to think about how to measure sun exposure. We could ask participants to estimate the number of previous sunburns. We could also ask them to estimate their sun exposure and sun protection habits over their life-time and combine the two questions or sets of questions to develop a sun exposure score. Both of these measures, sunburn and sun exposure score, are an attempt to measure a very similar effect "previous sun exposure". Which of them we consider the study factor and which the intermediate variable depends on our specific research hypothesis. Whichever we decide should be the study factor, the other characteristic should not be considered a potential confounder because it is too closely related to the study factor, exhibiting a very similar or identical effect on the outcome (skin cancer).

Intermediate variables very often describe altered physiology or pathological change that occurs as a result of the study factor. A variable that represents pathological change caused by the study factor could certainly be a risk factor for the outcome. It is also correlated with the study factor, as it is a direct result of the effect of the study factor. Nevertheless, it should not be considered a potential confounder as the effect of the study factor is mediated by the effect of the intermediate variable. Please note that when we consider an intermediate variable there is no mixing of effects. An intermediate variable is not an independent third variable because it does not introduce a separate effect. For example, if the intermediate variable is a direct logical consequence of the study factor there is only one effect on the outcome which is represented once by the study factor and a second time by the intermediate variable. Any variable that represents a logic step in the causal chain between the study factor and the outcome is an intermediate variable and not a potential confounder.

Sometimes we consider a study factor and also some of its components such as total alcohol consumption and beer consumption. If total alcohol consumption is our study factor, then beer consumption cannot be considered a potential confounder, because beer is part of the total alcohol consumption. Hence beer consumption does not have its own effect but is an intermediate variable of total alcohol consumption.

A different intermediate variable occurs when the intermediate variable logically determines the study factor. Consider a study in which we want to investigate the relationship between red meat consumption (= study factor) and colorectal cancer (= outcome). The characteristics "being a vegetarian" (no, yes) will be logically associated with the study factor as red meat consumption is "none" for all vegetarians. This kind of logic association is not confounding because there is no independent effect. "Being a vegetarian" is an intermediate variable.

Another type of intermediate variable can be identified in a study that investigates needle sharing behaviour (= study factor) and HIV sero-prevalence (= outcome) in intravenous drug users. A consequence of being HIV positive is persistent generalised lymphadenopathy. Persistent lymphadenopathy usually occurs in

response to a significant systemic disease including autoimmune diseases such as rheumatoid arthritis. Persistent lymphadenopathy is more likely to occur in HIV positive cases and those who more often shared needles - and who are in turn more likely to be HIV positive. That is, persistent lymphadenopathy is correlated with the study factor and the outcome, but is a direct consequence of HIV sero-positivity and therefore in this context an intermediate variable.

The problem with intermediate variables is that they are too closely related to either the study factor or the outcome. If we want to assess the relationship between study factor and outcome statistically we will have difficulties adjusting for an intermediate variable. Please note that special considerations are required to include an intermediate variable in an analysis.

Potential confounding

In most epidemiological studies age and gender are the first variables we should consider as potential confounders. One can single out age and gender as the two personal characteristics that regularly affect disease occurrences with female gender and younger age often correlated with more positive outcomes. Thus, age and gender are frequently associated with the outcome and should therefore always be considered as potential confounders.

Box 8.11: Potential confounding, or?

Assume we want to discuss potential confounding in a study which investigates the association between total alcohol consumption of pregnant women and birth weight of the newborn. Remember from the definition, that a variable has to be associated with the study factor and the outcome and not be intermediate to be called a potential confounder.

(1) Is the mother's religion a potential confounder?
No: Religion is not related to the birth weight of a child. Religion might be related to alcohol consumption, as some religions forbid the consumption of alcohol. However, in this case, alcohol consumption is logically related to religion and thus can be regarded as an intermediate variable but not as a potential confounder.

(2) Is the mother's height a potential confounder?
No: The height of the mother is probably correlated with birth weight of the child; tall mothers may have taller and therefore heavier babies. However, the height of the mother is not related to alcohol consumption. Therefore, height of the mother is a "risk factor" for birth weight but it should not introduce confounding bias.

(3) Is the mother's hair colour a potential confounder?
No: Hair colour is neither associated with alcohol consumption of the mother, nor with the birth weight of the child.

