Published in final edited form as:
Community Dent Oral Epidemiol. 2009 December ; 37(6): 477–484. doi:10.1111/j.1600-0528.2009.00494.x.

Towards a proposal for assessment of blinding success in clinical trials: up-to-date review

Jafar Kolahi[1], **Heejung Bang**[2], and **Jongbae Park**[3]

[1]Torabinejad Research Center, Isfahan University of Medical Sciences, Isfahan, Iran

[2]Division of Biostatistics and Epidemiology, Department of Public Health, Weill Medical College of Cornell University, New York, NY, USA

[3]Department of Physical Medicine and Rehabilitation, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Abstract

The CONSORT statement recommended that investigators should clearly report which key trial persons were blinded to treatment allocation and test for the success of blinding. Clinical researchers, however, more often than not overlook the assessment of the success of blinding. The severe under-reporting on the success of blinding may improve with awareness of existing quantitative methods. The two statistical methods, James' blinding index (BI) and Bang's BI, are currently available. Subjects could be asked to guess their treatment assignment, possibly with an option to express the degree of certainty. Assessments of blinding at various points may serve different purposes, i.e. to evaluate comparability between experimental versus control treatments before the trial by the third party; to examine further comparability and credibility of the control treatment and patients' expectation about treatment received in early stage of the trial, and to summarize the overall maintenance of the blinding success at the end of the trial. In this article, we review BIs and how to use these methods along with discussion of other issues in blinding assessment and reporting. We contend the two BIs that were independently developed but carry complementary properties would characterize blinding behaviours in clinical trials qualitatively as well as quantitatively, and may also shed some lights on the interpretation of the study findings. Finally we urge the Item 11b of the CONSORT statement to be revised to require the assessment/reporting of blinding success for all trials that adopt blinding schemes.

Keywords

blinding index; methods; randomized controlled trials; standards

In randomized controlled trials (RCTs), blinding is widely accepted as an important methodological component to protect internal validity, and significant bias may result from unsuccessful blinding. Because of the importance of the blinding success, the CONSORT statement has incorporated this issue as one (Item 11b) of 22 items to be included when investigators report their RCT (1). However, it is unclear what constitutes adequate evidence of blinding success.

Jafar Kolahi, No 10, Sayt 180, Shahin Shahr, Isfahan, Co 83188 65161, Iran. Tel.: +98 9131081322, Fax: +98 3125252209, jafar.kolahi@gmail.com.

Conflict of interests: The authors declare that they have no conflict of interests.

Hrobjartsson et al. (2) reported 31 of 1599 medical RCTs (2%) and Fergusson et al. showed only 7 of 97 (7%) general medicine RCTs and 8 of 94 (8%) psychiatric RCTs; overall only 15 of the 191 (8%) provided evidence on the success of blinding (3). Clearly, a vast majority of blinded trials do not report test or sufficient evidence for the success of blinding. Therefore, there is a failure among investigators who undertake clinical trials and journals that publish the results in reporting the success of blinding. The current highly unsatisfactory reporting on the success of blinding may be partly due to the lack of appropriate knowledge or guidance in this field including how to collect data, if there are available methods, what they are and how to use them. Hence, in this article, we aimed to provide an up-to-date review of important issues in blinding in RCTs and methods for the assessment and tests for blinding and to propose a set of guidelines for these tasks.

How to report the success of blinding

Blinding methods

Investigators undertaking clinical trials are recommended to explicitly describe the methods of blinding they used. Boutron et al. reviewed blinding techniques in pharmacological and nonpharmacological treatments (4,5). These reviews classified blinding methods that could help investigators overcome some barriers to successful implementation and maintenance of blinding.

