

Blinding in randomised trials: hiding who got what

Kenneth F Schulz, David A Grimes

Blinding embodies a rich history spanning over two centuries. Most researchers worldwide understand blinding terminology, but confusion lurks beyond a general comprehension. Terms such as single blind, double blind, and triple blind mean different things to different people. Moreover, many medical researchers confuse blinding with allocation concealment. Such confusion indicates misunderstandings of both. The term blinding refers to keeping trial participants, investigators (usually health-care providers), or assessors (those collecting outcome data) unaware of the assigned intervention, so that they will not be influenced by that knowledge. Blinding usually reduces differential assessment of outcomes (information bias), but can also improve compliance and retention of trial participants while reducing biased supplemental care or treatment (sometimes called co-intervention). Many investigators and readers naively consider a randomised trial as high quality simply because it is double blind, as if double-blinding is the sine qua non of a randomised controlled trial. Although double blinding (blinding investigators, participants, and outcome assessors) indicates a strong design, trials that are not double blinded should not automatically be deemed inferior. Rather than solely relying on terminology like double blinding, researchers should explicitly state who was blinded, and how. We recommend placing greater credence in results when investigators at least blind outcome assessments, except with objective outcomes, such as death, which leave little room for bias. If investigators properly report their blinding efforts, readers can judge them. Unfortunately, many articles do not contain proper reporting. If an article claims blinding without any accompanying clarification, readers should remain sceptical about its effect on bias reduction.

The rich history of blinding in clinical trials spans a couple of centuries.¹ Most researchers worldwide appreciate its meaning. Unfortunately, beyond that general appreciation lurks confusion. Terms such as single-blind, double-blind, and triple-blind mean different things to different people.² Moreover, many medical researchers confuse the term blinding with allocation concealment. The fact that such confusion arises suggests that both terms are misunderstood. Clear theoretical and practical differences separate the two. Blinding prevents ascertainment bias and protects the sequence after allocation.^{3,4} By contrast, researchers use methods of allocation concealment primarily to prevent selection bias and to protect an assignment sequence before and until allocation. Furthermore, in some trials, blinding cannot be successfully implemented, whereas allocation concealment can always be successfully implemented.^{4,5}

Blinding represents an important, distinct aspect of randomised controlled trials.³ The term blinding refers to keeping trial participants, investigators (usually health-care providers), or assessors (those collecting outcome data) unaware of an assigned intervention, so that they are not influenced by that knowledge. Blinding prevents bias at several stages of a trial, although its relevance varies according to circumstance. Although initial forays into blinding might have used a blindfold,¹ the processes have now become much more elaborate. In this article, we focus on the attributes and benefits of blinding.

Potential effects of blinding

If participants are not blinded, knowledge of group assignment can affect responses to the intervention

Lancet 2002; **359**: 696–700

Family Health International, PO Box 13950, Research Triangle Park, NC 27709, USA (K F Schulz PhD, D A Grimes MD)

Correspondence to: Dr Kenneth F Schulz (e-mail: KSchulz@fhi.org)

received.³ Participants who know that they have been assigned to a group who will receive a new treatment might harbour favourable expectations or increased apprehension. Those assigned to standard treatment, however, might feel deprived or relieved. Despite evidence to suggest that new treatments are as likely to be worse as they are to be better than standard treatments,⁶ participants probably assume that new treatments will be better than standard treatments—new means improved. In any case, knowledge of the intervention received, and perceptions of that treatment, can affect the psychological or physical responses of the participants. Knowledge of treatment allocation can also affect compliance and retention of trial participants (panel 1).

Blinding investigators—those who contribute to a broadly defined trial team including, but not limited to, trial designers, participant enrollers, randomisation implementors, health-care providers, intervention counsellors, and routine data collectors—is also important.³ Investigators especially pertinent to blinding include health-care providers (such as an attending physician or nurse) and intervention counsellors—eg, someone who delivers a behavioural prevention message—who might interact with the participants throughout the trial. If investigators are not blinded, their attitudes for or against an intervention can be directly transferred to participants.⁷ Their inclinations could also be manifested in differential use of ancillary interventions of supplemental care or treatment (co-interventions), differential decisions to withdraw participants from a trial, or differential adjustments to the medication dose (panel 1). Investigators might also encourage or discourage continuation in a trial on the basis of knowledge of the intervention group assignment.

