CHAPTER 3
Bias in randomized controlled trials

The main appeal of the randomized controlled trial (RCT) in health care comes from its potential to reduce selection bias. Randomization, if done properly, can keep study groups as similar as possible at the outset, so that the investigators can isolate and quantify the effect of the interventions they are studying. No other study design gives us the power to balance unknown prognostic factors at baseline. Random allocation does not, however, protect RCTs against other types of bias.

The existence of most biases related to RCTs is supported mainly by common sense. In recent years, however, important research efforts have used RCTs as the subject rather than the tool of research. These studies are usually designed to generate empirical evidence to improve the design, reporting, dissemination, and use of RCTs in health care.[1] They have confirmed that RCTs are vulnerable to many types of bias throughout their entire life span. Random allocation of the participants to different study groups increases the potential of a study to be free of allocation bias, but has no effect on other important biases.

In this chapter we will discuss the concept of bias in relation to RCTs and highlight some of its sources. We will also list a variety of biases, as well as some strategies that may help us recognize them and minimize their impact on the planning of research and health-related decisions.

What is bias?

An online dictionary[2] defines 'bias' as '*a partiality that prevents objective consideration of an issue or situation*'. In statistics it means '*a tendency of an estimate to deviate in one direction from a true value.*[3] This systematic deviation from the true value can result in either underestimation

or overestimation of the effects of an intervention. Because there is usually more interest in showing that a new intervention works than in showing that it does not work, biases in clinical trials most often lead to an exaggeration in the magnitude or importance of the effects of new interventions.

We should not jump to the conclusion that bias in health research is necessarily associated with a conscious or malicious attempt of investigators, funders, or readers to bend the results of a trial. Indeed, although bias may be introduced into a trial intentionally, it is probably more commonly unintentional, and often unrecognized even by the researchers themselves.

Why does bias in an RCT matter?

The true effects of any health care intervention are unknown. We try to anticipate, detect, quantify, and control bias to produce results from a sample of participants that can be generalized to the target population at large. It is impossible to ever know for sure whether the results of a particular study are biased, simply because it is impossible to establish whether those results depart systematically from a 'truth' that remains unknown.

What are the main types of bias in RCTs?

Most discussions on bias focus on biases that can occur during the actual course of a trial, from the allocation of participants to study groups, through the delivery of interventions, to the measurement of outcomes. Other types of bias can arise, however, even before the trial is carried out, in the choice of problem to study or type of research to use, or after the trial is carried out, in its analysis, and its dissemination. Bias can even be introduced by the person who is reading the report of a trial.[4] These biases, which can also have a profound influence on the way in which the results of RCTs are interpreted and used, tend to receive less attention.

To illustrate how biases can affect the results of an RCT, we invite you to think about the following hypothetical scenario:

> *'Imagine a new drug for the treatment of multiple sclerosis, which has shown promising results in animal studies and in phase I trials. These results, which suggest that the drug can delay the onset of severe motor*

compromise, have been widely publicized by the media during the past 3 months. Because of these results, patient advocacy groups are putting pressure on the government to make the new drug available as soon as possible. As multiple sclerosis is a debilitating disease that affects millions of people worldwide and for which there is no known cure, the investigators (all clinicians who have dealt with multiple sclerosis patients for years), the company producing the new drug (which has invested millions in developing the drug), the media (interested in confirming the results that they so widely publicized) and the potential participants (patients with multiple sclerosis who have been waiting for an effective treatment to be discovered) are all interested in demonstrating that the new compound is effective. After many intense sessions debating the course of action, a multidisciplinary task force created by the government, including consumer representatives, agrees that the next step should be a randomized clinical trial. A research protocol is produced by another multidisciplinary panel of investigators and consumers, and a well known research group at a large health care center is selected to conduct the study.'

We discuss the elements of this hypothetical scenario in the following sections.

Selection bias

With true randomization, all participants in the study are given the same opportunity to be allocated or assigned to each of the study groups. But even a perfectly randomized method to allocate participants to the study groups does not protect against selection bias, which can occur both in the way that individuals are accepted or rejected for participation in a trial, and in the way that the interventions are assigned to individuals once they have been accepted into a trial.