Who is blinded

It has been recommended that investigators should clearly report which of the following five categories of key trial persons were blinded to treatment allocation: (i) participants, (ii) healthcare providers, (iii) data collectors, (iv) outcome assessors (e.g. the data monitoring board or end-point committee or judicial assessors of outcome) and (v) data analysts (6). Furthermore, personnel writing the manuscript could be blinded to treatment allocation (possibly, those writing either of two drafts of a manuscript prior to breaking the randomization code, with draft 1 written assuming that group A is the treatment group and draft 2 assuming that group B is the treatment group) (7,8).

Authors frequently use the terms ‘single’, ‘double’ and ‘triple’ blind to describe the blinding strategy they adopted for their study. Indeed, these terms, particularly ‘double blind’, have become almost a matter of convention. Devereaux et al. have shown that both physicians and textbooks vary greatly in their interpretations and definitions of single, double and triple blinding (9). Haahr and Hrobjartsson also reported that only 2% of articles describing ‘double blind’ trials reported explicitly the blinding status of key trial persons (6). We believe explicit statements about who were actually blinded (i.e. the blinding status of key trial persons) should replace the use of current ambiguous or nonspecific terminologies (10).

Who is tested for the success of blinding

Investigators who undertake clinical trials are recommended to clearly report which of the five categories of key trial persons were tested. We must note that it does not mean that all key trial persons should be blinded and tested for the success of blinding. The level of blinding of a trial depends on the nature of the study such as the level of subjectivity of the outcome and precision needed for the study. However, Hrobjartsson et al. showed that 74% of trials tested only participants, 13% only data collectors, 10% both participants and data collectors, and 1% healthcare providers and data collectors (2).

Methods used to assess the success of blinding

Subjects could be asked to guess their treatment assignment, but they may be allowed to express treatment guess or uncertainty, i.e. subjects' guessing options could be ‘active versus

placebo (or control) versus do not know', we call this '2 × 3 format'. With successful blinding, we would expect the number of 'do not know' answers to be high or correct and incorrect guesses to be balanced. An example of this method is the study of Lao et al. (11).

Alternatively, the certainty of the guess could be rated on a 5-point scale (12). For example, LaRosa et al. asked subjects to rate the extent of certainty about their treatment on the following scale: 1) strongly believe the treatment is active, 2) somewhat believe the treatment is active, 3) somewhat believe the treatment is placebo, 4) strongly believe the treatment is placebo and 5) do not know (13); we call this '2 × 5 format'. An optional ancillary data can be additionally collected from the subjects who initially answered 'do not know' by re-asking them to choose one treatment (in 2×2 format). This datum can be useful for the purpose of validation of the 'do not know' data.

Another method is a pretrial evaluation of the potential for unblinding (14). This procedure involves conducting a test on an independent panel of volunteers who are not participating in the actual trial, in a preliminary study by asking them to try to distinguish different matching preparations for active versus control treatments.

Rees et al. described a fundamentally different method to ascertain the success of blinding (15). They emphasized that the assessment of belief regarding treatment assignment at a single cross-sectional session, especially at the end of the trial, does not capture the success or failure of blinding. This method shifts the focus away from the correctness of beliefs regarding treatment assignment; instead, it emphasizes the beliefs themselves and their change pattern during the trial. They concluded that the optimal means of assessing blinding is an inter-group comparison of the change in beliefs (and not necessarily their correctness) between the start and end of a trial.

Timing

The optimal timing and frequency of blinding assessments is a controversial issue. Some researchers believe that end-of-trial tests of blindness might actually be tests of hunches for efficacy or side effects rather than the success of blinding. This influence can be stronger if possible side effects have been mentioned as a part of the informed consent procedure (16). They recommend that the success of blinding should preferably be assessed in the early stages of the trial before the evidence of efficacy or side effects (17–19). On the other hand, some researchers recommend end-of-trial (or end of treatment administration if this is more relevant; for example, when the end point is mortality) tests of blinding. They believe unblinding can occur at any time during the study and testing at the end is more appropriate and will not or minimally affect the subject behaviours (20).