Perhaps most importantly, blinding helps to reduce differential assessment of outcomes (often called information or ascertainment bias) (panel 1). For example, if outcome assessors who know of the treatment allocation believe a new intervention is better than an old

Panel 1: Potential benefits accruing dependent on those individuals successfully blinded

Individuals blinded	Potential benefits
Participants	Less likely to have biased psychological or physical responses to intervention More likely to comply with trial regimens Less likely to seek additional adjunct interventions Less likely to leave trial without providing outcome data, leading to lost to follow-up
Trial investigators	Less likely to transfer their inclinations or attitudes to participants Less likely to differentially administer co-interventions Less likely to differentially adjust dose Less likely to differentially withdraw participants Less likely to differentially encourage or discourage participants to continue trial
Assessors	Less likely to have biases affect their outcome assessments, especially with subjective outcomes of interest

one they could register more generous responses to that intervention. Indeed, in a placebo-controlled trial in patients with multiple sclerosis⁸ the unblinded, but not the blinded, neurologists' assessments showed an apparent benefit of the intervention.

Subjective outcomes—eg, pain scores—present great opportunities for bias.³ Furthermore, some outcomes judged objective can be fraught with subjectivity, for example, salpingitis. In general, though, blinding becomes less important to reduce observer bias as the outcomes become less subjective, since objective (hard) outcomes leave little opportunity for bias. Knowledge of the intervention would not greatly affect measurement of a hard outcome, such as death.

Lexicon of blinding

Non-blinded (open or open label) denotes trials in which everyone involved knows who has received which interventions throughout the trial. Blinding (masking) indicates that knowledge of the intervention assignments is hidden from participants, trial investigators, or assessors.

The terminology single blind usually means that one of the three categories of individuals (normally participant rather than investigator) remains unaware of intervention assignments throughout the trial.⁹ A single-blind trial might also, confusingly, mean that the participant and investigator both know the intervention, but that the assessor remains unaware of it.

In a double-blind trial, participants, investigators, and assessors usually all remain unaware of the intervention assignments throughout the trial.³ In view of the fact that three groups are kept ignorant, the terminology double blind is sometimes misleading. In medical research, however, an investigator frequently also assesses, so in this instance the terminology accurately refers to two categories.

Triple blind usually means a double-blind trial that also maintains a blind data analysis.¹⁰ Some investigators, however, denote trials as triple-blind if investigators and assessors are distinct people and both, as well as participants, remain unaware of assignments. Investigators rarely use quadruple blind, but those that do use the term to denote blinding of participants, investigators, assessors, and data analysts.¹¹ Thus, quintuple blind must mean that the allocation schedule has been lost and nobody knows anything. Contrary to Mae West's claim that "too much of a good thing can be wonderful", such is not always the case in blinding.

Confused terminology of single, double, and triple blinding permeates the literature,³ with physicians, textbooks, and journal articles all offering different interpretations and definitions.² Not only do investigators

not define double-blind trials consistently, in particular, but they make matters worse by frequently failing to report their definitions clearly in their articles. Building on the original blindfolding efforts,¹ and the once common double blindfold terminology,¹² we further obfuscate by offering additional definitions of single and double blinding (figure 1). More seriously, when we use double-blind or its derivatives in this article, we mean that steps have been taken to blind participants, investigators, and assessors to group assignments. In reporting randomised controlled trials, we urge researchers to explicitly state what steps they took to keep whom blinded.