Selection bias can occur if some potentially eligible individuals are selectively excluded from the study, because the investigator knows the group to which they would be allocated if they participated. Let us suppose that the investigator in charge of recruiting patients for the multiple sclerosis trial (who at least subconsciously hopes that the drug will be found to be effective) thinks that depressed patients are less likely to respond to the new drug. If he has access to the allocation sequence (which has been generated by computer and is locked in his desk) this investigator could introduce bias into the trial by making it more difficult for depressive

patients to receive the new drug. He could, knowingly or unknowingly, exclude depressive patients who would be allocated to receive the new drug by making them fit the exclusion criteria more easily than if they had been allocated to the placebo group. He could also (again knowingly or unknowingly) present information on the trial to depressive patients allocated to receive the new drug in such a way that they would be discouraged from consenting to participate. At the end of the trial, if the investigator was right, and depressive patients were in truth less likely to respond to the new drug, the trial will show an exaggerated effect of the new drug during the treatment of multiple sclerosis, due to the disproportionate number of depressive patients in the placebo group.

How can selection bias be reduced?

There is empirical evidence to show that effects of new interventions can be exaggerated if the randomization sequence is not concealed from the investigators at the time of obtaining consent from prospective trial participants.[5] One study showed that trials with inadequate allocation concealment can exaggerate the estimate of the effect size of interventions by as much as 40% on average.[6] The irony is that *allocation concealment* is a very simple maneuver that can be incorporated in the design of any trial and that can always be implemented.

Despite its simplicity as a maneuver and its importance to reduce bias, allocation concealment is rarely reported, and perhaps rarely implemented in RCTs. Allocation concealment was reported in less than 10% of articles describing RCTs published in prominent journals in five different languages.[7] This does not necessarily mean that allocation is not concealed in 90% of RCTs; in some cases, allocation may have been concealed, but the authors, peer-reviewers, and journal editors were not aware of how important it is to mention it (it takes about a line in the report, so space limitation is not a good excuse). If, however, allocation concealment was not carried out in most cases in which it was not reported, the majority of RCTs are at risk of exaggerating the effects of the interventions they were designed to evaluate.

Even if the report of an RCT states that efforts were made to conceal the allocation sequence, there are many ways in which randomization can be subverted by investigators who want to break the allocation code before they obtain consent from prospective trial participants.[8] Even when the allocation codes are kept

in sealed opaque envelopes, for instance, investigators can (and sometimes do) look through the envelopes using powerful lights or even open the envelope using steam and reseal it without others noticing. Thus it is very easy to introduce selection bias into RCTs.

Users of RCTs should not get a false sense of security just because a study is randomized.

Ascertainment bias

Ascertainment bias occurs when the results or conclusions of a trial are systematically distorted by knowledge of which intervention each participant is receiving. Ascertainment bias can be introduced by the person administering the interventions, the person receiving the interventions (the participants), the investigator assessing or analyzing the outcomes, and even by the people who write the report describing the trial (Chapter 2).

The best way to protect a trial against ascertainment bias is by keeping the people involved in the trial unaware of the identity of the interventions for as long as possible. This is called blinding or masking. The strategies that can be used to reduce ascertainment bias can be applied during at least two periods of a trial: the time during which data are collected actively (from the administration of the interventions to the gathering of outcome data) and after data have been collected (from data analysis to the reporting of results).

It is important to recognize the difference between biases that are the result of lack of allocation concealment and biases that arise from lack of blinding. *Allocation concealment* helps to prevent selection bias, protects the randomization sequence *before* and *until* the interventions are given to study participants, and can *always* be implemented. *Blinding* helps prevent ascertainment bias, protects the randomization sequence *after* allocation, and cannot always be implemented.[6]

How can ascertainment bias be reduced during data collection?

Ascertainment bias can be introduced in different ways during data collection. For instance, the people administering the interventions can bias the results of a trial by altering systematically the co-interventions given to participants during the trial. Following our

example of the multiple sclerosis trial, the new drug may appear to be more effective at the end of the trial if patients allocated to the new drug received physiotherapy earlier and more intensively than patients allocated to placebo (*co-intervention bias*). If participants know that they have been allocated to the placebo group, they are likely to feel disappointed and less willing to report improvement at each of the study time points (*participant ascertainment bias*). In addition, if the people in charge of assessing and recording the outcomes know which patients are allocated to each of the study groups, they could, consciously or unconsciously, tend to record the outcomes for patients receiving the new drug in a more favorable way than for patients receiving placebo (*observer bias*).

