Chapter 7
Blinding

In any clinical trial bias in determining treatment effects is one of the main concerns. Bias may be defined as systematic error, or "difference between the true value and that actually obtained due to all causes other than sampling variability" [1]. It can be caused by conscious factors, subconscious factors, or both. Bias can occur at a number of places in a clinical trial, from the initial design through data analysis, interpretation and reporting. One general solution to the problem of bias is to keep the participants and the investigators blinded, or masked, to the identity of the assigned intervention. One can also blind several other aspects of a trial including the assessment, classification and evaluation of the response variables. A large sample size does not reduce bias.

The history of blind assessment in medicine goes back more than 200 years [2]. In its simplest form, the investigators used blindfolds or curtains so that the participants would not know the nature or timing of the intervention. Dummy interventions were also utilized from the inception. The first series of blind assessment was directed at mesmerism, an intervention based on a new "healing fluid" in nature called "animal magnetism." A group of women was involved in the first blindfold experiment. Its conclusion was the "while the woman was permitted to see the operation, she placed her sensations precisely in the part towards which it was directed; that on the other hand, when she did not see the operation, she placed them at hazard, and in parts very distant from those which were the object of magnetism." In another type of experiment, women were told that they were receiving mesmerism from an adjoining room through a paper curtain over a door. The knowledge of intervention produced sensations. When they received treatment but were not told they were mesmerized, nothing happened. Blinding eliminated the effects of mesmerism, and sham worked as well as "real" mesmerism.

The first clinical trial that in modern time applied the principle of blinding was published in 1931 by Amberson et al. [3]. This trial was probably also the first trial that employed a form of random assignment of participants to the study groups.

L.M. Friedman et al., *Fundamentals of Clinical Trials*,
DOI 10.1007/978-3-319-18539-2_7

Fundamental Point

A clinical trial should, ideally, have a double-blind design in order to limit potential problems of bias during data collection and assessment. In studies where such a design is impossible, other measures to reduce potential bias are advocated.

Who Is Blinded?

The blinding terminology is not well understood. A survey of 91 internal medicine physicians in Canada from 2001 [4] showed that 75% knew the definition of single-blind. Approximately 40% understood the proper definition of double-blind. A more recent survey showed that the understanding of the blinding terminology has not improved [5]. Among 66 single-blind trials, the investigators were asked who was blinded. Twenty-six said the patients, 22 the outcome assessors and 16 the data analysts/investigators. Viergever and Ghersi [5] also reviewed to what extent information of blinding was part of registered records of clinical trials. They concluded that this information was often not provided or was of poor quality in trial publications. The authors concluded that the term double-blind was found to be common despite the lack of clarity on its exact meaning.

The meaning of the term double-blind has been addressed in recent publications. Responders to a survey of 200 blinded RCTs from the Cochrane Central Register of Controlled Trials provided their operational meanings of the term [6]. The authors were asked which of the following six categories of key trial persons had been blinded: (1) patients, (2) health care providers responsible for care, (3) data collectors, (4) assessors of outcome (including the data monitoring committee), (5) data analysts or (6) manuscript writers. Fifteen different answers were given for the term "double-blind". The most common answers included patients (97%), health care providers (89%), data collectors (90%) and outcome assessors (69%).

The use of the terms single-blind and double-blind is particularly inconsistent in trials of non-pharmaceutical interventions [7].

Types of Blinding

Unblinded

In an unblinded or open trial, both the participant and all investigators know to which intervention the participant has been assigned. Some kinds of trials are primarily conducted in this manner and those include most surgical procedures, comparisons of devices and medical treatment, changes in lifestyle (e.g. eating habits, exercise, cigarette smoking) or learning techniques. Approaches to blinding elements of non-pharmacologic interventions are discussed below.

An unblinded study is appealing for two reasons. First, investigators are likely to be more comfortable making decisions, such as whether or not to continue a participant on the assigned study medication if they know its identity. Second, all other things being equal, it is often simpler to execute than other studies. The usual drug trial may be easier to design and carry out, and consequently less expensive, if blinding is not an issue. Also, it has been argued that it more accurately reflects clinical practice [8]. However, an unblinded trial need not be simple. For example, trials that simultaneously attempt to induce lifestyle changes and test drug interventions can be fairly complex. An example is the Women's Health Initiative [9] which had three distinct interventions: hormone replacement therapy, calcium and vitamin D supplementation and an unblinded dietary intervention.

The main disadvantage of an unblinded trial is the possibility of bias. Participant reporting of symptoms and side effects and prescription of concomitant or compensatory treatment are all susceptible to bias. Other problems of biased data collection and assessment by the investigator are addressed in Chap. 11. Since participants when joining a trial have sincere hopes and expectations about beneficial effects, they may become dissatisfied and drop-out of the trial in disproportionately large numbers if not on the new or experimental intervention. The benefit of blinding in trials of a short intervention (such as treatment with a fibrinolytic agent for acute myocardial infarction) where differential drop-out is unlikely, and with an outcome (like all-cause mortality) that is not subject to ascertainment bias can be debated. However, even in these trials assessment of other adverse events will be protected from bias with blinding.

