Commentary: Measuring the success of blinding in RCTs: don’t, must, can’t or needn’t?

David L Sackett

Elsewhere in this issue, Asbjorn Hrobjartsson and his colleagues at the Nordic Cochrane Center present an analysis of the success of blinding in a random sample of entries in the Cochrane Central Register of Controlled Trials, adding to the recent analyses of ‘blinded’ trials from ‘top journals’ published by Dean Fergusson and his colleagues and of trials identified through MEDLINE, Cochrane registries, and ‘high-impact-factor journals’ by Isabelle Boutron and her colleagues (No article appeared in all three reviews, and only three articles appeared in two reviews.) The first two teams found that only 2 and 8% of the ‘blinded’ trials they unearthed had tested for blindness, respectively. Furthermore, in all three studies, these tests concluded that blinding was successful half of the time or less. Thus, the vast majority of ‘blinded’ trials don’t report testing whether study participants or those who care for or assess them can identify which treatment was given.

Must blinded trials test for blindness? The authors of these three reviews all favour improved testing and reporting, but none outright insist that it must be measured in every blinded trial. Moreover, the CONSORT checklist of items to include when reporting a randomized trial is being modified to remove the sentence: ‘If done, how the success of blinding was evaluated.’ Thus, none of these sources insist that all blinded trials must test for blindness.

This reluctance is well-founded. Testing for ‘blindness’ may not, and often can’t, generate valid answers. As soon as study participants begin to have events, they, their fellow participants who remain event-free, and care-givers for both groups begin to generate ‘hunches’ about the efficacy of the treatments being tested. This dawned on our research group 30 years ago in a factorial trial of aspirin and sulfinpyrazone for preventing stroke in patients with transient ischaemic attacks (TIAs). At the end of the trial, but before we told them its results, study neurologists were asked to predict both the overall study results and the regimens for each of their patients. With four regimens, we would expect blind clinicians to guess the correct one for 25% of their patients. As it happened, they did statistically significantly worse than chance, correctly identifying regimens for only 18% of their patients! Our faulty reasoning was exposed when we examined their predictions of the overall study results: they had predicted that sulfinpyrazone was efficacious and aspirin was not, precisely the reverse of the trial’s actual result. It then dawned on us that we were not testing our neurologists for blindness, but for their hunches about efficacy. When their patient had done well they tended to predict they were on (what they mistakenly thought was effective) sulfinpyrazone, and when they had done poorly, on placebo or (what they mistakenly thought was ineffective) aspirin. How fortunate for us all that their hunches were wrong. If they had been correct, the interpretation of our end-of-study test for blindness would be that they had broken the randomization code. End-of-trial tests for ‘blindness’ can’t be done with validity, because they can’t distinguish blindness from hunches about efficacy. Moreover, although it often is appropriate to test regimens before a trial to be sure that they are indistinguishable, testing for blindness early in a trial, before any events and the consequent hunches have occurred, can’t predict its status later in the trial.

Might it be possible that we needn’t test for blindness, and that efforts to do so are better spent elsewhere? This alternate view begins by relegating blinding to the level of process, rather than an end in itself. It then redirects our attention away from testing for blindness to testing for the bias-generating consequences of its loss. The first consequence of concern is contamination: this can occur when participants or their care-givers discover they are ‘controls’, and obtain the experimental treatment outside the trial, often ‘just to be sure’ that they get it. The second bias-generating consequence is co-intervention: this can occur when sympathetic care-givers provide non-study, but effective, interventions to participants they know to be controls. The third consequence is biased event reporting, and brings us back to ‘hunches’, but from the other direction. The participant who suspects or knows they are on the ‘better’ treatment may downplay symptoms and deny mild events. Similarly, the study clinician who suspects or knows a participant is on the ‘better’ treatment may downplay that participant’s symptoms and under-report ‘soft’ clinical findings and mild events.

Trout Research and Education Center at Irish Lake, Canada.
E-mail: sackett@bmns.com

Accepted 29 March 2007
Keywords Randomised trials, Blinding, Bias, Contamination, Co-intervention
Of course, these biases can arise from other causes. For example, contamination can arise from allocation errors, co-intervention from treatments received outside a trial, and all three can be the result of hunches about efficacy.

Unlike blindness, these bias-generating consequences can be measured. Moreover, they can be measured repeatedly, at any time during or after a trial, whenever they matter. For example, in our TIA trial, we repeatedly looked for contamination by carrying out periodic blood tests that searched for distinctive aspirin effects on platelets and distinctive sulfinpyrazone effects on serum uric acid. Because the detection and treatment of hypertension can sharply reduce the risk of stroke, in our carotid endarterectomy trial we tested for co-intervention by monitoring all participants’ blood pressures at each follow-up visit, pestering the clinicians caring for both groups to equalize the control of their hypertension. Finally biased outcome reporting could be examined by determining whether the degree of agreement between a blind adjudicator and a study clinician that a participant had suffered an outcome event was the same for experimental and control participants.

Trialists (and those who appraise their results) need more and better measures of these bias-generating consequences, whether they arise from the loss of blindness, the development of hunches, or any other cause. Might current efforts directed at measuring blindness be better spent devising and applying these measures?

We neither can nor need to test for blindness during and after trials, but we must bear in mind the bias-generating consequences that result from its loss. These are what will give trialists the vision they need to decide how confident they should be about their results.

Acknowledgements

Doug Altman, Iain Chalmers, Mike Clarke, Paul Glasziou, Gordon Guyatt, Brian Haynes, Paul Karanicolas, Wayne Taylor, and Salim Yusuf provided valuable feedback on a draft of this editorial.

Conflict of interest: I refereed the paper by Asbjorn Hrobjartsson and colleagues, and am a co-author of a later paper with Isabelle Boutron. I have been wined, dined, supported, transported, and paid to speak by countless pharmaceutical firms for over 40 years, beginning with two research fellowships and interest-free loans that allowed me to stay to finish medical school. Dozens of my randomized trials have been supported in part (but never in whole) by pharmaceutical firms, who never received or analysed efficacy data and never had veto power over any reports, presentations or publications of the results. I have twice worked as a paid consultant to advise pharmaceutical firms whether their products caused lethal side effects; on both occasions I told them ‘yes’. I have testified as an unpaid expert witness for a stroke victim who successfully sued a manufacturer of oral contraceptives, and as a paid expert witness in suits against manufacturers of female hormones and prosthetic heart valves. I was paid by a pharmaceutical firm to develop ‘levels of evidence’ for determining the causation of adverse drug reactions. My wife inherited and sold stock in a pharmaceutical company. While head of a division of medicine I enforced the banning of drug-detail personnel from clinical teaching units (despite the threat of withdrawal of drug industry funding for resident research projects). I received the Pharmaceutical Manufacturers’ Association of Canada Medal of Honour (and cash) for ‘Contributions to Medical Science in Canada’ for the decade 1984–94. My most recent award (the 2005 Baxter International Foundation Prize for Health Services Research) was sponsored by the Baxter International Foundation.

References


4 Schulz KF. Personal communication to the author, 23 February 2007.