In ideal circumstances, ascertainment bias should be reduced by blinding all concerned: the individuals who administer the interventions, the participants who receive the interventions and the individuals in charge of assessing and recording the outcomes (Chapter 2).

The importance of blinding has been confirmed in empirical studies. It has been shown, for instance, that open studies are more likely to favor experimental interventions over the controls[9] and that studies that are not double-blinded can exaggerate effect estimates by 17%.[6] Despite the empirical evidence available, and common sense, only about half of the trials that could be double-blinded actually were.[10] Even when the trials are described as double-blind, most reports do not provide adequate information on how blinding was achieved or statements on the perceived success (or failure) of double-blinding efforts.[11,12]

The best strategy to achieve blinding during data collection is with the use of placebos. *Placebos* are interventions believed to be inactive, but otherwise identical to the experimental intervention in all aspects other than the postulated specific effect. Placebos are certainly easier to develop and implement successfully in drug trials, in which they should resemble the taste, smell and appearance of the active drug, and should be given using an identical procedure.

Placebo controls can also be used with non-drug interventions, such as psychological, physical, and surgical procedures, although they are more difficult to develop and implement successfully. For example, it is difficult, but not impossible to develop and implement placebo counseling, physiotherapy, acupuncture or electrical stimulation. In some cases it is impossible, unfeasible or simply

unethical to use placebos. It would be impossible, for example, to use a placebo intervention in a trial evaluating the effect on mothers and newborns of early versus late discharge from hospital after childbirth. It would be unfeasible or unethical to use a placebo in trials evaluating new or existing surgical interventions (although a strong case can still be made for trials in which sham surgery can successfully challenge the perceived effectiveness of surgical interventions).[13] Placebo controlled studies are not ethical to study a new or existing intervention when there is an effective intervention available (Chapter 8). Even in cases where the use of placebos is impossible, unfeasible or unethical, trials can be at least single blind. In a surgical or acupuncture trial, for instance, single-blinding can be achieved by keeping the investigators in charge of assessing the outcomes unaware of which participants receive which interventions.

How can ascertainment bias be reduced after data collection?

Ascertainment bias can be introduced easily after data collection, if the investigators in charge of analyzing or reporting the results of the trial are aware of which participants are receiving which interventions. The effects of a new intervention can be exaggerated, for instance, if the investigators in charge of analyzing the trial data select the outcomes and the time points that show maximum benefit from the new intervention and ignore outcomes and time points that show either no effect or harm from the new intervention. Similarly, investigators in charge of reporting the trial results can choose to emphasize the outcomes and time points that show the maximum effects of the new intervention, downplaying or ignoring findings that suggest that the new intervention is equivalent or less effective than the control.

This source of bias can be controlled by keeping the data analysts and the people in charge of reporting the trial results unaware of the identity of the study groups. In a study with two groups, for instance, the outcome data could be given to analysts coded as A and B, and once they complete the analysis, the results could be given to the person in charge of writing the report using the same codes. The codes would not be broken until after the data analysis and reporting phases were completed. These valuable strategies should be used and studied more often.

Other important sources of bias

What biases can occur during the planning phase of an RCT?

Choice-of-question bias

Perhaps one of the least recognized forms of bias in an RCT is hidden in the choice of the question that the trial intends to answer. This would not necessarily affect the internal validity of a trial, but may have profound effects on its external validity, or generalizability. This bias can take many forms.

Hidden agenda bias occurs when a trial is mounted, not in order to answer a question, but in order to demonstrate a pre-required answer. The unspoken converse may be 'Don't do a trial if it won't show you what you want to find'. This could be called the *vested interest bias*.[14] Closely related to this is the *self fulfiling prophecy bias*, in which the very carrying out of a trial ensures the desired result.