A trial of the possible benefits of ascorbic acid (vitamin C) in the common cold was designed as a double-blind study [10, 11]. However, it soon became apparent that many of the participants, most of whom were medical staff, discovered mainly by tasting whether they were on ascorbic acid or placebo. As more participants became aware of their medication's identity, the dropout rate in the placebo group increased. Since evaluation of severity and duration of colds depended on the participants' reporting of their symptoms, this unblinding was important. Among those participants who claimed not to know the identity of the treatment at the end of the trial, ascorbic acid showed no benefit over placebo. In contrast, among participants who knew or guessed what they were on, ascorbic acid did better than placebo. Therefore, preconceived notions about the benefit of a treatment, coupled with a subjective response variable, may have yielded biased reporting. The investigators' willingness to share this experience provided us with a nice illustration of the importance of maintaining blinding.

In a trial of coronary artery bypass surgery versus medical treatment [12], the number of participants who smoked was equal in the two study groups at baseline. During the early part of follow-up, there were significantly fewer smokers in the surgical group than in the medical group. A possible explanation could have been that the unblinded surgeons gave more anti-smoking advice to those randomized to surgery. The effect of this group difference on the outcome of the trial is difficult, if not impossible, to assess.

Single-Blind

The established definition of a single-blind study is that only the participants are unaware of which intervention they are receiving. The advantages of this design are similar to those of an unblinded study—it is usually simpler to carry out than a double-blind design, and knowledge of the intervention may help the investigators exercise their best judgment when caring for the participants. Indeed, certain investigators are reluctant to participate in studies in which they do not know the study group assignment. They may recognize that bias is partially reduced by keeping the participant blinded but feel that the participant's health and safety are best served if they themselves are not blinded.

The disadvantages of a single-blind design are similar to, though not so pronounced as, those of an unblinded design. The investigator avoids the problems of biased participant reporting, but she herself can affect the administration of non-study therapy, data collection, and data assessment. For example, a single-blind study reported benefits from zinc administration in a group of people with taste disorders [13]. Because of the possibility of bias in a study using a response variable as subjective and hard to measure as taste, the study was repeated, using a type of crossover, double-blind design [14]. This second study showed that zinc, when compared with placebo, did not relieve the taste disorders of the study group. The extent of the blinding of the participants did not change; therefore, presumably, knowledge of drug identity by the investigator was important. The results of treatment cross-over were equally revealing. In the single-blind study, participants who did not improve when given placebo as the first treatment, "improved" when placed on zinc. However, in all four double-blind, cross-over procedures (placebo to zinc, placebo to placebo, zinc to zinc, zinc to placebo), the participants who had previously shown no improvement on the first treatment did show benefit when given the second medication. Thus, the expectation that the participants who failed to respond to the first drug were now being given an active drug may have been sufficient to produce a positive response.

Another example comes from two noninferiority trials comparing ximelagatran, a novel oral direct thrombin inhibitor, to warfarin for the prevention of thromboembolic events in people with nonvalvular atrial fibrillation [15]. The first trial, SPORTIF III, was single-blind with blinded events assessment, while the second trial, SPORTIF V, was double-blind. The primary response variable was all strokes and systemic embolic events. The observed risk ratio in the single-blind SPORTIF III was 0.71 (95% CI, 0.48–1.07) while the result trended in the opposite direction in the double-blind SPORTIF V with a risk ratio of 1.38 (95% CI, 0.91–2.10). One cannot be sure how much bias may have played a role, but, in general, more confidence ought to be placed on trials with a double-blind design. A more recent example is the open label trials of renal artery denervation for resistant hypertension that reported large treatment benefits not seen in a subsequent sham-controlled blinded trial [16].

Both unblinded and single-blind trials are vulnerable to another source of potential bias introduced by the investigators. This relates to group differences in *compensatory* and *concomitant* treatment. Investigators may feel that the control group is not being given the same opportunity as the intervention group and, as a result, may prescribe additional treatment as "compensation." This may be in the form of advice or therapy. For example, several studies have attempted blood pressure lowering as either the sole intervention, or as part of a broader effort. In general, the investigators would make an intensive effort to persuade participants in the intervention group to take their study medication. To persuade successfully the investigators themselves had to be convinced that blood pressure reduction was likely beneficial. When they were seeing participants who had been assigned to the control group, this conviction was difficult to suppress. Therefore, participants in the control group were likely to have been instructed about non-pharmacological ways by which to lower their blood pressure or other preventive treatments. The result of compensatory treatment is a diminution of the difference in blood pressure or even hypertension-related outcomes between the intervention group and the "untreated," control group. This may also have been a factor in a heart failure trial of dronedarone, an antiarrhythmic drug, in which the intervention group had a higher mortality rate [17]. This finding in the dronedarone group would in part be due to a lower use of ACE-inhibitors, a drug class known to reduce mortality.

Working against this is the fact that investigators typically prefer to be associated with a study that gives positive findings. Favorable results published in a reputable journal are likely to lead to more invitations to present the findings at scientific meetings and grand rounds and can also support academic promotions. Investigators may, therefore, subconsciously favor the intervention group when they deal with participants, collect data, and assess and interpret results, although this may perhaps be less of an issue in multicenter trials.

Concomitant treatment means any non-study therapy administered to participants during a trial. If such treatment is likely to influence the response variable, this needs to be considered when determining sample size. Of more concern is the bias that can be introduced if concomitant treatment is applied unequally in the two groups. In order to bias the outcome of a trial, concomitant treatment must be effective, and it must be used in a high proportion of the participants. When this is the case, bias is a possibility and may occur in either direction, depending on whether the concomitant treatment is preferentially used in the control, or in the intervention group. It is usually impossible to determine the direction and magnitude of such bias in advance or its impact after it has occurred.

