Fool’s gold? Why blinded trials are not always best


Published in:
BMJ

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
Copyright 2019 BMJ. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Fool’s gold? Why blinded trials are not always best

Blinding is intended to reduce bias but can make studies unnecessarily complex or lead to results that no longer address the clinical question, argue Rohan Anand and colleagues

Rohan Anand, doctoral research student1, John Norrie, professor2, Judy M Bradley, professor1, Danny F McAuley, clinical professor1, Mike Clarke, professor3

1Wellcome-Wolfson Institute for Experimental Medicine, School of Medicine, Dentistry, and Biomedical Sciences, Queen’s University Belfast, Belfast, UK; 2Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK; 3Northern Ireland Clinical Trials Unit and Methodology Hub, Centre for Public Health, Queen’s University Belfast, Belfast, UK

The essence of blinding is withholding information about treatment assignment from people involved in the trial. Trials in which patients, clinicians, and researchers are blinded to the allocated intervention are usually regarded as the gold standard of clinical research and evidence. However, blinding’s illustrious reputation brings with it the danger that it is regarded as essential for a trial to be “good,” especially if users place an uncritical reliance on hierarchies of evidence in which blinded evaluations are near the top. Given that the number of new trials is increasing every year, with 25 000 registered since the start of 2019, we are concerned that a substantial amount of time, energy, and funding may be going into considering and implementing blinding without a sound rationale for it.

Past, present, and future trials contain vast amounts of important data. If trials without blinding are inappropriately judged to be of lower quality than blinded trials then we may not be making best use of their data to improve healthcare, while blinded trials may be producing results that are more difficult to interpret than they need be. In this article, we seek to stimulate debate by challenging some of the prevailing beliefs on the benefits of blinding.

Purpose of blinding

The first recorded instance of blinding is from Paris in 1784. This was a study into an unknown, mystical force called mesmerism and the claims that patients could be healed simply by encountering a “mesmeriser” or an object that had been mesmerised. People believed to have been mesmerised would sometimes show hysterical behaviour, but the French Academy of Sciences was sceptical. It assembled a distinguished scientific team, including Antoine Lavoisier and Benjamin Franklin, to conduct experiments to test mesmerism. The participants were blinded to which objects were mesmerised or whether a mesmeriser was behind a curtain and then observed for their reactions. When they were blinded, mesmerism lost its power, and the study instead established the power of blinding.

Blinding is used in trials to reduce bias by ensuring that knowledge of which intervention a given trial participant received does not influence the judgments of trial participants or investigators. This allows the identification of the “true effect” of the new intervention, as distinct from any effect arising simply from the participant’s knowledge or expectation of receiving an intervention. In placebo controlled trials, any placebo effect of the new treatment would be discounted when comparing the intervention and control group to determine the effect of the active properties.

Blinding is also used to reduce bias in which the measured effect is not the true effect. Blinding aims to minimise response and observer bias. Response bias occurs when participants respond inaccurately, either intentionally or unintentionally. Observer bias occurs when researchers assessing the effects of the interventions have presumptions about them and so may inaccurately measure outcomes, leading to different effect estimates. Blinding also aims to minimise co-intervention bias, in which non-trial interventions may be taken differently by the groups being compared if participants know what they have been allocated.

Evidence on the ability of blinding to minimise the “placebo effect” (which can occur regardless of whether an actual placebo is used) and reduce bias comes from comparisons between selected trials and from systematic reviews of methodology research, with lack of blinding leading to an exaggerated treatment effect of up to 68%. However, there are several negative consequences that can arise from blinding.