The *cost and convenience bias* can seriously compromise what we choose to study. When we study what we can afford to study, or what is convenient to study, rather than what we really want to study, or should study, we take resources away from what we know is important. Closely related to this is the *funding availability bias* where studies tend to concentrate on questions that are more readily fundable, often for a vested or commercial interest. We should always look for the *secondary gains search bias* which can influence the choice of study, the methodology used, and the ascertainment and dissemination of the results.

Regulation bias

This is sometimes referred to as the *IRB bias* or the *Bureaucracy bias*. It occurs when institutional review boards are overly restrictive, and block the study of important questions. It also occurs when they are overly permissive and allow or even encourage studies that may not be scientifically or socially valid, but may bring either funding or prestige to the institution. Complicated 'informed consent' regulations may block the participation of many otherwise eligible subjects, and hence bias the results (Chapter 8).

Wrong design bias

The perceived value of an RCT may sometimes induce researchers to use this design for questions that may be better (or can only be

answered) with a different design, such as outcome research.[15] The wrong research design can produce misleading answers.

What biases can occur during the course of an RCT?

Population choice bias

The sample population studied can have a major effect on the generalizability of an RCT. If the sample is overly restricted by not including women (*gender bias*) or people over (or under) a specific age group (*age bias*), the results may not be generalizable to people who do not belong to the groups. *Pregnancy bias*, (excluding pregnant women) may sometimes be necessary for reasons of safety to the fetus, but the exclusion must be carefully noted. The same reasoning is required when trials are restricted to, or exclude, people in special circumstances (*special circumstances bia*s).

Population choice may be restricted when potential participants are approached (*recruitment bias*) or during registration of participants. Eligible patients may be kept out of a trial because they do not understand the consent form (*informed consent bias, literacy bias, language bias*).

Severity of illness bias is an important subgroup of the sample choice bias. Patients with a mild form of an illness may not respond in the same way as those with a more severe form.

Intervention choice bias

The nature of the intervention chosen can have a major effect on the results obtained. The stage at which an intervention is studied can be very important. The *too early bias* and the *too late bias* can determine the effects found.[16] This holds particularly true for surgical trials where there can be a *learning curve bias* for new operators, or improvements (or regression) in the techniques or contexts in which they are used. Similar concerns may hold for medical interventions, when dose or timing of a medication may be important determinants of the outcome.

Complexity bias can occur when a trial is used to study complex interventions, with a number of components, or where outcomes may depend on multiple contingencies outside of the control of the investigator (e.g. the skill of the surgeons or the resources of the community).[17]

Comparison choice (or control group) bias

If an intervention is compared to a poorly chosen control group, it can erroneously appear to be more (or less) effective than it really is. If a study compares an experimental intervention with a placebo control, the results will only tell us whether the intervention has a specific effect or not. It will not imply that the experimental intervention has a different or better effect than existing alternatives. An obvious way to make an intervention appear to be more effective than it really is would be to choose an ineffective comparison group.

Unfortunately, current regulatory bodies that mandate placebo controls lead to carrying out studies with this limited clinical value.

Outcome choice bias

Sometimes RCTs evaluate outcomes that are easy to measure, rather than the outcomes that are relevant (*measurement bias*). One variant of this is the *time term bias* in which short-term outcomes are measured rather than the important long-term outcomes. It is not surprising that researchers sometimes yield to the temptation to study outcomes that are readily measured rather than those that are important.

What biases can occur during the reporting of a trial?

Withdrawal bias: bias introduced by inappropriate handling of withdrawals, drop outs, and protocol violations

Ideally, all participants in a trial should complete the study, follow the protocol, and provide data on all the outcomes of interest at all time points. In reality, however, most trials have missing data. Data can be missing because some of the participants drop out before the end of the trial, because participants do not follow the protocol either deliberately or accidentally, or because some outcomes are not measured correctly or cannot be measured at all at one or more time points.

Regardless of the cause, inappropriate handling of the missing information can lead to bias. For instance, if in the multiple sclerosis trial patients who do not obtain benefit from the new drug withdraw more frequently because of adverse effects, their exclusion from analysis would lead the investigators to exaggerate the benefit and underestimate the harm of the new drug. This bias can

occur independently of whether or not the investigators are aware of the identity of the interventions received by the participants. If the decisions on withdrawals have been made because of knowledge of the interventions received by the participants, this constitutes yet another cause of ascertainment bias.