On the other hand, as blinding conveys stories during the entire course of the trial, a cross-sectional assessment will not be able to capture the dynamics of unblinding. Desbiens recommended the blinding test at several points during the study depending on the duration of the trial (21). Interestingly, Rees et al. (15) showed that the difference between the six-point assessment of blinding success during the trial and the two-point model was not significant. This is important because of concerns that repeated questioning can draw attention to the issue and may cause additional unblinding and bias.

In conclusion, we recommend assessments of blinding before the trial by the third party for independent credibility, in the early stage of the trial with great caution for credibility at specific trial setting, and at the end of the trial for overall assessments of trial participants.

Data analysis

Descriptive statistics of blinding data can be useful but may not be sufficient and a formal statistical analysis is encouraged. Hrobjartsson et al. found that 58% of RCTs that claimed the success of blinding presented no statistical analysis (2). Until now, some traditional methods such as the chi-squared test or the kappa statistic are often employed.

However, specialized statistical methods for blinding assessment, namely blinding index (BI), are available these days. James et al. (22) proposed a BI, as a variation of the kappa coefficient, that is sensitive not to the degree of agreement but to the degree of disagreement, by placing the highest weight on ‘do not know’ responses. This index ranges from 0 to 1, 0 being total lack of blinding, 1 being complete blinding and 0.5 being completely random blinding (i.e. 50% correct and 50% incorrect guesses). If the upper bound of the confidence interval (CI) of BI is below 0.5 (i.e. CI does not cover the null value), the study is regarded as lacking blinding. Otherwise, one may conclude that there is insufficient evidence for unblinding. An example of usage of this method is the study of Colford et al. (23).

An important assumption in the construction of James' BI is that when a respondent says that she or he does not know the treatment identity this represents an honest response, not just a socially desirable response or one that avoids making a judgement. It is therefore important that investigators encourage the respondents to report their suspicions honestly when such suspicions exist.

However, in practice, the drug and placebo arms of a clinical trial can exhibit distinct blinding behaviours not only in magnitude but also in direction. As James' BI is a single index value that combines blinding data from all arms, it cannot distinguish this difference, if any, between treatments and can even lead us to a misleading conclusion due to cancelling out of positive and negative estimates. Moreover, this and other standard methods do not provide the estimate of the proportion of unblinded participants beyond random chance level. Investigators may want to know how many participants are unblinded regardless of the validity and effectiveness of blinding.

Therefore, Bang et al. (24) proposed a new BI to address these issues. Bang's BI can be directly interpreted as the proportion of the unblinding in each arm and has the ability to detect different behaviours in different treatment arms including the ‘wishful thinking’ scenario (20). It takes a value between −1 and 1, with 0 as a null value, which indicates the most desirable situation under random blinding. A positive value may imply failure in blinding above random guessing (i.e. a majority of participants guess their treatment allocation correctly), and a negative value may suggest the success of blinding or failure of blinding in the other direction (i.e. more individuals mistakenly name the alternative treatment). An example of the usage of this method is the study of Park et al. (25). Furthermore, the Bang et al. approach allows us to distinguish nine possible blinding scenarios, determined by the combination of three possible results for each of the two arms (see Table. 1). The three possible results for each arm are ‘random’ (e.g. indicated by values of BI sufficiently close to 0), ‘unblinded’ (indicated by BI values significantly larger than 0) and ‘opposite’ (indicated by BI values significantly lower than 0). Alternatively, these scenarios can be classified based on *a priori* chosen (fixed) threshold value for BI (e.g. BI $\lesssim$0.2). Clinical investigators may be interested in knowing which category their trials fall into. Bang et al. emphasized estimation of BI rather than hypothesis testing in evaluating blinding success due to subjectivity and inherent confounding issues.

Computing modules for the two BIs are available for Stata (StataCorp LP, College Station, TX, USA) since March 2008, which are available at http://ideas.repec.org/c/boc/bocode/s456898.html (20,26).