Sparse reporting on blinding, however, is common. Many investigators neglect to report whether or not their trial was blinded. For example, reports of 51% of 506 trials in cystic fibrosis,¹³ 33% of 196 trials in rheumatoid arthritis,¹⁴ and 38% of 68 trials in dermatology¹⁵ did not state whether blinding was used. When researchers have reported their study as double-blind, they frequently have not provided much further clarification.^{14,16-18} For example, of 31 double-blind trials in obstetrics and gynecology, only 14 (45%) reports indicated the similarity of the treatment and control regimens (for example, appearance, taste, administration) and only 5 (16%) provided statements to indicate that blinding was successful.¹⁸

Masking or blinding

Some people prefer the term masking to blinding to describe the same procedure. Masking might be more appropriate in trials that involve participants who have impaired vision, and could be less confusing in trials in which blindness is an outcome.³ Blinding, however, conveys a strong bias prevention message. Apparently, blinding terminology emerged when Benjamin Franklin and colleagues¹⁹ actually blindfolded participants to shield them from knowledge in their assessments of the



Figure 1: The authors: double blinded versus single blinded



Figure 2: **The authors blinded and masked**

therapeutic claims made for Mesmerism. The imagery of blindfolding, a total covering of the eyes, conveys stronger bias prevention than masking, where eye holes could permit viewing (figure 2). Blinding also suggests a more secure procedure to some. The International Conference on Harmonization (ICH) guidance,²⁰ for example, primarily uses blinding terminology. (The ICH is an intensive tripartite collaboration between regulatory authorities in Europe, Japan, and the USA to develop common guidelines for the design, implementation, and reporting of clinical trials). We prefer blinding because it has a long history, maintains worldwide recognition, creates strong imagery, and permeates the ICH guidelines.³

Placebos and blinding

Interventions (treatments) sometimes have no effect on the outcomes being studied.³ When an ineffective intervention is administered to participants in the context of a well-designed randomised controlled trial, however, beneficial effects on participants' attitudes sometimes occur, which in turn affect outcomes.¹⁰ Researchers refer to this phenomena as the placebo effect.

A placebo refers to a pharmacologically inactive agent that investigators administer to participants in the control group of a trial.³ The use of a placebo control group balances the placebo effect in the treatment group, allowing for independent assessment of the treatment effect. Although placebos can have a psychological effect, they are administered to participants in a trial because they are otherwise inactive. An active placebo is a placebo with properties that mimic the symptoms or side-effects—eg, dry mouth, sweating—that might otherwise reveal the identity of the (pharmacologically) active test treatment. Most researchers agree that placebos should be administered, whenever possible, to controls when assessing the effects of a proposed new treatment for a condition for which no effective treatment already exists.^{9,10} Indeed, blinding frequently necessitates the use of placebos.

However, a proven effective standard treatment, if such exists, is usually given to the control group for comparison against a new treatment.³ Thus, investigators might compare two active treatment groups without a placebo group. Even then, however, investigators frequently attempt to achieve blinding by use of the double-dummy method, in essence two placebos.^{11,21} For example, for comparison of two agents, one in a blue capsule and the other in a red capsule, the investigators would prepare blue placebo capsules and red placebo capsules. Then both treatment groups would receive a blue and a red capsule, one active and one inactive.

Does blinding prevent bias?

Some investigators, readers, and editors overstate the importance of blinding in prevention of bias. Indeed, some consider a randomised trial as high quality if it is double blind—ie, as if double blinding is the sine qua non of a randomised controlled trial.³ Unfortunately, scientific life is not that simple. A randomised trial can be methodologically sound and not be double blind or, conversely, double blind and not methodologically sound. Lasagna¹² captured that notion long ago: “Let us examine the placebo somewhat more critically, however, since it and ‘double blind’ have reached the status of fetishes in our thinking and literature. The Automatic Aura of Respectability, Infallibility, and Scientific Savoir-faire which they possess for many can be easily shown to be undeserved in certain circumstances.”¹² Although double blinding suggests a strong design, it is not the primary indicator of overall trial quality. Moreover, many trials cannot be double blinded. Such trials must, therefore, be judged on overall merit rather than an inapplicable standard based on double blinding.