Double-Blind

In a double-blind study, neither the participants nor the investigators or more specifically the team of investigators responsible for following the participants, collecting data, and assessing outcomes should know the identity of the intervention

assignment. Such designs are usually restricted to trials of drugs or biologics. It is theoretically possible to design a study comparing two surgical procedures or implantation of two devices in which the surgeon performing the operation knows the type of surgery or device, but neither the study investigator nor the participant knows. Similarly, one might be able to design a study comparing two diets in which the food looks identical. However, such trials are uncommon.

The main advantage of a truly double-blind study is that the risk of bias is reduced. Preconceived ideas of the investigators will be less important, because they will not know which intervention a particular participant is receiving. Any effect of their actions, therefore, would theoretically occur equally in the intervention and control groups. As discussed later, the possibility of bias may never be completely eliminated. However, a well designed and properly run double-blind study can minimize bias. As in the example of the trial of zinc and taste impairment, double-blind studies have at times led to results that differ from unblinded or single blind studies. Such cases illustrate the role of bias as a factor in clinical trials.

In a double-blind trial certain functions, which in open or single-blind studies could be accomplished by the investigators, might sometimes be taken over by others in order to maintain the blinding. These functions include participant care if it is important for patient care to know the intervention, collection of efficacy and safety data that might disclose the nature of the intervention, and assessment and monitoring of treatment outcomes. Typically, an outside body needs to monitor the data for toxicity and benefit, especially in long-term trials. Chapter 17 discusses data monitoring in greater detail. A person other than the investigator who sees the participants needs to be responsible for assigning the interventions to the participants. Treatments that require continuous dose adjustment, such as warfarin, are difficult to blind, but it can be accomplished. In one trial [18], an unblinded pharmacist or physician adjusted the warfarin doses according to an algorithm for maintaining the International Normalized Ratio (INR), a measure of anticoagulation, within a pre-specified range but also adjusted the placebo doses randomly. The authors concluded that "placebo warfarin dose adjustment schedules can protect blinding adequately" for participants and investigators and recommended their use for future trials of warfarin. A similar approach was employed in the Coumadin Aspirin Reinfarction Study [19]. An INR control center adjusted the doses in the three treatment arms to keep the INR values below the prespecified safety limits and to maintain the double-blind. In another trial [20] a point of care device was used that encrypted result that was a true INR for participants on warfarin and a sham INR for those not on warfarin. These INR values were used for dose adjustments. The system seemed to work well to maintain blinding.

The double-blind design is no protection against imbalances in use of concomitant medications. A placebo-controlled trial of a long-acting inhaled anticholinergic medication in participants with chronic obstructive pulmonary disease allowed the use of any other available drug treatment for this condition as well as a short-acting inhaled anticholinergic agent for acute exacerbations [21]. The extent of this co-intervention is likely to differ between the actively treated and the placebo groups, but the findings regarding concomitant drug use by study group were not

presented. Moreover, it may have influenced symptomology as well as risks of disease events and made it very difficult to determine the true effects of the long-acting anticholinergic inhaler. Reporting the proportion of participants given a co-intervention at any time over the four years of the trial by treatment group would have helped the interpretation of results, even if the frequency and intensity of its use were not reported.

In many single- and double-blind drug trials the control group is placed on a matched placebo. Much debate has centered on the ethics of using a placebo. See Chap. 2 for a further discussion of this issue.

Triple-Blind

A triple-blind study is an extension of the double-blind design; the committee monitoring response variables is not told the identity of the groups. The committee is simply given data for groups A and B. A triple-blind study has the theoretical advantage of allowing the monitoring committee to evaluate the response variable results more objectively. This assumes that appraisal of efficacy and harm, as well as requests for special analyses, may be biased if group identity is known. However, in a trial where the monitoring committee has an ethical responsibility to ensure participant safety, such a design may be counterproductive. When hampered in the safety-monitoring role, the committee cannot carry out its responsibility to minimize harm to the participants, since monitoring is often guided by the constellation of trends and their directions. In addition, even if the committee could discharge its duties adequately while being kept blinded, many investigators would be uneasy participating in such a study. Though in most cases the monitoring committee looks only at group data and can rarely make informed judgments about individuals, the investigators still rely on the committee to safeguard their study participants. This may not be a completely rational approach because, by the time many monitoring committees receive data, often any emergency situation has long passed. Nevertheless, the discomfort many investigators feel about participating in double-blind studies would be magnified should the data monitoring committee also be kept blinded.

Finally, people tend not to accept beneficial outcomes unless a statistically significant difference has been achieved. Rarely, though, will investigators want to continue a study in order to achieve a clearly significant difference in an adverse direction; that is, until the intervention is statistically significantly worse or more harmful than the control. Therefore, many monitoring committees demand to know which study groups are on which intervention.

A triple-blind study can be conducted ethically if the monitoring committee asks itself at each meeting whether the direction of observed trends matters. If it does not matter, then the triple-blind can be maintained, at least for the time being. This implies that the monitoring committee can ask to be unblinded at any time it chooses. In the Randomized Aldactone Evaluation Study (RALES), the Data and

Safety Monitoring Board was split and several members argued against being blinded [22]. However, triple-blind was employed initially. For most outcome variables, the treatment groups were labeled A and B. Since increased rates of gynecomastia and hyperkalemia, which might occur with aldactone, would unmask the A and B assignments, these adverse events were labeled X and Y. Using different labels for these adverse events prevented unblinding of the other outcome variables.