Reporting of the results

Subjects' answers could be reported for relevant parties of the five categories of key trial persons who were tested for the success of blinding. Also subjects' answers could be reported for each trial arms. We recommend investigators report blinding data in 2 × 3 or 2 × 5 tables or adequate details that allow readers to reconstruct one of these tables. Table 3 in the study of Park et al. (25) is a good example.

Interpretation of the impact of blinding

The author's conclusion about the success of blinding and its ‘potential’ (not definite) effects or any implications on the study results should be described clearly. Hrobjartsson et al. showed that the conclusion was unclear or not reported in 32% of RCTs that reported blinding test results (2).

James' BI could be used to infer the totality of the blinding success in RCTs and hypothesis testing assuming ‘do not know’ is credible data. Bang's BI could be used to characterize and evaluate the blinding situation in each trial arm separately, classifying the result into one of nine qualitative different blinding scenarios, and estimating the percentage of unblinding beyond chance in each arm.

Most importantly, if significant unblinding occurred, investigators should try to find how (e.g. physical characteristics such as appearance, taste and smell, therapeutic or side effects, or accidental disclosure) and when this happened and to make efforts to avoid or minimize making the same mistakes or misconducts if possible in the current and future trials. For example, Hertzberg et al. (27) observed frequent unblinding in warfarin arm, compared to aspirin arm, in two RCTs and provided an explanation that the observed trend may be due to the number of dose change and haemorrhage, etc. in the warfarin arm.

Examples

Below are three examples of blinding data and tests in scientific research. The first two examples highlight the use of the 2 × 3 format and the third the use of the 2 × 5 format and shows usage of ancillary data.

The two other methods, pretrial evaluation of the potential for unblinding (14) and study of change pattern of beliefs during the trial (15), are not included in these examples. Indeed, they can provide a significant amount of blinding data. However, it is important to note that the omission of these methods of blinding assessment are limits to the given research examples. Ideally, pretrial and posttrial evaluations should be carried out in order to ensure the blinding, measure any amount of unblinding that took place and generally add to the scientific knowledge base.

Example 1

An RCT investigated the anti-plaque efficacy of chlorhexidine-containing chewing gum (28). Placebo was formulated to be identical in size, appearance, taste and package to the active gum. Study participants, data collector and data analysts were blinded to treatment allocation. In this study, healthcare providers and outcome assessors were not blinded. To assess participant blinding, they were asked to guess their treatment assignment (active, placebo or do not know) on the second day of each test period. The data were originally

analysed using McNemar's test, yet we re-analysed the results with the two BIs using Stata 10.0 (StataCorp LP) here. Responses of participants are summarized in Table 2. James' BI was 0.72 (95% CI 0.60–0.84), indicating the success of blinding (P = 0.999). Bang's BI for active intervention was 0.11 (95% CI −0.17 to 0.39) and for placebo it was 0.05 (95% CI −0.18 to 0.29), indicating the success of blinding in both active ($P = 0.261$) and placebo ($P = 0.352$) interventions.

Nevertheless, the assessment of blinding in this RCT involves some limitations including: (i) Pretrial evaluation of the potential for unblinding was not carried out; (ii) Blinding of data collector and data analysts were not examined; and (iii) End-of-trial tests of blindness were not conducted and the dynamics of unblinding was unclear.