(4) Is the mother's smoking habit a potential confounder?
Yes: Individuals who smoke also tend to drink more alcohol, or people who drink alcohol are also more likely to smoke. We know that smokers exhibit higher "risk taking behaviour" (Jenks, R.J., 2001). It is also known that maternal smoking is a risk factor for low birth weight (Australian Government Department of Health and Ageing, 2006). In addition, smoking is not an intermediate variable, that is, smoking is not logically preceding alcohol consumption, neither is it a *logical* consequence of alcohol consumption and smoking is also not a logical pre-requisite or consequence of having a baby with low birth weight. Therefore, the smoking habit of the mother is a potential confounder in this context.

(5) Is the mother's wine consumption potential confounder?
No: Wine consumption is part of the mother's total alcohol consumption. Hence wine consumption will not introduce a separate effect but wine consumption and total alcohol consumption will have the same effect (not in size maybe but in principle) on birth weight. Wine consumption is an intermediate variable.

(6) Is the mother's age a potential confounder?
Yes: Age and smoking habits are related. Recent Australian statistics show that people aged 25 to 29 years are more likely to smoke than older people (Scollo, M.M. and Winstanley, M.H., 2008). Age of the mother and birth weight of the newborn are also related (Australian Bureau of Statistics, 2007). Women under the age of 19 years and older than 40 years are more likely to give birth to under weight babies. Age is not logically related to either alcohol consumption or birth weight and therefore not an intermediate variable. Hence the mother's age is a potential confounder.

The example in Box 8.11 demonstrates that an "outsider" might not always know about specific associations between the study factor, the outcome and potential confounders. If we worked in maternal and child health, we would probably know these associations and therefore the discussion for or against potential confounding would be easier.

Please note that in contrast to selection bias and information bias, it is standard practice to control for confounding bias during data analysis. However, this can only be accomplished if the required information is available. The importance of identifying any potential confounding variables prior to undertaking the study and recording the relevant information cannot be overstated!

We have already mentioned that one of the three potential confounders, identified via our three criteria, does not necessarily induce genuine confounding bias. The actual presence of confounding bias can only be assessed during data analysis. The most frequent ways to deal with confounding are multivariable statistical analysis and stratified analysis.

In multivariable statistical analysis confounding bias is identified by comparing the crude effect measure, such as the odds-ratio, with the adjusted effect measure which corrects for the potential distortion by the confounder. Confounding bias occurs when the crude and adjusted effect measures differ in value. Please note that the crude

and adjusted effect measures are almost always slightly different; we need to decide how much they can differ before we call the variable a confounder. Often, a difference of more than 5% or 10% will define confounding.

Deciding whether or not a variable is a confounder should not be based on a statistical test. The issue of confounding is an issue of systematic rather than random error and thus should not be based on statistical significance testing. Even if a distortion caused by a confounder is not statistically significant, it still needs to be controlled for if the distortion is reasonably large. Multivariable analysis is widely used in epidemiology as it allows us to adjust simultaneously for the effects of different confounders.

Control of Bias

Bias is a central concern in epidemiological research. Researchers aim to gain valid results and, therefore, want to avoid or at least minimize bias. In general, bias can be controlled *a priori* by using a well-planned study design followed by meticulous attention to detail while conducting the study or *a posteriori* during statistical analyses.

If appropriate information is available (!) it is possible to numerically adjust for selection, information and confounding bias during the analysis phase of the study. However, this *a posteriori* approach is only standard procedure for confounding bias. Often additional studies would be required to gather the information necessary to adjust for selection and information bias. Consequently it is preferable to avoid or minimize these types of biases using a carefully planned study design and being fastidious when conducting the study. It is necessary to record information about all potential confounders to be able to numerically adjust for confounding bias.

Control of bias during design phase of study

If we want to control bias during the design phase of our study we first have to identify potential sources of bias in the specific study and its design. This requires us to have a sound understanding of the subject matter we are investigating. In addition, common sense and some expertise and experience in epidemiological study designs are needed. This experience cannot be learned from a textbook! However, the different study designs have their specific weak points and when you plan a study, control of bias could start by targeting the most likely sources of bias. In the following section, we have provided a brief summary of the most common types of bias in the different basic study designs and ways to avoid them.

- **Experiments**

Selection bias: The source of the participants, setting and location, are important to consider. For example, if patients are recruited from an hospital, are these patients representative of all patients with the disease in this location, or only of more severe cases? Inclusion and exclusion criteria for participants need to be carefully stated as they define the target population. A high response rate and complete follow-up are crucial.

Information bias: Blinding and placebo controls should be used whenever possible and ethically justifiable to achieve optimal observational uniformity. Always consider blinding the outcome assessor.

Confounding: Randomisation should be used if possible and ethically justifiable, to maximize structural uniformity. Remember that randomisation potentially controls for known and unknown confounders. It is very important to check whether the randomisation for known confounders was actually successful! It is essential to collect data about all known potential confounders, so that we can check the success of randomisation and/or are able to control confounding bias by stratification or multivariable analysis.

