FEATURE ARTICLE

Assessing Blinding in Randomised Controlled Trials of Acupuncture: Challenges and Recommendations

Ann K Hopton and Hugh MacPherson

ABSTRACT The revision of CONSORT guidelines for reporting blinding in randomised controlled trials is the subject of controversy and criticism. To determine whether the criticism is justified, in this short communication paper we discuss the problems encountered in the methodology of the assessment of blinding, and the reporting of blinding in randomised controlled trials and the standards of reporting on blinding with reference to their usage in clinical trials of acupuncture for chronic pain. To conclude we recommend two simple guidelines: the development of sound clinical protocols that anticipate potential difficulties and reinforce overall internal validity, and secondly, the accurate reporting of the methodologies used to ensure a clear view of blinding procedures.

KEYWORDS blinding assessment, methodology, reporting, acupuncture

Introduction

The need for clarity and accuracy in reporting the design and methodology of randomised controlled trials is central for the scientific evaluation of the clinical efficacy of treatment\(^7\). In trials of acupuncture compared to sham acupuncture, we need to be certain that the different outcomes between treatment arms of the trial are due to differences in the treatment. The aim of blinding is to strengthen the internal validity of a trial by reducing the sources of bias brought about by patients and researchers’ expectations, prejudices and beliefs that are likely to create inconsistencies in the assessment of outcomes. Examples from systematic reviews of acupuncture for chronic pain\(^2-6\) indicate the challenges of assessing blinding. Unfortunately, blinding is not always successful. Unless the blinding is assessed, the quality of the trial cannot be established, as the readers cannot judge whether the results are due to the treatments received, or from a source of bias. Nevertheless, members of the CONSORT executive\(^7,8\) are now “unconvinced that all trialists should carry out a blinding assessment exercise”. The recent change in the CONSORT guidelines presents a Catch 22 situation to research. We want people blinded to allocation for internal validity but it is hard to establish if blinding itself is effective. If there is no necessity to report how blinding was assessed, how can we be sure that steps were taken to minimize sources of bias? The decision to eliminate the recommendation to report “how the success of blinding was evaluated” is controversial, and has been criticized as a philosophy of “let us give up because it is difficult to do or to interpret” especially when there are statistical methods available\(^9\). In this paper we aim to assess whether the criticism is justified and discuss what else can be done to confer internal validity.

Problems Encountered in the Assessment of Blinding

Without clear direction in the method of assessment, the evaluation of blinding has been erratic and beset by methodological problems. In a review of blinding in randomized controlled trials Boutron, Estellat and Ravaud\(^10\) found considerable uncertainty of how best to assess blinding. Consequently there has been a lack of consensus on quantitative procedures for evaluating the success of blinding in the literature\(^11\). The success of blinding is usually evaluated by direct questioning of the participants or by using a validated blinding index scale\(^12,13\). Subjects may be forced to guess...
their allocation, or allowed to express uncertainty, and in theory, blinding is believed successful if the proportion of guesses was no better than chance\(^{14,15}\). The problem with the forced guess method is that we cannot distinguish between those who respond because they are certain of their allocation and those who are genuinely guessing their allocation, and have no alternative. Therefore it weakens the confidence we have in the success of the blinding. However, even where the "do not know option" is allowed, the majority of trial reports examined exclude the data from the analysis\(^{16}\). The requirement to assess blinding in the previous CONSORT guidelines\(^{17,18}\) did not prescribe how blinding should be assessed or recommend any particular time point when assessment of blinding should take place\(^7\), and the amendments in the revised CONSORT statement\(^{8,19}\) still do not address these methodological issues.

**Treatment Effects**

The treatment effects and outcome may determine the success of blinding in several ways. Where the active treatment delivered a noticeable effect or good outcome, the participant or researcher is more likely to guess correctly their allocation from their treatment response. Similarly, distinctive side effects associated with the intervention may improve the likelihood of correctly guessing the allocation, and consequently compromise the integrity of the trial. Patients who experience minimal effect of the intervention and no side effects are less likely to guess their allocation to treatment and under these circumstances, blinding is more likely to be declared effective. Consequently we are faced with a circular argument where blinding is more likely to be declared inadequate if treatment is effective, and if the treatment did not work then the blinding is more likely to be declared effective. In addition, the present situation continues to ignore the question of how participants became unblinded.

**Timing**

The question of the optimum time to test for the integrity of blinding within a trial poses a major problem. Subjects may be asked to guess their allocation early on when treatment has started, or at time points midway or at the end. Sackett argues that a patient's hunch about allocation to treatment can have an influence at any potential time point for testing. Early testing before hunches about allocation develop does not predict the integrity of blinding later in the trial. Mid trial testing is affected by hunches that participants develop about their allocation as soon as treatment effects occur. Assessment of blinding at the end of a trial is compromised because researchers and participants cannot differentiate blindness from their hunches about efficacy\(^{20}\). In summary, there is never an optimum time to ask the participant to guess their allocation to treatment as means of assessing blinding status.

**Standards of Reporting**

The differences and difficulties that are experienced in determining how the assessment of blinding is conducted are likely to be a contributing factor in the poor quality of reporting and has also left the results open to interpretation\(^21\). Fergusson\(^{22}\) examined a random sample of 191 double blind randomized controlled trials and found that only 15 of trialists provided qualitative or quantitative information on methods of assessment. Of those 15, only five reported that blinding as successful, four assessed the success of blinding in participants, outcome assessors or the investigators, and three did not present supporting data. Hrobjartsson contacted 130 trial report authors to challenge their lack of reporting the methods and assessment of blinding\(^{16}\). Of the 130 contacted, 15 responded to confirm that assessment of blinding had taken place, 11 reported successful blinding, 2 reported partial loss of blinding, and 9 trial authors divulged that no formal statistical test has been used to assess the success of blinding. Based on their results, Fergusson, et al raise the call for an urgent need for a resolution to incorporate and report the success of blinding, and Hrobjartsson et al demand an urgent need to improve the methodology of testing the success of blinding\(^{22}\).

A key problem of reporting blinding methods accurately are the terms single, double and triple blinding. The terminology of single, double and triple blinding refers to methods of keeping the participants, investigators and assessors unaware of the allocated treatment. Although the terms are in common use, the lack of usage is not standardized and the terminology has been subject to misinterpretation. To overcome the problem, CONSORT 2010 have added the specification of how blinding was done by recommending reporting who was blinded after assignment to the interventions\(^{19}\) for example,