Example 2

An RCT investigated the analgesic efficacy of acupuncture following surgical removal of mandibular third molar with partial bony impaction (29). Placebos were sham noninsertion adjacent to acupuncture points; distal sham shallow insertion at nonpoints and sham shallow insertion adjacent to acupuncture points; distal sham noninsertion. The surgeon (healthcare provider), research assistant (data collector), patient and statistician were blinded to group assignment. However, blindness of outcome assessors was not reported. To assess participant blinding, they were asked to guess their treatment assignment (acupuncture, placebo or uncertain) following acupuncture treatment. The blinding data originally were not analysed and presented as descriptive statistics. Responses of participants are reconstructed in Table 3. Based on our formal analyses, James' BI was 0.69 (95% CI 0.66–0.73), indicating the success of blinding ($P = 1$). Bang's BI for placebo arm was −0.11 (95% CI −0.19 to −0.02), indicating opposite guessing or successful blinding ($P = 0.987$). Yet, Bang's BI for active intervention arm was 0.39 (95% CI 0.29–0.48), indicating tendency to unblinding ($P < 0.001$). How to interpret the data? Totality of blinding success in this RCT may be accepted, yet it showed that subjects in both arms are more likely to report they received the active treatment. This situation may be classified as the ‘unblinding in active treatment and opposite guessing in placebo arm’ scenario according to the rule outlined in Table 1 and we found that this phenomenon is not uncommon in acupuncture studies (25). Interestingly, the authors reported, although no statistically significant analgesic effect was observed between the acupuncture and placebo groups, participants in both experiments who believed they received real acupuncture reported significantly less pain than patients who believed that they received a placebo. In order to understand this intriguing finding more objectively, further research should be underway with more blinding data from similar settings.

Example 3

This example shows the usage of the blinding data in the 2 × 5 format and ancillary data which can be additionally collected from the subjects who initially answered ‘do not know’. Data of this example are from a study of cholesterol medication in the elderly (13). At the end of the trial, participants were asked to rate the extent to which they were certain about their treatment assignment on the 2 × 5 format (Table 4). Also, subjects who answered ‘do not know’ were re-asked about their treatment guess one more time (Table 4). Results of statistical analysis of the primary blinding data are shown in Table 5, with or without ancillary data. Although conclusions were not changed, the BIs changed as expected when ancillary data were incorporated, indicating it is possible that, in some RCTs, ancillary data can be influential, especially when many subjects answered ‘do not know’ initially.

Implications for future RCTs

As recommended by Boutron et al. (12), we believe that a recommendation on the appropriate methods to assess the success of blinding should be added to the CONSORT statement. In accordance with Fergusson et al. (3), we would like to see Item 11b of the CONSORT statement revised to require the assessment of the blinding success for all blinded RCTs. Furthermore, the addition of 'test for blinding success' to item 9.4.6 of FDA guideline (please see FDA guidelines at: http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM129456.pdf) and section 4 of ADA clinical trial protocols (please see the ADA clinical trials protocols at: (http://www.ada.org/ada/seal/standards/guide_clinical_trial.pdf) are suggested. Also, we would like to see the other international regulatory organizations such as The European Agency for the Evaluation of Medicinal Products (EMEA), Human Subject Protections - Office of Human Subjects Research, NIH (OHSR), International Committee of Medical Journal Editors, etc. pay more attention to the blinding issues in RCTs that they sponsor or publish.

Moreover, researchers are faced with difficulties in identifying RCTs that assessed the success of blinding in the absence of MeSH terms (http://www.ncbi.nlm.nih.gov/sites/entrez?db=mesh) or consistent text words in international electronic databases. Addition of relevant MeSH terms would be a great step in the dissemination and generalization of assessment of blinding success. Finally, as recommended by Desbiense (30), we believe a test for blinding success should become a standard feature of RCTs whenever blinding is implemented in a trial in order to be designated a 'well-designed and well-conducted trial'. Investigators undertaking clinical trials, protocol reviewers and journal editors should make a concerted effort to incorporate, report and publish such information for good clinical research practice and eventually to understand its potential impacts on study results (i.e. treatment effectiveness/efficacy).

Conclusion

In conclusion, we suggest a set of guidelines for researchers in clinical trials.