- **Cohort studies**

Selection bias: The source of the participants, setting and location, should be considered. For example, if participants are recruited from a community, are these participants representative of all people in this location? Inclusion and exclusion criteria for participants need to be carefully reported as they define the target population.
A high response rate and complete follow-up are crucial.

Information bias: In a prospective cohort study, information bias should only be a minor issue as investigators can control the quality of the collected information.
In a retrospective cohort study, information bias can be a major concern because the quality of previously recorded data cannot be directly controlled by the researchers. Recall bias may also be a problem in this context. Retrospective cohort studies are often occupational studies using company records for individual exposure data and the quality of these records is of concern.

Confounding: It is essential to collect data about all potential confounders so they can be used later during statistical analysis (controlling). Matching, that is finding a similar unexposed person for each exposed person, could be applied but is in reality not frequently used in cohort designs.

- **Case-control studies**

Selection bias: The source of the cases, setting and location, should be considered. For example, if patients are recruited from an hospital, are these patients representative of all patients with the disease in this location, or only of the more severe cases? Inclusion and exclusion criteria for cases need to be carefully reported as they define the target population of cases.
The source of the controls, setting and location, are also very important. For example, controls should represent the

community from which the cases are recruited. So if cases are recruited from an hospital, the controls should be recruited from the community which is served by this hospital. Population-based controls are preferred over hospital-based controls because of Berkson's fallacy (a special form of selection bias). If cases are difficult to recruit, two or more controls per case increases the statistical power.
Inclusion and exclusion criteria for cases and controls need to be carefully reported as they define the target population.
A high response rate is crucial.

Information bias: Case-control studies rely on recall of exposure, for example, by questioning people or retrieving records, thus information bias is a critical issue.

Confounding: It is essential to collect data about all potential confounders for later use during statistical analysis (controlling). Matching is often used in case-control studies to reduce the effect of strong known potential confounders. Individual matching is the best method. Matching tries to produce similar outcome groups, that is cases and controls, only! We may still need to control for matching variables during statistical analysis when we compare the exposed group with the unexposed group.

- **Cross-sectional studies**

Selection bias: The most critical issue for a cross-sectional study is that the sample is a probability sample of the target population. This means that the probability of being selected into the sample is known for each individual of the target population. Unfortunately, probability samples such as random samples are rather difficult to accomplish in the real world.
A high response rate is crucial.

Information bias: If a cross-sectional study relies on recall of previous exposures, information bias is likely. If exposure and outcome are assessed at the time of the study, information bias may be a minor issue.

Confounding: It is essential to collect data on all known potential confounders for later use during statistical analysis (adjusting). Alternatively, we can opt for stratified sampling, stratifying according to one or two known strong confounders. However, this latter approach effectively leads to numerous studies, as a research hypothesis and a sample size calculation are required for each stratum.

From the above summary it becomes clear that an experimental design generally delivers the most powerful tools to control for or even avoid bias. While selection and confounding bias can be handled by thorough planning and statistical analysis,

information bias seems to remain an unresolved problem in observational and, to a lesser extent, experimental designs.

Control of bias during statistical analysis

Selection bias can only be controlled for during statistical analysis when there is additional information about the selection probabilities. For example, if a case-control study was conducted in an hospital setting and hospitalisation rates for cases and controls are known, the effect measure can be adjusted accordingly. Similarly, misclassification of the disease status can be corrected, if the sensitivity and specificity of the diagnostic procedure are known. However, both adjustments are rarely performed as the prerequisite data are usually not available. For details please see refer to Lash, T.L. et al. (2009).

In contrast confounding bias is routinely controlled for using statistical analysis. Confounding bias caused by known confounders can be controlled for during statistical analysis if information about these confounders has been collected. The simplest way to control for confounding is by calculating the effect measure in the strata of the confounder. However, stratification may only be feasible if the confounder does not have too many categories and if there are not too many confounders. Another method of controlling for confounding effects is standardization, but as with stratification, standardization can only be undertaken for one or two confounders because of sample size restrictions. Nowadays, the method of choice used to control for confounding bias is multivariable statistical analysis. Multivariable statistical analysis allows you to control for both categorical and numerical confounders and for several confounders simultaneously. Alvan Feinstein (1996) has written a thorough introduction to the world of multivariable analysis, which we recommend to the interested reader; the book avoids mathematics where possible.