Triple blinding may at times be useful, but if trends in important clinical outcomes or an imbalance in adverse effects develop, it is no longer appropriate.

Protecting the Double-Blind Design

Double-blind studies are more difficult to carry out than other trials. One must ensure that the investigator team remains blinded and that any data which conceivably might endanger blinding be kept from them during the study. An effective data monitoring scheme must be set up, and emergency unblinding procedures must be established. These requirements pose their own problems and can increase the cost of a study. Page and Persch [23] discuss strategies for blinding health care providers and data collectors. The latter ought to be different from those providing medical care for the participants.

An old illustration is the Aspirin Myocardial Infarction Study [24], a double-blind trial of aspirin in people with coronary heart disease, in which the investigators wished to monitor the action of aspirin on platelets. A postulated beneficial effect of aspirin relates to its ability to reduce the aggregation of platelets. Therefore, measuring platelet aggregation provided both an estimate of whether the aspirin treated group was getting a sufficient dose and a basis for measurement of participant adherence. However, tests of platelet aggregation needed to be performed shortly after the blood sample was drawn. The usual method used to have a laboratory technician insert the specimen in an aggregometer, add a material such as epinephrine (which, in the absence of aspirin, causes platelets to aggregate) and analyze a curve which is printed on a paper strip. In order to maintain the blind, the study needed to find a way to keep the technician from seeing the curve. Therefore, a cassette tape-recorder was substituted for the usual paper strip recorder and the indicator needle was covered. These changes required a modification of the aggregometer. All of the 30 clinics required this equipment, so the adjustment was expensive. However, it helped ensure the maintenance of the blind.

A double-blind design is a particular problem in clinical trials of treatments other than the use of pharmaceuticals. Methods of blinding procedures in 123 reports of nonpharmacological trials were systematically reviewed by Boutron et al. [25]. Three categories were classified: surgical or technical procedures, participative interventions, and devices. Most of the reports used some form of sham procedure. For surgical interventions the sham procedure was a simulating intervention. The controls in participative interventions were either an attention-control

intervention or a differently administered placebo. The device trials used sham prosthesis, identical apparatus, and simulation of a device. A small number of nonpharmacological trials blinded the participants to the study hypothesis. A critical approach employed in one-third of the trials was a blinded, centralized assessment of the primary outcome.

Protecting the double-blind can be a special problem in active-control trials, i.e. trials comparing active interventions. The adverse drug effect patterns for the drugs being compared can be distinctly different. When the selective serotonin receptor inhibitors were introduced they were compared to tricyclic antidepressants. The latter are anticholinergic and commonly cause dryness of mouth, blurred vision and tachycardia. The occurrence of these adverse effects unblinded treatment in a large number of participants in 20 comparative trials [26].

Naturally, participants want to be on the "better" intervention. In a drug trial, the "better" intervention usually is presumed to be the new one; in the case of a placebo-control trial it is presumed to be the active medication. Investigators may also be curious about a drug's identity. For these reasons, consciously or subconsciously, both participants and investigators may try to unblind the medication. Unblinding can be done deliberately by going so far as to have the drug analyzed, or in a less purposeful manner by "accidentally" breaking open capsules, holding pills up to the light, carefully testing them, or by taking any of numerous other actions. In the first case, which may have occurred in the vitamin C study discussed earlier, little can be done to ensure blinding absolutely. Curious participants and investigators can discover many ways to unblind the trial, whatever precautions are taken. Probably, however, the less purposeful unblinding is more common.

We strongly recommend that the assessment of trial outcomes be as objective as possible. This means that the person at the clinic making these assessments be blinded. At times, this may be done at a central location.

Matching of Drugs

Drug studies, in particular, lend themselves to double-blind designs. One of the surest ways to unblind a drug study is to have dissimilar appearing medications. When the treatment identity of one participant becomes known to the investigator, the whole trial is unblinded. Thus, matching of drugs is essential.

Proper matching has received little attention in the literature. A notable exception is the vitamin C study [10, 11] in which of the double-blind was not maintained throughout the trial. One possible reason given by the investigators was that, in the rush to begin the study, the contents of the capsules were not carefully produced. The lactose placebo could easily be distinguished from ascorbic acid by taste, as the study participants quickly discovered. An early report showed similar concern [27]. The authors noted that, of 22 studies surveyed, only five had excellent matching between the drugs being tested. A number of features of matching must be considered. A review of 191 randomized placebo-controlled trials from leading

general medicine and psychiatry trials showed that 81 (42%) trials reported on the matching of drug characteristics [28]. Only 19 (10%) commented on more than one of the matching features and appearance was, by far, the most commonly reported characteristic. Thus, most reports of drug studies do not indicate how closely tablets or capsules resembled one another, or how great a problem was caused by imperfect matching.

Cross-over studies, where each subject sees both medications, require the most care in matching. Visual discrepancies can occur in size, shape, color, and texture. Ensuring that these characteristics are identical may not be simple. In the case of tablets, dyes or coatings may adhere differently to the active ingredient than to the placebo, causing slight differences in color or sheen. Agents can also differ in odor. The taste and the local action on the tongue of the active medication are likely to be different than those of the placebo. For example, propranolol is a topical anesthetic which causes lingular numbness if held in the mouth. Farr and Gwaltney reported on problems in matching zinc lozenges against placebo [29]. Because zinc lozenges are difficult to blind, the authors questioned whether studies using zinc for common cold prevention were truly valid. They conducted trials illustrating that if a placebo is inadequately matched, the "unpleasant side effects of zinc" may reduce the perception of cold symptoms.