- Clearly report methods of blinding.
- Clearly report which of the five categories of key trial persons were blinded to treatment allocation and tested for the success of blinding.
- Conduct a pretrial evaluation of the potential for unblinding by the third party.
- Ask subjects to guess their treatment assignment using three or five choices for response in early stage of the trial with great caution, and at the end of the trial.
- Report the subjects' answers for relevant parties of the five categories of key trial persons who were tested. Also report the subjects' answers for each trial arms. We recommend blinding data be presented in 2 × 3 or 2 × 5 tables.
- Analyse the data using James' (22) and Bang's (24) BIs.
- Carefully interpret the impact of blinding and its potential effects on the study results, if possible. One may use James' BI to conclude the totality of the blinding success of the RCT and hypothesis testing and use the Bang's BI to conclude the blinding situation in each trial arm, classifying the result into one of nine blinding scenarios, and to estimate the percentage of unblinding beyond chance in each arm.

- If significant unblinding occurred, try to find out how (e.g. physical characteristics such as appearance, taste and smell, therapeutic or side effects, or accidental disclosure) and when this happened.

Acknowledgments

We thank Isabel Canette, PhD, and Jiefeng Chen, MS, for development and free dissemination of a STATA module in the World Wide Web. We also thank Dr Anna Katie Smith for her linguistic revision. Heejung Bang was partially supported by a CTSC grant at Weill Cornell Medical College (UL1-RR024996).

References

1. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised consort statement for reporting randomized trials: Explanation and elaboration. Ann Intern Med. 2001; 134:663–94. [PubMed: 11304107]
2. Hrobjartsson A, Forfang E, Haahr MT, Als-Nielsen B, Brorson S. Blinded trials taken to the test: an analysis of randomized clinical trials that report tests for the success of blinding. Int J Epidemiol. 2007; 36:654–63. [PubMed: 17440024]
3. 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. BMJ. 2004; 328:432. [PubMed: 14761905]
4. Boutron I, Guittet L, Estellat C, Moher D, Hrobjartsson A, Ravaud P. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. PLoS Med. 2007; 4:e61. [PubMed: 17311468]
5. Boutron I, Estellat C, Guittet L, Dechartres A, Sackett DL, Hrobjartsson A, et al. Methods of blinding in reports of randomized controlled trials assessing pharmacologic treatments: a systematic review. PLoS Med. 2006; 3:e425. [PubMed: 17076559]
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; 3:360–5. [PubMed: 17060210]
7. Gotzsche PC. Blinding during data analysis and writing of manuscripts. Control Clin Trials. 1996; 17:285–90. discussion 90–3. [PubMed: 8889343]
8. Kolahi J, Abrishami M. Multiple-blind: towards a new blinding protocol for future generations of clinical trials. Med Hypotheses. 2009 May 30. Epub ahead of print. 10.1016/j.mehy.2009.04.040
9. Devereaux PJ, Manns BJ, Ghali WA, Quan H, Lacchetti C, Montori VM, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. JAMA. 2001; 285:2000–3. [PubMed: 11308438]
10. Park J, White AR, Stevinson C, Ernst E. Who are we blinding? A systematic review of blinded randomised controlled trials. Perfusion. 2001; 14:296–304.
11. Lao L, Bergman S, Hamilton GR, Langenberg P, Berman B. Evaluation of acupuncture for pain control after oral surgery: a placebo-controlled trial. Arch Otolaryngol Head Neck Surg. 1999; 125:567–72. [PubMed: 10326816]
12. Boutron I, Estellat C, Ravaud P. A review of blinding in randomized controlled trials found results inconsistent and questionable. J Clin Epidemiol. 2005; 58:1220–6. [PubMed: 16291465]
13. LaRosa JC, Applegate W, Crouse JR 3rd, Hunninghake DB, Grimm R, Knopp R, et al. Cholesterol lowering in the elderly. Results of the cholesterol reduction in seniors program (crisp) pilot study. Arch Intern Med. 1994; 154:529–39. [PubMed: 8122946]
14. Walter SD, Awasthi S, Jeyaseelan L. Pre-trial evaluation of the potential for unblinding in drug trials: a prototype example. Contemp Clin Trials. 2005; 26:459–68. [PubMed: 16054578]
15. Rees JR, Wade TJ, Levy DA, Colford JM Jr, Hilton JF. Changes in beliefs identify unblinding in randomized controlled trials: a method to meet consort guidelines. Contemp Clin Trials. 2005; 26:25–37. [PubMed: 15837450]
16. Schulz KF, Chalmers I, Altman DG. The landscape and lexicon of blinding in randomized trials. Ann Intern Med. 2002; 136:254–9. [PubMed: 11827510]