On occasion, it is impossible to know the status of participants at the times when the missing information should have been collected. This could happen, for example, if participants move to different areas during the study or fail to contact the investigators for an unknown reason. If the reasons for excluding these participants or specific outcome measurements from the final analysis were in any way related to the intervention, this could also lead to bias.

There are two strategies that can confidently be assumed to eliminate bias in these circumstances. One is known as *intention-to-treat analysis*, which means that all the study participants are included in the analyses as part of the groups to which they were randomized, regardless of whether they completed the study or not. The second method is a *worst-case scenario* or *sensitivity analysis*. This is performed by assigning the worst possible outcomes to the missing patients or time points in the group that shows the best results, and the best possible outcomes to the missing patients or timepoints in the group with the worst results. We can then see whether the new analysis contradicts or supports the results of the initial analysis excluding the missing data.

Selective reporting bias

A major and common source of bias in an RCT is selective reporting of results, describing those outcomes with positive results, or which favor the studied intervention. This is not always consciously done. The investigator may even unconsciously be attracted more to certain outcomes than others. Variants of this have been named the *social desirability bias* in which the items that are desired, or the *optimism bias* in which the items hoped for, are more likely to be reported.

The *data dredging bias* is another variant of the selective reporting bias. Having looked at all the data, the investigators can report the outcomes they wish to stress, and not mention the less desirable outcomes. A variant is the *interesting data bias*, in which the authors report the data that they find most interesting. The acme of data dredging can be in the selective analysis of data. If unethically contrived, all trials can be made to appear to have positive results.[18]

Fraud bias

Intentional fraud is perhaps the most important, most serious, and most difficult to detect source of bias. We hope that it is rare, but the extent to which fraudulent results are reported may be underestimated, and may be increasing under the pressure to produce results, no matter how.

What biases can occur during the dissemination of the trials?

What is publication bias?

Investigators and sponsors are more likely to write and submit, and peer-reviewers and editors to accept, manuscripts with positive results for publication. This tendency has been called *publication bias*.[19,20] A systematic review of five empirical methodological studies published mostly during the previous 10 years confirmed that the failure to publish is not a random event, but is heavily influenced by the direction and strength of research findings, whereby manuscripts with statistically significant (positive) results are published preferentially over manuscripts reporting nonsignificant (negative) results.[21] Publication bias may be the main factor behind the systematic differences found between studies funded by industry and their counterparts.[14,22]

Efforts have been made to eliminate publication bias through compulsory registration of trials at inception, and publication of the results of all trials. These have been the focus of intense debate and controversy for several years, fueled by strong ethical and economic interests. Many major journals now refuse to publish the results of studies that had not been registered at inception. Even so, readers must be aware that by relying on published studies to guide their decisions they are always at risk of overestimating the effect of interventions[23–25] (see Chapter 5).

What is language bias?

Recently, a variation of publication bias has been described as *language bias*, to indicate that manuscripts may be submitted to and published by journals in different languages depending on the direction of their results. More studies with positive results may be published in English.[26] A variant of this is the *country of publication bias*, the tendency by some countries to publish a disproportionate number of positive trials.[27]

What is time lag bias?

This bias occurs when the speed of publication depends on the direction and strength of the trial results. In general, it seems that trials with 'negative' results take twice as long to be published as 'positive' trials.[28,29]

What biases can occur during the uptake phase?

Up to this point we have focused on the biases introduced by the investigators who plan and carry out randomized trials, or those who publish and disseminate the results. As this book is primarily a user's guide, rather than a manual for researchers, we felt that we should emphasize the responsibility of the reader of research studies.

Different types of reader biases were described many years ago.[4] At the time in which they were reported, the existence of these biases was supported only by common sense and experience. Recently, there have been empirical studies that support the existence of reader bias, showing that there are systematic differences in the way readers assess the quality of RCTs depending on whether the assessments are conducted under masked or open conditions.[11,30] These studies, however, do not focus on any specific type of reader bias.

The following are some of the biases that we believe are most common and pertinent:

Relation to the author bias, with its subgroups *Rivalry bias* (underrating the strengths or exaggerating the weaknesses of studies published by a rival) and *I owe him one* bias (favoring flawed results from a study by someone who did the same for the reader).