Drug preparations should be pretested if it is possible. One method is to have a panel of observers unconnected with the study compare samples of the medications. Perfect matches are almost impossible to obtain and some differences are to be expected. Preparing placebos for trial of herbal medicines can be a challenge. One way is the use of appropriately matched placebo capsules, an approach applied successfully [30]. However, beyond detecting differences, it is important to assess whether the observers can actually identify the agents. If not, slightly imperfect matches may be tolerated. The investigator must remember that, except in cross-over studies, the participant has only one drug and is therefore not able to make a comparison. On the other hand, participants may meet and talk in waiting rooms, or in some other way compare notes or pills. Of course, staff always have the opportunity to compare different preparations and undermine the integrity of a study.

Differences may become evident only after some time, due to degradation of the active ingredient. Freshly prepared aspirin is relatively odor free, but after a while, tell-tale acetic acid accumulates. Ginkgo biloba has a distinct odor and a bitter taste. In one trial of Ginkgo, the investigators used coated tablets to mask both odor and taste [31]. The tablets were placed in blister packs to reduce the risk of odor. Quinine was added to the placebo tablets to make them as bitter as the active drug. This approach prevented any known blind-breaking.

Use of substances to mask characteristic taste, color, or odor, as was done in the ginkgo biloba trial mentioned above, is often advocated. Adding vanilla to the outside of tablets may mask an odor; adding dyes will mask dissimilar colors. A substance such as quinine or quassin will impart a bitter taste to the preparations. Not only will these chemical substances mask differences in taste, but they will also effectively discourage participants from biting into a preparation more than once.

However, the possibility that they may have toxic effects after long-term use or even cause allergic reactions in a small percent of the participants must be considered. It is usually prudent to avoid using extra substances unless absolutely essential to prevent unblinding of the study.

Less obviously, the weight or specific gravity of the tablets may differ. Matching the agents on all of these characteristics may be impossible. However, if a great deal of effort and money are being spent on the trial, a real attempt to ensure matching makes sense. The investigator also needs to make sure that the containers are identical. Bottles and vials need to be free of any marks other than codes which are indecipherable except with the key.

Sometimes, two or more active drugs are being compared. The ideal method of blinding is to have the active agents look alike, either by formulating them appropriately or possibly by enclosing them in identical capsules. The former may not be possible, and the latter may be expensive or require capsules too large to be practical. In addition, enclosing tablets in capsules may change the rate of absorption and the time to treatment response. In a comparative acute migraine trial, one manufacturer benefitted from encapsulating a competitor's FDA-approved tablet in a gelatin capsule [32]. A better, simpler and more common option is to implement a "double-dummy." Each active agent has a placebo identical to it. Each study participant would then take two medications. This is a good approach when the administration of the two drugs being compared is different, for example, when a once daily drug is being compared to a twice daily drug. A pharmaceutical sponsor may sometimes have problems finding a matching placebo for a competitor's product.

Sometimes, if two or more active agents are being compared against placebo, it may not be feasible to make all drugs appear identical. As long as each active agent is not being compared against another, but only against placebo, one option is to create a placebo for each active drug or a so-called "double-dummy." Another option is to limit the number of placebos. For example, assume the trial consists of active drugs A, B, and C and placebo groups. If each group is the same size, one third of placebo groups will take a placebo designed to look like active drug A, one third will take a placebo designed to look like drug B, and one third, like active drug C. This design was successfully implemented in at least one reported study [33].

Coding of Drugs

By drug coding is meant the labeling of individual drug bottles or vials so that the identity of the drug is not disclosed. Coding is usually done by means of assigning a random set of numbers to the active drug and a different set to the control. Each participant should have a unique drug code which remains with him for the duration of the trial. If only one code is used for each study group, unblinding a single participant would result in unblinding everybody. Furthermore, many drugs

have specific side effects. One side effect in one participant may not be attributable to the drug, but a constellation of several side effects in several participants with the same drug code may easily unblind the whole study.

In large studies it is possible through use of computer programs to make up and stock drugs under hundreds or thousands of unique codes. Bar coding of the bottles with study medication is now common. This type of coding has no operational limits on the number of unique codes, it simplifies keeping an accurate and current inventory of all study medications and helps assure that each participant is dispensed his assigned study medication.

Official Unblinding

A procedure should be developed to break the blind quickly for any individual participant at any time should it be in his best interest. Such systems include having labels on file in the hospital pharmacy or other accessible locations, or having an "on call" 24 hour-a-day process so that the assignment can be decoded. In order to avoid needless breaking of the code, someone other than the investigator could hold a list that reveals the identity of each drug code. Alternatively, each study medication bottle label might have a sealed tear-off portion that would be filed in the pharmacy or with the participant's records. In an emergency, the seal could be torn and the drug identity revealed. In one study, the sealed labels attached to the medication bottles were transparent when held up to strong light. Care should be taken to ensure that the sealed portion is of appropriate color and thickness to prevent reading through it.