We do not, however, suggest that blinding is unimportant.³ Intuitively, blinding should reduce bias, and available evidence supports that impression. Methodological investigations tend to show that double blinding prevents bias but is less important, on average, in prevention of bias than is adequate allocation concealment.^{4,22,23}

What to look for in descriptions of blinding

In general, if researchers describe a trial as double-blind, readers can assume that they have avoided bias. Empirical evidence lends support to this recommendation. As suggested in the CONSORT guidelines,^{24,25} however, investigators should not use only the single-blind, double-blind, or triple-blind terminology, but should also explicitly state who was blinded, and how. Moreover, if the researchers contend that the trial investigators, participants, and assessors were blinded—ie, double blind—then they should provide information about the mechanism (capsules, tablets, film, &c), similarity of treatment characteristics (appearance, taste, administration), and allocation schedule control—eg, location of the schedule during the trial, when the code was broken for the analysis, and circumstances under which the code could be broken for individual instances. Such additional information can lend support to or undermine claims of double-blinding (panel 2).^{26–29}

If researchers properly report their blinding efforts, readers can judge those efforts. Unfortunately, many articles will not contain proper reporting. If a researcher claims to have done a blinded study, but does not provide accompanying clarification, readers should remain sceptical about its effect on bias reduction. For example, one trial³⁰ of prophylactic antibiotics claimed to be blinded, but the methods section of the report revealed that little or no blinding occurred.

Ideally, researchers should also relate if blinding was successful. Investigators can theoretically assess the success of blinding by directly asking participants, health-care providers, or outcome assessors which intervention they think was administered (panel 3). In principle, if blinding was successful, these individuals should not be able to do better than chance when guessing the intervention, for example. In practice, however, blinding might be totally successful, but participants, health-care providers, and outcome assessors might nevertheless guess the intervention because of ancillary information. Disproportionate levels of adverse side-effects might

Panel 2: Descriptions of blinding

"No patient, research nurse, investigator, or any other medical or nursing staff in the ICU was aware of the treatment assignments for the duration of the study. All statistical analysis was also done with masking maintained. Randomisation authorities were instructed to report any suspected breach of the masking procedures. No report was filed . . . The drug or placebo (vehicle without active drug) was prepared for syringe pump infusion or for volumetric pump infusion in indistinguishable syringes or bags."²⁶

". . . in a double-blind, placebo-controlled manner . . . Neither the patients nor doctors could distinguish the placebo from sibutramine capsules. The taste of the capsules was identical provided they were swallowed whole as instructed. . . Results of biochemical analyses were completed before the randomisation code was broken at the end of the completed trial."²⁷

"The study was double-blinded—that is, neither the women nor the study staff, including the biostatisticians at Family Health International, knew which group was using the nonoxynol 9 film. The nonoxynol 9 film contained . . . The placebo film contained . . . The films were identical in appearance, packaging, and labeling."²⁸

"The doxycycline and placebo were in capsule form and identical in appearance . . . The randomization code was kept in the USA." (Note: the trial was conducted in Kenya) "Thus, all administration and assessments were done blinded to treatment assignment, and the investigators and patients were also blinded to the ongoing results of the study. The code was broken only after data collection had been completed."²⁹

ICU=intensive care unit.

provide strong hints as to the intervention. Irrespective of painstaking efforts to do double-blinded trials, some interventions have side-effects that are so recognisable that their occurrence will unavoidably reveal the intervention received to both the participants and the health-care providers.^{11,24} Even more fundamental than hints from adverse effects are the hints from clinical outcomes. Researchers usually welcome large clinical effects (except perhaps in equivalence trials). If they arise, health-care providers and participants would likely deduce—not always accurately of course—that a participant with a positive outcome received the active (new) intervention rather than control (standard). If indeed the active (new) intervention materialises as helpful (highly desirable) then their deductions would be correct more often than chance guesses.^{24,31} Irrespective of their suspicions, end-of-trial tests of blindness might actually be tests of hunches for adverse effects or efficacy.^{32,33}