Negative aspects of blinding

The substantial challenges of recruitment and retention in clinical trials have been highlighted as priorities for research. Poor recruitment leads to prolonged study times and underpowered results. These challenges are made worse by blinding, especially in trials using a placebo control. Trials with
nested components that were blinded and unblinded found that blinded designs discouraged people from participating.\textsuperscript{15-17} Key reasons given by patients for not wanting to enrol in these trials were that they wanted a named medication or wanted to know what was in the tablets. This suggests that achieving blinding and using a non-active comparator discourages people from joining a trial. In another study about 25% of patients expressed concerns about receiving placebos.\textsuperscript{18} Successful retention of patients is equally important,\textsuperscript{19} and the use of a placebo might be damaging if patients who suspect that they have been allocated to receive it withdraw from the study. A meta-analysis investigating retention in trials of antipsychotic interventions concluded that a placebo controlled design significantly increased dropout.\textsuperscript{20} Patient preference or resentful demoralisation can be a problem if patients in a placebo group lose motivation when they suspect or discover they are not receiving an active treatment. This could result in bias from differential loss to follow-up between groups.\textsuperscript{21} Patients may feel frustrated because they believe they are receiving inadequate treatment and so exaggerate negative answers on questionnaires or even withdraw from the trial.\textsuperscript{22}

The production and packaging of interventions in ways that will ensure their identity is blinded to participants, including use of placebo controls, can also cause difficulties. Not only do control interventions need to look identical to the intervention, any characteristic taste, texture, smell, colour, or viscosity of the intervention needs to be matched as well, which can be expensive.\textsuperscript{23} Money spent on blinding has opportunity costs if it reduces funding to optimise other features that would have more influence on the trial's robustness such as the training of trial staff, boosting the sample size, and comprehensively measuring outcomes.\textsuperscript{24} Even if the blinded control is designed to be physically identical to the intervention, any signature side effects associated with the intervention(s) may lead to unblinding.\textsuperscript{25} Examples come from the IMOP trial of isosorbide mononitrate for cervical ripening\textsuperscript{26} and the IMAGES trial of magnesium for acute stroke,\textsuperscript{27} both of which had high rates of specific side effects in the intervention arms; even though no formal unblinding of researchers and patients occurred, those involved in such trials may have had a good idea about the groups that patients were in.

In addition to this passive association, others might actively look for signs that they believe to be linked to the interventions. The online community group PatientsLikeMe was set up to enable people to share information on their illnesses. Members who were enrolled in blinded clinical trials shared their outcomes, including side effects, on online platforms outside of the official protocol or any trial regulations, even before the trial's completion.\textsuperscript{28} Their aim was to help each other deduce their allocated intervention, showing their frustration in the blinded approach. This highlights that maintaining blinding may be increasingly difficult in the age of social media and online networks.\textsuperscript{29} Researchers have also been found to break blinding by comparing pills and searching through the restricted notes of patients.\textsuperscript{30} Box 1 describes other problems that can arise from blinding.

### Box 1: Problems associated with blinding

#### Emergency unblinding

If an individual's allocation has to be unblinded for clinical reasons, there is the potential that this can cascade and unblind others in the trial. A simple example would be an adverse event needing treatment that is reported by blinded trial staff, who then code break to identify which intervention the patient received. Although the trial staff are officially unblinded to only this single case, they might now associate this event or related symptoms with the specific intervention. Even worse, if all the interventions had been coded in the same way (such as “drug A” and “drug B”) those who unblind themselves to one patient, effectively unblind themselves to all patients. Even in the absence of such coding, unblinding of patients in a trial using blocked randomisation might reveal the allocations of patients from the same block or strata.\textsuperscript{31}

#### Testing for blinding

Testing for the success of blinding in trials has been reported in about 2% of trials,\textsuperscript{32} usually by asking those blinded to guess treatment allocation.\textsuperscript{33} In theory, any significant difference over chance suggests that blinding was compromised. However, measuring blinding is highly challenging. Asking people to say which treatment was allocated after outcomes have been accumulated makes them likely to base their answer on assumptions related to the effects of the intervention. This was observed in a 2x2 factorial trial of aspirin and sulfinpyrazone for stroke prevention in which blinded clinicians were asked to guess treatment groups and did significantly worse than chance.\textsuperscript{34} Their guesses seemed to be influenced by their prior assumptions that sulfinpyrazone was more effective than aspirin and that patients who did well must have been on sulfinpyrazone, when in fact the trial showed the opposite.\textsuperscript{35} This essentially confounds testing for the success of blinding with expectations about treatment efficacy.