Official breaking of the blind may be necessary. There are bound to be situations that require disclosures, especially in long-term studies. Perhaps the study drug requires tapering the dosage. Children may get hold of study pills and swallow them. In an emergency, knowledge that a participant is or is not on the active drug would indicate whether tapering is necessary. Usually, most emergencies can be handled by withdrawing the medication without breaking the blind. When the treating physician is different from the study investigator, a third party can obtain the blinded information from the pharmacy or central data repository and relate the information to the treating physician. In this way, the participant and the study investigator need not be unblinded. Knowledge of the identity of the study intervention seldom influences emergency care of the participant. This information is important for treating physicians to know since it can help reduce the frequency of unnecessary unblinding. When unblinding does occur, the investigator should review and report the circumstances which led to it in the results paper.

In summary, double-blind trials require careful planning and constant monitoring to ensure that the blind is maintained and that participant safety is not jeopardized.

Inadvertent Unblinding

The phrase "truly double-blind study" was used earlier. While many studies are designed as double- or single-blind, it is unclear how many, in fact, are truly and completely blind. Drugs have side effects, some of which are fairly characteristic. Known pharmaceutical effects of the study medication may lead to unblinding. Inhalation of short-acting beta-agonists causes tremor and tachycardia within minutes in most users. Even the salt of the active agent can cause side effects that lead to unblinding. For example, the blinded design was broken in a clinical trial comparing the commonly used ranitidine hydrochloride to a new formulation of ranitidine bismuth citrate. The bismuth-containing compound colored the tongue of its users black [34]. Rifampin, a treatment for tuberculosis, causes the urine to change color. Existence or absence of such side effects does not necessarily unblind drug assignment, since not all people on drugs do develop reactions and some people on placebo develop events which can be mistaken for drug side effects. In trials of warfarin vs. oral anticoagulants, healthcare providers often check the INR in the event of significant bleeding. Since it is elevated with warfarin, this is likely to lead to unblinding. It is well known that aspirin is associated with gastrointestinal problems. In the Women's Health Study [35], 2.7% of the participants in the low-aspirin group developed peptic ulcer. On the other hand, 2.1% of the placebo participants had the same condition. This difference is highly significant ($p < 0.001$), but for a participant having an ulcer, in itself, would not unblind.

Occasionally, accidental unblinding occurs. In some studies, a special center labels and distributes drugs to the clinic where participants are seen. Obviously, each carton of drugs sent from the pharmaceutical company to this distribution center must contain a packing slip identifying the drug. The distribution center puts coded labels on each bottle and removes the packing slip before sending the drugs to the investigator. In one instance, one carton contained two packing slips by mistake. The distribution center, not realizing this, shipped the carton to the investigator with the second packing slip enclosed.

Laboratory errors have also occurred. These are particularly likely when, to prevent unblinding, only some laboratory results are given to the investigators. Occasionally investigators have received the complete set of laboratory results. This usually happens at the beginning of a study before "bugs" have been worked out, or when the laboratory hires new personnel who are unfamiliar with the procedures. If a commercial laboratory performs the study determinations, the tests should be done in a special area of the laboratory, with safeguards to prevent study results from getting intermingled with routine work. Routine laboratory panels obtained during regular clinical care of patients may include laboratory results that could lead to unblinding. In a large, long-term trial of a lipid-lowering drug, the investigators were discouraged from getting serum cholesterol determination on their coronary patients. It is difficult to know how many complied.

In addition, monitoring the use of study medication prescribed outside the study is essential. Any group differences might be evidence of a deficiency in the blind.

Another way of estimating the success of a double-blind design is to monitor specific intermediate effects of the study medication. The use of platelet aggregation in the Aspirin Myocardial Infarction Study is an example. An unusually large number of participants with non-aggregating platelets in the placebo group would raise the suspicion that the blind had been broken.

Assessment and Reporting of Blinding

The importance of blinding in avoiding bias is well established in clinical trials. However, the assessment and reporting of blinding do not always receive proper attention. Readers of trial reports are often given incomplete information about the type of blinding during the trial. This is a potential concern since randomized trials with inadequate blinding, on average, show larger treatment effects than properly blinded trials [36].

In their systematic review of 819 articles of blinded randomized trials assessing pharmacologic treatment, Boutron et al. [25] considered three blinding methods—(1) the initial blinding of participants and investigators, (2) the maintenance of this blinding and, (3) the blinding of those assessing trial outcomes. Overall, only 472 of the blinded reports (58%) described the method of blinding, while 13% gave some information and 29% none at all. The methods to establish blinding were presented in 41% of the reports. These included different types of matching, the use of a "double-dummy" procedure, sham interventions and masking of the specific taste of the active treatments. The methods for blinded assessment were described in 14% of the reports. They are especially useful in trials when blinding cannot be established. The main method was a centralized assessment of the primary outcome by blinded classification committees.

In a survey of 191 placebo-controlled double-blind trials published in 1998–2001, the authors evaluated how often the success of blinding was reported [28]. Only 15 (8%) reported evidence of success, and of these 15 trials, blinding was imperfect in nine. A similar survey of 1,599 blinded randomized trials from 2001 reported that only 2% of the trials reported tests for the success of blinding [37]. Interestingly, many investigators had conducted, but not published such tests. A report on the quality of reporting of randomized clinical trials in medical oncology between 2005 and 2009 according to the 2001 CONSORT statement showed that numerous items remained unreported. Only 41% of the 347 trials clearly reported whether and how blinding was applied [38]. A similar review of 442 trials in the psychiatric literature showed that the reporting of how blinding was accomplished and evaluated decreased following the publication of the CONSORT statement [39].