Furthermore, individuals might be reluctant to expose any unblinding efforts by providing accurate responses to

Panel 3: Assessment of the success of blinding

"We asked 126 staff members their opinions of which film was the placebo. Eighteen percent thought film A (the placebo) was the placebo, 13 percent thought film B (nonoxynol 9) was the placebo, and 69 percent had no opinion about which film was the placebo. Of the 68 peer educators (the staff members most likely to reflect the opinion of the participants), 16 percent thought film A was the placebo, 13 percent thought film B was the placebo, and 71 percent had no opinion."²⁸

the queries—in other words, if they have deciphered group assignments, they might provide responses contrary to their deciphering findings to disguise their actions. That difficulty, along with interpretation difficulties stemming from adverse side-effects and successful clinical outcomes, leads us to question the usefulness of tests of blinding in some circumstances. Investigators should carefully consider the usefulness of assessing the success of their blinding efforts, but if they proceed, should provide the results of any assessments. At the very least, they should report any failure of the blinding procedure, such as non-identical placebo or active preparations. Published reports rarely contain assessments of blinding, but, if provided, readers should sceptically assess the information presented.

Double blinding proves difficult or impossible in many trials. For instance, in general, surgical trials cannot be double blinded. Specifically, a trial that compares degrees of pain associated with sampling blood from the ear or thumb cannot be double-blinded.³⁴ If researchers do not describe their trial as double-blind or the equivalent, it could still be scientifically strong. Apart from assessment of the other methodological aspects of the trial, readers would have to assess how much bias might have ensued due to absence of blinding. Readers should identify if anybody was blinded in the trial and what benefits might have accrued (panel 1). Indeed, blinding of outcome assessors is often possible and advisable, even in open trials.¹¹ For example, lesions can be photographed before and after treatment and assessed by someone not involved in the study.¹¹ We recommend placing greater credence in results when someone unaware of treatment assignments judges outcome measures.

Even that recommendation, however, is not absolute. As noted earlier, some hard outcomes, such as death, leave little room for ascertainment bias. In other words, blinding the assessor to hard outcomes might have little effect.

Conclusion

Blinding embodies a rich history spanning over two centuries. Most researchers worldwide understand blinding terminology, but confusion lurks beyond a general comprehension. Investigators should clearly explicate those blinded and not blinded in their trial, rather than only labeling their trial as single-blind, double-blind, or triple-blind. Readers should expect such clarity when reading and judging a trial report.

We thank Willard Cates and David L Sackett for their helpful comments on an earlier version of this report. Much of this material stems from our 15 years of teaching the Berlex Foundation Faculty Development Course.

References

- 1 Kaptchuk TJ. Intentional ignorance: a history of blind assessment and placebo controls in medicine. *Bull Hist Med* 1998; **72**: 389–433.
- 2 Devereaux PJ, Manns BJ, Ghali WA, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA* 2001; **285**: 2000–03.
- 3 Schulz KF, Chalmers I, Altman DG. The landscape and lexicon of blinding in randomized trials. *Ann Intern Med* 2002; **136**: 254–59.
- 4 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. *JAMA* 1995; **273**: 408–12.
- 5 Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA* 1994; **272**: 125–28.
- 6 Chalmers I. What is the prior probability of a proposed new treatment being superior to established treatments? *BMJ* 1997; **314**: 74–75.