### Blinding and risks to patient safety

When blinding might compromise patient safety, it is paramount to consider whether it is necessary. For example, a placebo controlled trial of fibrinogen for postpartum haemorrhage required a moratorium on the use of any new treatments for 15 minutes after the randomly allocated treatment was given, with the sole purpose of maintaining the blind, potentially creating an unacceptable risk for the women.\textsuperscript{36} Similarly, adjusting doses creates problems in blinded trials and in such situations, using a fixed dose of a drug with a narrow and volatile therapeutic range could compromise patient safety. Clinical trials with anticoagulants\textsuperscript{37} and antipsychotics have been historically difficult to blind because of the need for dose adjustments.\textsuperscript{25} Use of a placebo or other sham therapy might lead to adverse effects that would not have happened if an open control group had been used. These could be direct harms from the procedures intended to ensure blinding, such as infection from piercing the skin to give a placebo injection or muscular problems from sham physiotherapy.

In considering these concerns about patient safety, Franklin G Miller outlined key questions that might help when deciding whether to use placebos in surgical trials.\textsuperscript{38-41} It seems reasonable to apply a similar but expanded set of questions, as listed in box 2, when considering using blinding in all clinical trials. These questions are context dependent and would be determined by those designing the trial; if the disadvantages outweigh the benefits in one of the questions then a blinded trial might not be appropriate.

#### Box 2: Questions to consider before using blinding

- Is blinding needed for a scientifically sound result? (Will the intervention blind have a placebo effect which needs to be separated from true effect?)
- How likely is it that patients or clinicians will behave differently if they know the intervention and would this change in behaviour bias the results?
- Are the potential harms to patients of using blinding excessive?
- Does the anticipated social value of the study results justify any potential harms of blinding?
- Does the financial cost of blinding compromise spending on other methodological aspects of trial integrity?
Pragmatism and what happens in the real world

At its simplest, a randomised trial is a comparative effectiveness study that aims to obtain as unbiased an estimate as possible of the difference in the outcomes for patients in the treatment group compared with those in the control group. Beyond this, the ultimate aim is to generate evidence that can be used to make assumptions about what will happen to future patients who receive the treatment after the trial. Blinding might help to reduce bias but hamper the evidence generated. Minimising biases with blinding might weaken the ability to predict the future accurately, because blinding is unlikely to be used in routine practice. There is a continuum from explanatory to pragmatic trials, and blinding influences where a trial is on this continuum.43

Some of the types of blinding that would be contemplated only in a research setting are inconsistent with the desire for pragmatism in large, phase III pragmatic effectiveness trials. Pragmatic trials strive to generate situations that are as close as possible to routine practice, when patients and clinicians will not be blinded to the intervention. Outside trial settings the intervention is known and this will have a legitimate effect on the patient, because blinding is unlikely to be used in routine practice. There is a continuum from explanatory to pragmatic trials, and blinding influences where a trial is on this continuum.43

Some of the types of blinding that would be contemplated only in a research setting are inconsistent with the desire for pragmatism in large, phase III pragmatic effectiveness trials. Pragmatic trials strive to generate situations that are as close as possible to routine practice, when patients and clinicians will not be blinded to the intervention. Outside trial settings the intervention is known and this will have a legitimate effect on the patient, because blinding is unlikely to be used in routine practice. There is a continuum from explanatory to pragmatic trials, and blinding influences where a trial is on this continuum.43