Direction of Bias

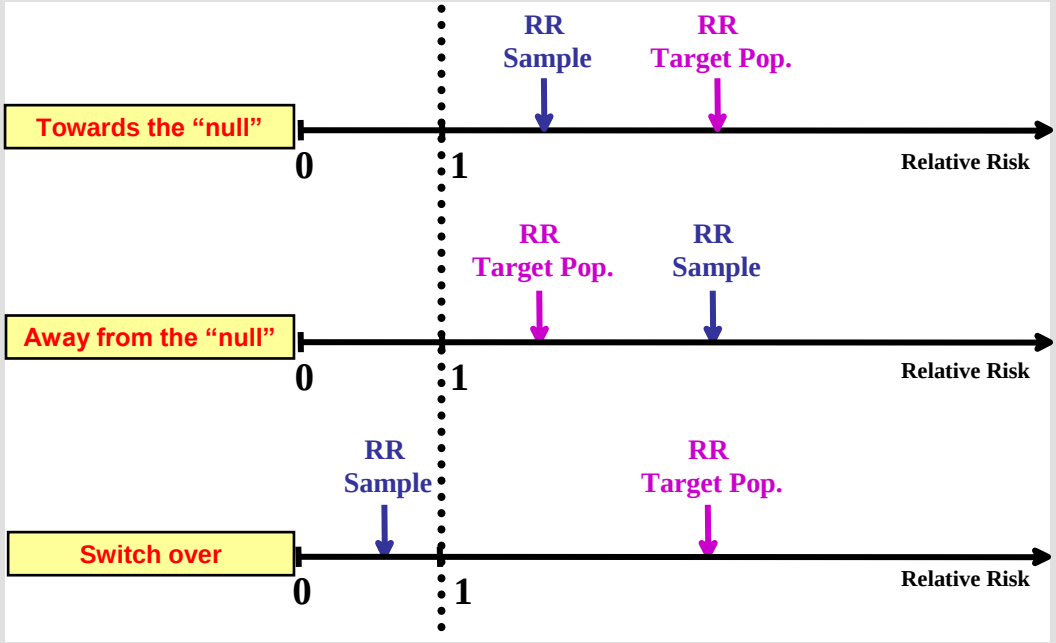
Let us emphasis again that it is very difficult to plan and conduct an epidemiological study that is completely free of bias. Therefore, when reading a publication which describes the results of an epidemiological study we need to be alert for possible biases. For instance, we need to carefully consider whether: the participants in the study are representative of the intended target population; all the information was appropriately collected; and all potential confounders were dealt with either in the design or the analysis phase of the study. When we identify a potential bias in a published study, the next step is to try to find out what effect this bias is likely have had on the results. This leads us to the last topic of this chapter, the “direction of a bias”.

Box 8.15: Direction of bias

The **direction of bias** refers to the deviation of the estimated effect measure in the sample from the effect measure in the target population caused by bias. The direction of bias can be **towards the null-value**, **away from the null-value**, or

switch over (Figure 8).

Figure 8: Direction of bias using relative risk as the effect measure.



A bias is called **towards the null-value**, if the estimate of the effect measure of the sample is closer to the null-value than the true effect measure in the target population. Bias towards the null implies that we are under-estimating the true effect measure in the target population.

A bias is called **away from the null-value**, if the estimate of the effect measure of the sample is further away from the null-value than the true effect measure in the target population. Bias away from the null means that we are over-estimating the true effect measure in the target population.

A bias causes a **switch-over**, if the estimate of the effect measure of the sample is one side of the null-value and the true effect measure of the target population is on the other side.

Please note: if the effect measure is a ratio, such as relative risk or odds ratio, then the null value is 1. If the effect measure is a difference, such as attributable risk, then the null value is 0.

Comments:
(1) The examples in Figure 8 refer to "risk" associations. But the statements above apply similarly to "protective" effects.

(2) Direction of bias is a theoretical issue and requires us to think logically about a situation and then consider possible scenarios and choose the most likely. Thus we might come up with different scenarios leading to different directions of the bias.

(3) It is often impossible to even be sure about the direction of a bias (let alone its size) simply because the effect measure in the target population is unknown. The

situation becomes especially complex when several potential biases are identified which may lead in different directions.