- 7 Wolf S. Effects of suggestion and conditioning on action of chemical agents in human subjects: pharmacology of placebos. *J Clin Invest* 1950; **29**: 100–09.
- 8 Noseworthy JH, Ebers GC, Vandervoort MK, Farquhar RE, Yetisir E, Roberts R. The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology* 1994; **44**: 16–20.
- 9 Meinert CL. Clinical trials: design, conduct, and analysis. New York: Oxford University Press, 1986.
- 10 Pocock SJ. Clinical trials: a practical approach. Chichester: Wiley, 1983.
- 11 Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. *BMJ* 2000; **321**: 504.
- 12 Lasagna L. The controlled trial: theory and practice. *J Chronic Dis* 1955; **1**: 353–67.
- 13 Cheng K, Smyth RL, Motley J, O’Hea U, Ashby D. Randomized controlled trials in cystic fibrosis (1966–1997) categorized by time, design, and intervention. *Pediatr Pulmonol* 2000; **29**: 1–7.
- 14 Gøtzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Control Clin Trials* 1989; **10**: 31–56.
- 15 Adetugbo K, Williams H. How well are randomized controlled trials reported in the dermatology literature? *Arch Dermatol* 2000; **136**: 381–85.
- 16 Mosteller F, Gilbert JP, McPeck B. Reporting standards and research strategies for controlled trials: agenda for the editor. *Controlled Clin Trials* 1980; **1**: 37–58.
- 17 DerSimonian R, Charette LJ, McPeck B, Mosteller F. Reporting on methods in clinical trials. *N Engl J Med* 1982; **306**: 1332–37.
- 18 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. *BMJ* 1996; **312**: 742–44.
- 19 Franklin B, Bailly JS, Lavoisier A. Rapport des commissaires chargés par le roi, de l’examen du magnétisme animal. Nice: Chez Gabriel Floteron, 1785.
- 20 Department of Health and Human Services, FDA. International conference on harmonisation: guidance on statistical principles for clinical trials. *Fed Regist* 1998; **63**: 49583–98.
- 21 Altman DG. Practical statistics for medical research. London: Chapman and Hall, 1991.
- 22 Khan KS, Daya S, Collins JA, Walter SD. Empirical evidence of bias in infertility research: overestimation of treatment effect in crossover trials using pregnancy as the outcome measure. *Fertil Steril* 1996; **65**: 939–45.
- 23 Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998; **352**: 609–13.
- 24 Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001; **134**: 663–94.
- 25 Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports or parallel-group trials. *Lancet* 2001; **357**: 1191–94.
- 26 Bellomo R, Chapman M, Finfer S, Hickling K, Myburgh J. Low-dose dopamine in patients with early renal dysfunction: a placebo-controlled randomised trial. Australian and New Zealand Intensive Care Society (ANZICS) Clinical Trials Group. *Lancet* 2000; **356**: 2139–43.
- 27 James WP, Astrup A, Finer N, et al. Effect of sibutramine on weight maintenance after weight loss: a randomised trial. STORM Study Group. *Lancet* 2000; **356**: 2119–25.
- 28 Roddy RE, Zekeng L, Ryan KA, Tamoufe U, Weir SS, Wong EL. A controlled trial of nonoxynol 9 film to reduce male-to-female transmission of sexually transmitted diseases. *N Engl J Med* 1998; **339**: 504–10.
- 29 Sinei SK, Schulz KF, Lamptey PR, et al. Preventing IUCD-related pelvic infection: the efficacy of prophylactic doxycycline at insertion. *Br J Obstet Gynaecol* 1990; **97**: 412–19.
- 30 Baker KR, Drutz HP, Barnes MD. Effectiveness of antibiotic prophylaxis in preventing bacteriuria after multichannel urodynamic investigations: a blind, randomized study in 124 female patients. *Am J Obstet Gynecol* 1991; **165**: 679–81.
- 31 Quitkin FM, Rabkin JG, Gerald J, Davis JM, Klein DF. Validity of clinical trials of antidepressants. *Am J Psychiatry* 2000; **157**: 327–37.
- 32 Sackett DL, Gent M, Taylor DW. Tests for the blindness of randomized trials may not. *Clin Res* 1986; **34**: 711A.
- 33 The Canadian Cooperative Study Group. A randomized trial of aspirin and sulfinpyrazone in threatened stroke. *N Engl J Med* 1978; **299**: 53–59.
- 34 Carley SD, Libetta C, Flavin B, Butler J, Tong N, Sammy I. An open prospective randomised trial to reduce the pain of blood glucose testing: ear versus thumb. *BMJ* 2000; **321**: 20.