Methods that increase trial integrity

The prospective randomised open blinded endpoint evaluation (PROBE) is an established method for trials.44-47 It emphasises randomisation (with secure concealment until the allocation is revealed) and blinded outcome assessment, two facets that protect against bias. The blinding is implemented while evaluating defined endpoints during a trial. Trials using PROBE are regarded as open label with respect to patients and clinicians but implement the blinding of outcome assessors or the blinded evaluation of the trial’s endpoints. This approach of keeping outcome assessors blind to the random allocation can be used in most trials, including pragmatic effectiveness trials in which outcomes are either subjective or objective. Blinding the outcome assessors throughout a trial or using blinded evaluation of endpoints by a committee at set points, reduces the effect of observer and response bias, which can cause substantial reported differences between treatments. Such methods would increase rigour when double blinding of patients and clinicians is
dropped, might be simpler to deliver, and can avoid the many challenges we have outlined. For example, although the outcome assessor is blinded, the study replicates routine practice in that patients and clinicians know which intervention is being used for a particular participant.

If blinded outcome assessment cannot be used in a trial, bias can still be substantially reduced by using objective (eg, death) rather than subjective (eg, quality of life) outcomes. This is supported by a large meta-epidemiological study that found little evidence of bias in unblinded trials that used objective outcomes for both drug and non-drug interventions. Another option to reduce bias is to modify the outcome to make it less subjective. This can include avoiding surrogate markers and limiting the size of any effect on a given clinical measure (eg, using a 5 point Likert scale rather than 10 point Likert scale).

Blinding in clinical trials can increase the reliability of a trial's results but has consequences for the practicality, safety, and results of some trials. We suggest that the key elements for clinical trials seeking to minimise bias when comparing the effects of interventions should be adequate randomisation, allocation concealment, use of objective outcomes, independent blinded adjudication of outcomes, and, when possible, blinded assessment of outcomes. The traditional double blinding of participants and clinicians should not be regarded as a gold standard to strive for and should be used only if the negative effects are considered carefully and are outweighed by the potential benefits.

Key messages
- Blinding of participants, clinicians, and others avoids bias in clinical trials but can sometimes be detrimental to their integrity
- Some trials without blinding are appropriately judged as poor quality
- Blinding participants and clinicians can affect recruitment, retention, and applicability to routine practice as well as causing potential harm to patients
- Double blinded designs are not always ideal for providing a reliable answer to the trial's research question
- People using such designs should rationalise their use of blinding
- A more nuanced approach, using blinded outcome assessment and independent blinded adjudication of outcomes, alongside adequate randomisation and objective outcome measures, should reduce the main forms of bias

Contributors and sources: The authors have compiled extensive experience in the design, conduct, management, and analysis of clinical trials from clinical and methodological aspects. MC is director of the Northern Ireland Clinical Trials Unit and coordinating editor of the Cochrane Methodology Review Group with over 30 years’ experience in trials and systematic reviews. JN is chair of medical statistics and director of Edinburgh Clinical Trials Unit. JMB is a physiotherapist, director of the Wellcome Trust-Wolfson Northern Ireland Clinical Research Facility, and co-founder of the Northern Ireland Clinical Research Network for Respiratory Health. DFM&J is a consultant in intensive care medicine, director of the MRC/NHR efficacy and mechanism evaluation programme and co-director of research for the Intensive Care Society. RA’s PhD is exploring how trial methods, such as the use of placebo, can affect the outcomes. All authors contributed to conceptualisation and writing of the paper. RA prepared the original draft, and MC is guarantor.

Competing interests: All authors have read and understood BMJ policy on declaration of interests and declare no conflicts of interest.

Provenance and peer review: Not commissioned; externally peer reviewed.

3 Burns PB, Rohrich RJ, Chung KC. The levels of evidence and their role in evidence-based medicine. Plast Reconstr Surg 2011;128:305-10. 10.1097/PRS.0b013e31820e1f51

38 Sackett DL. Commentary: measuring the success of blinding in RCTs: don't, must, can't or needn't? Int J Epidemiol 2007;36:884-5. 10.1093/ije/dyq044 1765306


Published by the BMJ Publishing Group Limited. For permission to use (where not already granted under a licence) please go to http://group.bmj.com/group/rights-licensing/permissions.