Summary

- When conducting an epidemiological study, if a systematic error occurs which leads to misinterpretation of the effect measure (for example, relative risk, exposure odds-ratio, or attributable rate) this misinterpretation is called bias. If there is no misinterpretation the effect measure is called valid (or unbiased).
- There are three main types of bias: selection bias, information bias, and confounding bias.
- Selection bias refers to a distortion in the effect measure, resulting from the manner in which the people are selected for the sample. Thus, selection bias may be introduced if sampling techniques are inappropriate. If selection bias occurs, the sample(s) do not represent the target population.
- Information bias refers to a distortion in the effect measure, which is due to measurement error or misclassification of participants for one or more variables. Information bias occurs when the measurement of either the study factor or the outcome is systematically inaccurate.
- We distinguish between two types of misclassification: non-differential misclassification and differential misclassification. In non-differential misclassification, the misclassification acts identically on exposed and unexposed groups or diseased and disease free groups. In differential misclassification, the misclassification acts differently on diseased and disease free groups or on exposed and unexposed groups, or the misclassification occurs only with certain combinations of the exposure and the outcome.
- Confounding bias is a type of bias that may occur when the effect that the study factor has on the outcome is mixed in the data with the effect of a third variable (= confounder). A potential confounder has to be (1) related to the study factor, (2) related to the outcome, and (3) not an intermediate variable.
- Confounding bias means that the stratum-specific estimates are different from the overall crude estimate of the effect measure. A confounder is called an effect modifier, if the stratum-specific estimates of the effect measure are different from each other.
- A bias is called towards the null-value, if the estimate of the effect measure of the sample is closer to the null-value than the true effect measure in the target population. Bias towards the null means an underestimation of the true effect measure in the target population. A bias is called away from the null-value, if the estimate of the effect measure from the sample is further away from the null-value than the true effect measure in the target population. Bias away from the null means an overestimation of the true effect measure in the target population. A bias causes a switch-over if the estimate of the effect measure of the sample is one side of the null-value and the true effect measure of the target population is on the other side.
- A non-differential misclassification bias is always towards the null-value.

References

Australian Bureau of Statistics (ABS) (2007) Australian social trends. Australia's babies: www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/04FEBEF9C81FE6BACA25732C002077A2?opendocument.

Australian Government Department of Health and Ageing (2006) Smoking harms unborn babies: www.health.gov.au/internet/quitnow/publishing.nsf/Content/warnings-b-babies.

Caldwell, G.G., Kelley, D.B., Heath, C.W. Jr (1980) Leukemia among participants in military maneuvers at a nuclear bomb test: A preliminary report. JAMA 244(14):1575–8.

Feinstein, A.R. (1985) Clinical epidemiology: The architecture of clinical research. W.B. Saunders Co., Philadelphia.

Feinstein A.R. (1996) Multivariable analysis. An introduction. Yale University Press, New Haven.

Greenland, S. and Neutra, R. (1981) An analysis of detection bias and proposed corrections in the study of estrogens and endometrial cancer. J Chronic Dis 34(9–10):433–8.

Horwitz, R.I and Feinstein, A.R. (1978) Alternative analytic methods for case-control studies of estrogens and endometrial cancer. N Engl J Med 299:1089–94.

Hutchison, G.B. and Rothman, K.J. (1978) Correcting a bias? N Engl J Med 299:1129–30.

Jenks, R.J. (2001) Attitudes, perceptions and risk-talking behaviors of smokers, ex-smokers, and non-smokers. J Social Psychol 132(5):569–75.

Jones, M., Onslow, M., Packman, A., Williams, S., Ormond, T., Schwarz, I., Gebski, V. (2005) Randomised controlled trial of the Lidcombe programme of early stuttering intervention. Br Med J 331(7518):659.

Kleinbaum, D.G., Kupper, L.L., Morgenstern, H. (1982) Epidemiologic research: Principles and quantitative methods. John Wiley & Sons, New York.

Lash, T.L., Fox, M.P., Fink, A.K. (2009) Applying quantitative bias analysis to epidemiologic data. Springer Publishers, New York.

McMichael, A.J., Gerber, W.S., Gamble, J.F., Lednar, W.M. (1976) Chronic respiratory symptoms and job type within the rubber industry. J Occup Med 18(9):611–17.

Richters, J., Butler, T., Schneider, K., Yap, L., Kirkwood, K., Grant, L., Richards, A., Smith, A.M., Donovan, B. (2009) Consensual sex between men and sexual violence in Australian prisons. Arch Sex Behav 41(2):517–24.

Rothman, K.J., Greenland, S., Lash, T.L. (2008) Modern epidemiology. 3rd edn. Lippincott Williams & Wilkins, Philadelphia.

Scollo, M.M. and Winstanley, M.H. (eds) (2008) Tobacco in Australia: Facts and issues. 3rd edn. Cancer Council Victoria, Melbourne.

Spilker, B.L. (1991) Guide to clinical trials. Raven Press, New York.