Personal habit bias occurs when readers overrate or underrate a study depending on their own habits (e.g. a reader who enjoys eating animal fat overrating a study that challenges the adverse effects of animal fat on health). This is similar to the *moral bias*, in which readers overrate or underrate a study depending on how much it agrees or disagrees with their moral views (e.g. a reader who regards abortion as immoral overrating a study showing a relationship between abortion and breast cancer). This is closely related to the *values bias* (depending on how important you consider the outcomes of the study to be).

Clinical practice bias takes place when readers judge a study according to whether it supports or challenges their current or past clinical practice (e.g. a clinician who gives lidocaine to patients with acute myocardial infarction underrating a study that suggests that lidocaine may increase mortality in these patients). This is similar to

the *institution bias* (that is, or is not, the way that we do it in our hospital), and the *territory bias* which can occur when readers overrate studies that support their own specialty or profession (e.g. a surgeon favoring a study that suggests that surgery is more effective than medical treatment, or obstetricians underrating a study that suggests that midwives can provide adequate care during uncomplicated pregnancies and deliveries). *Tradition bias* happens when a reader rates a study depending on whether it supports or challenges traditional procedures (e.g. underrating a study that challenges episiotomy during normal vaginal deliveries).

Do something bias means overrating a study that suggests that an intervention is effective, particularly when there is no alternative effective intervention available. This bias may be common among clinicians and patients (e.g. a patient with AIDS overrating a study describing a cure for AIDS).

In this general heading we can include the *technology bias*, which relates to judging a study according to the reader's attraction or aversion for technology in health care. *Resource allocation bias* happens when readers have a strong preference for one type of resource allocation. This bias may be one of the most frequently found in health care, as it can emanate from consumers, clinicians, policy makers, researchers, and fund holders.

Printed word bias occurs when a study is overrated because of undue confidence in published data. Subgroups of the printed word bias include the *prestigious journal bias* (the results of studies published in prestigious journals are overrated), and its opposite, the *non-prestigious journal bias*. Similar to this is the *peer review bias*, which comes into play when readers have an unwarranted belief in the ability of peer review to guarantee the validity of a study.

Prominent author bias occurs when the results of studies published by prominent authors are overrated, and, of course has its converse in the *unknown or non-prominent author bias*. This has been called the '*who is s/he? bias*'.[4] Similar to these are the *famous institution bias*, the *credential or professional background bias* (e.g. physicians underrating research done by nurses or vice versa; basic scientists underrating research done by clinicians or vice versa; PhDs underrating studies published by MDs and vice versa; readers overrating research by authors with many letters after their names and vice versa). Their variants include the *esteemed author bias*, *esteemed professor bia*s, and the *friendship bias*; when the reader overrates results obtained by a close friend or mentor.

We are not through yet!

Geography bias occurs when studies are judged according to the country or region where it was conducted, and is closely related to the *language bias* (e.g. the belief that studies published in languages other than English are of inferior quality than those published in English).[26]

The *trial design bias* can go in either direction. The *favored design bias* occurs when a study that uses a design supported, publicly or privately, by the reader (e.g. a consumer advocate overrating an RCT that takes into account patient preferences). Its converse is the *Disfavored design bias*. Somewhat related are the *large trial bias*, in which the results of large trials are overrated, and the *multicentre trial bias* when the results of multicentre collaborative trials are overrated. The *small trial bias* occurs when the results of trials with small sample size are underrated, particularly when they contradict the opinion of the reader (i.e. attributing to chance any statistically or clinically significant effect found by a small trial, or any lack of significant effects to low power).

Complementary medicine bias refers to the systematic overrating or underrating of studies that describe complementary medicine interventions, particularly when the results suggest that the interventions are effective.

Flashy title bias occurs when the results of studies with attractive titles are overrated (particularly by patients or journalists) or underrated (particularly by academics if they regard them as sensationalist!). Other rather tricky biases include the *substituted question bias,* when a reader substitutes a question for the question that the study is designed to answer and regards the results of the study as invalid if they do not answer the substituted question.