Although the success of blinding may be important, it is not easy to access and few trial publications provide this information. If done, there are different views as to whether and when to assess blinding—early after randomization, throughout the trial or at the end [40]. Early assessment in a double-blind trial would be a measure of the initial success of blinding. Repeated questioning may trigger the curiosity of

the study participants. Assessment at the end "confounds failures in blinding with successes in pre-trial hunches about efficacy" [41]. If study participants do well, there is a tendency for them to predict that they received active treatment; if they have suffered events or perceived no improvement, their prediction is more likely to be placebo. Similarly, investigators' hunches about efficacy can also be influenced by their preconceived expectations as illustrated by Sackett [42]. He concluded that "We neither can nor need to test for blindness during and after trial,...". The CONSORT 2010 statement eliminated the 2001 recommendation to assess how the success of blinding was assessed [43].

The CONSORT 2010 [43] and the SPIRIT 2013 [44] guidelines have a checklist of items recommended to be included in trial protocols Both have a similar item for blinded trials asking who was blinded after assignment to interventions (e.g. participants, care providers, outcome assessors, data analysts) and how. The former has a second item asking for "if relevant, description of the similarity of interventions". The latter has a second item asking for a description of "If blinded, circumstances under which unblinding is permissible and procedure for revealing a participant's allocated intervention during the trial".

Debriefing of Participants

Typically, participants randomized to blinded trials are never informed which treatment they received [39]. Various reasons are given by the investigators for not debriefing trial participants. There is strong and consistent evidence from recent large randomized clinical trials that trial participants would welcome being told which treatment they received. Whether investigators have ethical or other obligations to provide feedback is debated. No guidelines exist regarding debriefing of treatment received. However, it has been emphasized that clinically useful findings ought to be shared.

Debriefing to placebo allocation appears to raise more issues. Three theoretical concerns have been brought up: first, participants who benefited from placebo treatment might relapse on debriefing; second, the debriefing may engender mistrust and harm future doctor-patient relationships; and, third, the debriefing may have negative consequences for participants. However, the support for these concerns has been mixed [45].

A survey of participants in 14 randomized clinical trials in Parkinson's disease reported that 54% remembered being surprised or shocked [46]. Twenty-eight percent felt "disappointed". However, the respondents were overall positive and, most importantly, were willing to consider participating in future trials.

We favor that the investigators debrief the trial participants in person at the trial completion about the trial findings and their treatment group assignments. This ought to be part of transfer of post-trial care (see Chap. 20 on Closeout).

References

1. Mausner JS, Bahn AK. *Epidemiology: An Introductory Text*. Philadelphia: W.B. Saunders, 1974.
2. Kaptchuk TJ. Intentional Ignorance: A history of blind assessment and placebo controls in medicine. *Bull Hist Med* 1998;72:389–433.
3. Amberson JB, Jr, McMahon BT, Pinner M. A clinical trial of sanocrysin in pulmonary tuberculosis. *Am Rev Tuberc* 1931;24:401–435.
4. Devereaux PJ, Manns BJ, Ghali WA, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA* 2001;285:2000–2003.
5. Viergever RF, Ghersi D. Information on blinding in registered records of clinical trials. *Trials* 2012;13:210.
6. Haahr MT, Hrobjartsson A. Who is blinded in randomized clinical trials? A study of 200 trials and a survey of authors. *Clin Trials* 2006;3:360–365.
7. Park J, White AR, Stevinson C, Ernst E. Who are we blinding? A systematic review of blinded clinical trials. *Perfusion* 2001;14:296–304.
8. Hansson L, Hedner T, Dahlöf B. Prospective randomized open blinded end-point (PROBE) study. A novel design for intervention trials. Prospective Randomized Open Blinded End-Point. *Blood Press* 1992;1:113–119.
9. The Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. *Control Clin Trials* 1998;19:61–109.
10. Karlowski TR, Chalmers TC, Frenkel LD, et al. Ascorbic acid for the common cold. A prophylactic and therapeutic trial. *JAMA* 1975;231:1038–1042.
11. Lewis TL, Karlowski TR, Kapikian AZ, et al. A controlled clinical trial of ascorbic acid for the common cold. *Ann NY Acad Sci* 1975;258:505–512.
12. European Coronary Surgery Study Group. Coronary-artery bypass surgery in stable angina pectoris: survival at two years. *Lancet* 1979;*i*:889–893.
13. Schechter PJ, Friedewald WT, Bronzert DA, et al. Idiopathic hypogeusia: a description of the syndrome and a single-blind study with zinc sulfate. *Int Rev Neurobiol* 1972;(suppl 1):125–140.
14. Henkin RI, Schechter PJ, Friedewald WT, et al. A double blind study of the effects of zinc sulfate on taste and smell dysfunction. *Am J Med Sci* 1976;272:285–299.
15. Halperin JL, and The Executive Steering Committee on behalf of the SPORTIF III and V Study Investigators. Ximelagatran compared with warfarin for prevention of thromboembolism in patients with nonvalvular atrial fibrillation: Rationale, objectives, and design of a pair of clinical studies and baseline patient characteristics (SPORTIF III and V). *Am Heart J* 2003;146:431–438.
16. Bhatt DL, Kandzari DE, O'Neill WW, et al. for the SYMPLICITY HTN-3 Investigators. A controlled trial of renal denervation for resistant hypertension. *N Engl J Med* 2014;370:1393–1401.
17. Køber L, Torp-Pedersen C, McMurray JJV, et al. for the Dronedarone Study Group. Increased mortality after dronedarone therapy for severe heart failure. *N Engl J Med* 2008;358:2678–2682.
18. Hertzberg V, Chimowitz M, Lynn M, et al. Use of dose modification schedules is effective for blinding trials of warfarin: evidence from the WASID study. *Clin Trials* 2008;5:25–30.
19. Coumadin Aspirin Reinfarction Study (CARS) Investigators. Randomised double-blind trial of fixed low-dose warfarin with aspirin after myocardial infarction. *Lancet* 2008;350:389–396.
20. SPORTIF Executive Steering Committee for the SPORTIF V Investigators. Ximelagatran vs warfarin for stroke prevention in patients with nonvalvular atrial fibrillation. A randomized trial. *JAMA* 2005;293:690–698.
21. Tashkin DP, Celli B, Senn S, et al. for the UPLIFT Study Investigators. A 4-year trial of tiotropium in chronic obstructive pulmonary disease. *N Engl J Med* 2008;359:1543–1554.
22. Wittes J, Boissel J-P, Furberg CD, et al. Stopping the Randomized Aldactone Evaluation Study early for efficacy. In DeMets DL, Furberg CD, Friedman LM, (eds.). *Data Managing in Clinical Trials. A Case Study Approach.* New York: Springer, 2006, pp. 148–157.