17. Henneicke-von Zepelin HH. Assessment of blinding in clinical trials. Contemp Clin Trials. 2005; 26:512. author reply 14–5. [PubMed: 15936250]
18. Hemila H. Assessment of blinding may be inappropriate after the trial. Contemp Clin Trials. 2005; 26:512–4. author reply 14–5. [PubMed: 15951244]
19. Sackett DL. Turning a blind eye: why we don't test for blindness at the end of our trials. BMJ. 2004; 328:1136. [PubMed: 15130997]
20. Park J, Bang H, Canette I. Blinding in clinical trials, time to do it better. Complement Ther Med. 2008; 16:121–3. [PubMed: 18534323]
21. Desbiens NA. In randomized controlled trials, should subjects in both placebo and drug groups be expected to guess that they are taking drug 50% of the time? Med Hypotheses. 2002; 59:227–32. [PubMed: 12208144]
22. James KE, Bloch DA, Lee KK, Kraemer HC, Fuller RK. An index for assessing blindness in a multicentre clinical trial: Disulfiram for alcohol cessation – a VA cooperative study. Stat Med. 1996; 15:1421–34. [PubMed: 8841652]
23. Colford JM Jr, Rees JR, Wade TJ, Khalakdina A, Hilton JF, Ergas IJ, et al. Participant blinding and gastrointestinal illness in a randomized, controlled trial of an in-home drinking water intervention. Emerg Infect Dis. 2002; 8:29–36. [PubMed: 11749745]
24. Bang H, Ni L, Davis CE. Assessment of blinding in clinical trials. Control Clin Trials. 2004; 25:143–56. [PubMed: 15020033]
25. Park J, White AR, James MA, Hemsley AG, Johnson P, Chambers J, et al. Acupuncture for subacute stroke rehabilitation: a sham-controlled, subject- and assessor-blind, randomized trial. Arch Intern Med. 2005; 165:2026–31. [PubMed: 16186474]
26. Chen, J. Software component provided by Boston College Department of Economics in its series Statistical Software Components with number S456898. 2008 January. Blinding: Stata module to compute blinding indexes.
27. Hertzberg V, Chimowitz M, Lynn M, Chester C, Asbury W, Cotsonis G. Use of dose modification schedules is effective for blinding trials of warfarin: evidence from the WASID study. Clin Trials. 2008; 5:23–30. [PubMed: 18283076]
28. Kolahi J, Soolari A, Ghalayani P, Varshosaz J, Fazilaty M. Newly formulated chlorhexidine gluconate chewing gum that gives both anti-plaque effectiveness and an acceptable taste: a double blind, randomized, placebo-controlled trial. J Int Acad Periodontol. 2008; 10:38–44. [PubMed: 18564727]
29. Bausell RB, Lao L, Bergman S, Lee WL, Berman BM. Is acupuncture analgesia an expectancy effect? Preliminary evidence based on participants' perceived assignments in two placebo-controlled trials. Eval Health Prof. 2005; 28:9–26. [PubMed: 15677384]
30. Desbiens NA. Lessons learned from attempts to establish the blind in placebo-controlled trials of zinc for the common cold. Ann Intern Med. 2000; 133:302–3. [PubMed: 10929173]

Table 1

Nine different blinding scenarios and their frequencies as reported in the Park et al. literature review (20)