Vested interest bias has a number of subgroups. *Bankbook bias* occurs when a study is rated depending on the impact of its results on the income of the reader (e.g. a surgeon underrating a study that questions the need for surgery to relieve back pain in patients with spinal stenosis, or a pharmaceutical company overrating the results of a study that supports the use of one of its products). *Cherished belief bias* reminds us that there are other competing interests besides the financial ones.

Reader attitude biases include the *Belligerence bias* which results in underrating studies systematically just for the sake of being difficult; the *Empiricism bias* (overrating or underrating a study because it challenges the clinical experience of the reader), or the *I am an*

epidemiologist bias in which the reader repudiates a study that contains any flaw, albeit minor, in its design, analysis or interpretation.

Finally, *careless reading bias* occurs when a study is overrated or underrated because the reader neglected to read a key section. Unfortunately, far too common.

Musings

This has been a difficult chapter to write. We approached it with fear and trepidation, feeling part of a 'no win' situation. We know that the control of bias is the *raison d'etre* for clinical trials, and accept that control of bias is the most important factor in diminishing inevitable error. We know that allocation bias is a major source of potential error in clinical comparison studies, and we know that randomization, if properly done, can control for allocation bias. We want to stress the value of randomization for this purpose, and the vital importance of RCTs.

But we also realize that randomization *per se* can control *only* for allocation bias, and this does not even completely control for selection bias. Other biases can also subvert the validity of conclusions at any stage in the planning, conduct, analysis, or interpretation of the results. As we worked together on this chapter, as we uncovered an increasing number of biases, our fears mounted. We started to feel very discouraged. What is the big deal, if this seemingly powerful tool is so vulnerable? Why should we believe in trials if they can be subverted so easily and at so many levels? If biases cannot be controlled, what is left? We are not sufficiently naïve to think that by finding biases and naming them that we can overcome them. Can we run the risk that by drawing attention to the biases we would attack the very foundation of RCTs, and appear to advocate nihilism?

We believed (and still believe) in the value of RCTs. We felt like heretics, not for the first time.[31] Both of us were, and are, strong and enthusiastic proponents of RCTs. Indeed our support for RCTs has become even stronger as we have become more aware of their limitations. But it is no longer a blind faith, rather one that has been through and survived the crises of doubt.

We are concerned with the danger that RCTs may be perceived as a sort of talisman, to protect us from the evil of bias. But randomized trials are not divine revelations, they are human constructs, and like all human constructs, are fallible. They are valuable, useful

tools that should be used wisely and well. We believe that a strong belief in the strength of randomized trials, without acknowledging their weaknesses, runs the risk of fundamentalism and intolerance of criticism, or alternative views. In this way, it can discourage innovation.

Our list of biases is far from exhaustive. The number of possible biases is practically infinite, as is the names that can be given to them, or the ways in which they can be classified or categorized. RCTs can never be completely objective. They should be carried out with humility; the investigator should be as up front, explicit, and transparent as possible about his or her motivations for choosing to carry out the trial, the methods used, the outcomes looked for as well as the outcomes found. Journalists have an important responsibility to assume, because of their influence on public understanding. At present they tend to bring to public attention the results of trials purporting beneficial effects of a new intervention for incurable diseases, while they ignore the results of previous (or concurrent) trials in which the same intervention showed no benefit.[32] This media coverage may influence the decisions of clinicians and patients who are not aware of the other studies. The same onus must be put on the reader, the one who will be making use of the information gleaned from the trial. It can be far too easy to criticize an RCT, or to read into it what we want, to find rather than what the results actually show.

Our bottom line is that a new sense of freedom can emerge, as we free ourselves from a false sense of objectivity, and can recognize and use RCTs as the valuable tools that they are, when they are the right tool in the right place.

References

1. Jadad AR, Rennie D. The randomized controlled trial gets a middle-aged checkup. *Journal of American Medical Association* 1998;279:319–320.
2. http://www.wordreference.com/definition/bias (accessed December 22, 2006).
3. *Webster's Third New International Dictionary*, Unabridged. G&C Merriam Co, 1976.
4. Owen R. Reader bias. *Journal of American Medical Association* 1982; 247:2533–2534.
5. Chalmers TC, Celano P, Sacks HS, Smith H. Bias in treatment assignment in controlled clinical trials. *New England Journal of Medicine* 1983; 309:1359–1361.

6. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *Journal of American Medical Association* 1995;273:408–412.
7. Moher D, Fortin P, Jadad AR, Juni P, Klassen T, Le Lorier J, Liberati A, Linde K, Penna A. Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. *Lancet* 1996;347:363–366.
8. Schulz KF. Subverting randomization in controlled trials. *Journal of American Medical Association* 1995;274:1456–1458.
9. Colditz GA, Miller JN, Mosteller F. How study design affects outcomes in comparisons of therapy. I: Therapy. *Statistics in Medicine* 1989;8:441–454.
10. Schulz KF, Grimes DA, Altman DG, Hayes RJ. Blinding and exclusions after allocation in randomised controlled trials: survey of published parallel group trials in obstetrics and gynaecology. *British Medical Journal* 1996;312:742–744.
11. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds JM, Gavaghan DJ, McQuay DM. Assessing the quality of reports on randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials* 1996;17:1–12.
12. Moher D, Jadad AR, Tugwell P. Assessing the quality of randomized controlled trials. *International Journal of Technology Assessment in Health Care* 1996;12:195–208.
13. Freeman TB. Use of placebo surgery in controlled trials of a cellular-based therapy for Parkinson's disease. *New England Journal of Medicine* 1999;34:988–992.
14. Fries JF, Krishnan E. Equipoise, design bias, and randomized controlled trials: the elusive ethics of new drug development. *Arthritis Research and Therapy* 2004;6:R250–R255.
15. Sackett DL, Wennberg JE. Choosing the best research design for each question: It's time to stop squabbling over the 'best' methods. *British Medical Journal* 1997;315:1636.
16. Lilford RJ, Braunholtz DA, Greenhalgh R, Edwards SJL. Trials and fast changing technologies: the case for tracker studies. *British Medical Journal* 2000;320;43–46.
17. Kotaska A. Inappropriate use of randomised trials to evaluate complex phenomena: case study of vaginal breech delivery. *British Medical Journal* 2004;329:1039–1042
18. Martin G. Munchausen's statistical grid, which makes all trials significant. *Lancet* 1984;ii:1457.
19. Dickersin K. The existence of publication bias and risk factors for its occurrence. *Journal of American Medical Association* 1990;263: 1385–1389.
20. Rennie D, Flanagin A. Publication bias – the triumph of hope over experience. *Journal of American Medical Association* 1992;267:411–412.

21. Dickersin K. How important is publication bias? A synthesis of available data. *AIDS Education Prevention* 1997;9(Suppl A):15–21.
22. Smith R. Medical Journals are an extension of the marketing arm of pharmaceutical companies. http://medicine.plosjournals.org/perlserv/?request=get-document&doi=10.1371/journal.pmed.0020124 (accessed November 21, 2006)
23. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D, Stroup DF. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *Journal of American Medical Association* 1996;276:637–639.
24. Grimes DA. The 'CONSORT' guidelines for randomized controlled trials in Obstetrics and Gynecology. *Obstetrics and Gynecology* 2002;100:631–632.
25. Rennie D. How to report randomized controlled trials. The CONSORT statement. *Journal of American Medical Association* 1996;276:649.
26. Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technology Assessment* 2003;7:1–90.
27. Vickers A, Goyal N, Harland R, Rees R. Do certain countries produce only positive results? A systematic review of controlled trials. *Controlled Clinical Trials* 1998;19:159–166.
28. Ioannidis JPA. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials: a survival analysis. *Journal of American Medical Association* 1998;279:281–286.
29. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *British Medical Journal* 1997;315:640–645.
30. McNutt RA, Evans AT, Fletcher RH, Fletcher SW. The effects of blinding on the quality of peer review. A randomized trial. *Journal of American Medical Association* 1990;263:1371–1376.
31. Enkin MW, Jadad AR. Using anecdotal information in evidence-based health care: heresy or necessity? *Annals of Oncology* 1998;9:963–966.
32. Koren G, Klein N. Bias against negative studies in newspaper reports of medical research. *Journal of American Medical Association* 1991;266:1824–1826.