23. Page SJ, Persch AC. Recruitment, retention, and blinding in clinical trials. *Am J Occup Ther* 2013;67:154–161.
24. Aspirin Myocardial Infarction Study Research Group. A randomized, controlled trial of aspirin in persons recovered from myocardial infarction. *JAMA* 1980;243:661–669.
25. Boutron I, Guittet L, Estellat C, et al. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. *PLoS Med* 2007;4:370–378.
26. von Wolff A, Hölzel LP, Westphal A, et al. Selective serotonin reuptake inhibitors and tricyclic antidepressants in the acute treatment of chronic depression and dysthymia: a systematic review and meta-analysis. *J Affect Disord* 2013;144:7–15.
27. Hill LE, Nunn AJ, Fox W. Matching quality of agents employed in "double-blind" controlled clinical trials. *Lancet* 1976;*i*:352–356.
28. Fergusson D, Glass KC, Waring D, Shapiro S. Turning a blind eye: the success of blinding reported in a random sample of randomised, placebo controlled trials. *Br Med J* 2004;328:432.
29. Farr BM, Gwaltney JM Jr. The problems of taste in placebo matching: an evaluation of zinc gluconate for the common cold. *J Chronic Dis* 1987;40:875–879.
30. Fai CK, De Qi G, Wei DA, Chung LP. Placebo preparation for the proper clinical trial of herbal medicine – requirements, verification and quality control. *Recent Pat Inflamm Allergy Drug Discov* 2011;5:169–174.
31. DeKosky ST, Williamson JD, Fitzpatrick AL, et al. for the Ginkgo Evaluation of Memory (GEM) Study Investigators. *Ginkgo biloba* for prevention of dementia. A randomized controlled trial. *JAMA* 2008;300:2253–2262.
32. Mathew NT, Schoenen J, Winner P, et al. Comparative efficacy of eletriptan 40 mg versus sumatriptan 100 mg. *Headache* 2003;43:214–222.
33. The CAPS Investigators. The Cardiac Arrhythmia Pilot Study. *Am J Cardiol* 1986;57:91–95.
34. Pacifico L, Osborn JF, Anania C, et al. Review article: bismuth-based therapy for helicobacter pylori eradication in children. *Aliment Pharmacol Ther* 2012;35:1010–1026.
35. Ridker PM, Cook NP, Lee I-M, et al. A randomized trial of low-dose aspirin in the primary prevention of cardiovascular disease in women. *N Engl J Med* 2005;352:1293–1304.
36. Schulz KF, Chalmers I, Hayes R, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408–412.
37. Hróbjartsson A, Forfang E, Haahr MT, et al. Blinded trials taken to the test: an analysis of randomized clinical trials that report tests for the success of blinding. *Int J Epidemiol* 2007;36:654–663.
38. Peron J, Pond GR, Gan HK, et al. Quality of reporting of modern randomized controlled trials in medical oncology: a systematic review. *J Natl Cancer Inst* 2012;104:982–989.
39. Han C, Kwak K, Marks DM, et al. The impact of the CONSORT statement on reporting of randomized clinical trials in psychiatry. *Comtemp Clin Trials* 2009;30:116–122.
40. Boutron I, Estellat C, Ravaud P. A review of blinding in randomized controlled trials found results inconsistent and questionable. *J Clin Epidemiol* 2005;58:1220–1226.
41. Sackett DL. Turning a blind eye. Why we don't test for blindness at the end of our trials. (Letter) *Br Med J*;328:1136.
42. Sackett DL. Commentary: Measuring the success of blinding in RCTs: don't, must, can't or needn't? *Int J Epidemiol* 2007;36:664–665.
43. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *BMJ* 2010;340:698–702.
44. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;346:e7586.
45. Bishop FL, Jacobson EE, Shaw J, Kaptchuk TJ. Participants' experiences of being debriefed to placebo allocation in a clinical trial. *Qual Health Res* 2012;22:1138—1149.
46. Goetz CG, Janko K, Blasucci L, Jaglin JA. Impact of placebo assignment in clinical trials of Parkinson's disease. *Movement Disorders* 2003;18:1146–1149.