Active group	Placebo group	Frequencies, *n* (%)	95% Confidence interval (%)[b]
Random guess	Random guess	10 (15.9)	8–27
Random guess	Opposite guess[a]	2 (2.3)	0.5–11
Random guess	Unblinded	5 (7.9)	2–18
Unblinded	Unblinded	18 (28.6)	18–41
Unblinded	Opposite guess[a]	11 (17.5)	9–29
Unblinded	Random guess	15 (23.8)	14–36
Opposite guess	Guess[a]	1 (1.6)	0.08–9
Opposite guess	Random guess	0 (0)	0–7
Opposite guess	Unblinded	1 (1.6)	0.08–9
Total		63 (100)	92–100

Random guess: −0.2 < BI < 0.2; unblinded: BI >0.2; opposite guess: BI <−0.2.

[a] Wishful thinking, RCT participants tend to think they are allocated to the active group even if not in reality.

[b] The confidence intervals were calculated with a correction for continuity.

Table 2

Assessment of blinding success in the study of Kolahi et al. (28)

	Subject's guess, *n* (%)				
Intervention	**Active**	**Placebo**	**Do not know**	**Total**	**Question could not be asked, *n* (%)**
Active	6 (16.6)	4 (11.1)	8 (22.2)	18 (50)	0
Placebo	3 (8.3)	4 (11.1)	11 (30.5)	18 (50)	0
Total	9 (25)	8 (22.2)	19 (52.7)	36 (100)	0

Table 3

Assessment of blinding success in the study of Bausell et al. (29)

	Subject's guess, *n* (%)				
Intervention	**Active**	**Placebo**	**Do not know**	**Total**	**Question could not be asked, *n* (%)**
Active	43 (14.3)	4 (1.3)	53 (17.6)	100 (33.3)	0
Placebo	61 (20.3)	39 (13.0)	100 (33.3)	200 (66.6)	0
Total	104 (34.6)	43 (14.3)	153 (5)	300 (100.0)	0

Table 4

Assessment of blinding success in the study of LaRosa et al. (13): main data (upper) and ancillary data (lower) for 'Do not know' responses

	Subject's guess, *n* (%)						
Intervention	**1**	**2**	**3**	**4**	**5**	**Total**	**Question could not be asked, *n* (%)**
Active	38 (9)	44 (10)	170 (40)	21 (5)	4 (1)	277 (66)	0
Placebo	11 (2)	16 (4)	83 (20)	21 (5)	8 (2)	139 (33)	0
Total	49 (12)	60 (14)	253 (61)	42 (10)	12 (3)	416 (100)	0

	Subject's guess, n (%)			
Intervention	**Active**	**Placebo**	**Total**	**Question could not be asked, n (%)**
Active	79 (31.2)	86 (33.9)	170 (67.1)	5 (1.9)
Placebo	36 (14.2)	45 (17.7)	83 (32.8)	2 (0.7)
Total	115 (45.4)	131 (51.7)	253 (100)	7 (2.7)

1 = strongly believe the treatment is active; 2 = somewhat believe the treatment is active; 3 = Do not know; 4 = somewhat believe the treatment is placebo; 5 = strongly believe the treatment is placebo.

Table 5

Results of data analysis[a] of Table 4

Methods	Index	*P*-value	95% Confidence interval	Conclusion
James	0.74	1	0.71–0.78	Blinded
Bang – Drug/2 × 5	0.16	<0.001	0.12–0.20	Unblinded
Bang – Drug/2 × 5[b]	0.15	<0.001	0.11–0.20	Unblinded
Bang – Placebo/2 × 5	−0.00	0.53	−0.06 to 0.05	Blinded
Bang – Placebo/2 × 5[b]	0.01	0.38	−0.05 to 0.08	Blinded

[a] STATA commands used were (the '↵' symbol indicates push of Enter key):
matrix define *A* = (38,44,21,4,170\11,16,21,8,83) ↵ blinding *A* ↵
matrix define *B* = (79,86,5\36,45,2) ↵ blinding *A*, ancillary (*B*) ↵.

[b] Showed results with incorporation of ancillary data for 'Do not know' responses.