Physician Interpretations and Textbook Definitions of Blinding Terminology in Randomized Controlled Trials

P. J. Devereaux, MD
Braden J. Manns, MD
William A. Ghali, MD, MPH
Hude Quan, MD, PhD
Christina Lacchetti, MHSc
Victor M. Montori, MD
Mohit Bhandari, MD, MSc
Gordon H. Guyatt, MD, MSc

To assess the validity of a randomized controlled trial (RCT), clinicians should note whether participants, health care providers, data collectors, and those assessing occurrence of target events of interest (judicial assessors of outcomes) are blind to participant allocation to treatment or control. Some methodologists have suggested that allocation should also be concealed from data analysts and personnel writing the paper. Authors of RCTs frequently use the terms “single,” “double,” and “triple” blind to communicate the blinding status of persons involved in the trials. We suspected that physicians and textbooks vary in their interpretations and definitions of these terms. Therefore, we surveyed physicians and systematically reviewed textbooks to test this hypothesis. A review of 200 RCTs provided an estimate of the clarity with which investigators identify who is blinded in studies they describe as single-, double-, and triple-blind.


Methods

Target Population and Data Collection

We surveyed attending physicians between February and May 1999 within departments of medicine at 3 Canadian university tertiary care centers: Dalhousie University, McMaster University, and University of Calgary Foothills Hospital. Respondents completed a survey that defined the 6 groups who are potential candidates for blinding in an RCT.

Context

When clinicians assess the validity of randomized controlled trials (RCTs), they commonly evaluate the blinding status of individuals in the RCT. The terminology authors often use to convey blinding status (single, double, and triple blinding) may be open to various interpretations.

Objective

To determine physician interpretations and textbook definitions of RCT blinding terms.

Design and Setting

Observational study undertaken at 3 Canadian university tertiary care centers between February and May 1999.

Participants

Ninety-one internal medicine physicians who responded to a survey.

Main Outcome Measures

Respondents identified which of the following groups they thought were blinded in single-, double-, and triple-blinded RCTs: participants, health care providers, data collectors, judicial assessors of outcomes, data analysts, and personnel who write the article. Definitions from 25 systematically identified textbooks published since 1990 providing definitions for single, double, or triple blinding.

Results

Physician respondents identified 10, 17, and 15 unique interpretations of single, double, and triple blinding, respectively, and textbooks provided 5, 9, and 7 different definitions of each. The frequencies of the most common physician interpretation and textbook definition were 75% (95% confidence interval [CI], 65%-83%) and 74% (95% CI, 52%-90%) for single blinding, 38% (95% CI, 28%-49%) and 43% (95% CI, 24%-63%) for double blinding, and 18% (95% CI, 10%-28%) and 14% (95% CI, 0%-58%) for triple blinding, respectively.

Conclusions

Our study suggests that both physicians and textbooks vary greatly in their interpretations and definitions of single, double, and triple blinding. Explicit statements about the blinding status of specific groups involved in RCTs should replace the current ambiguous terminology.
RCT: participants, health care providers, data collectors, data analysts, judicial assessors of the outcomes (those ultimately deciding whether a participant meets the criteria for a study’s outcome), and personnel writing the paper (those writing either of 2 drafts of a paper prior to the breaking of 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). We randomized the order in which the 6 groups were listed. Respondents offered their opinions concerning which group(s) is blinded in single-, double-, and triple-blinded trials. Physicians could include as many groups as they thought appropriate for each blinding term but they were only allowed to provide 1 definition per term.

Using the terms clinical epidemiology, randomized controlled trial, and evidence-based medicine, we systematically identified textbooks published since 1990 with definitions for single, double, and/or triple blinding at the university libraries of McMaster and Dalhousie Universities. Two of us (P.J.D., C.L.) independently evaluated the textbooks, recorded all definitions of the blinding terminology, and resolved disagreements by consensus. The chance-correct agreements, assessed through means of a $\kappa$ statistic, were near perfect ($\kappa \geq 0.9$).

Starting with June 2000, we systematically identified the 40 most recent RCTs published in the Annals of Internal Medicine, BMJ, JAMA, The Lancet, and The New England Journal of Medicine. Two of us (V.M.M., M.B.) independently evaluated the RCTs for reporting of single, double, and triple blinding and the blinding status of the 6 groups identified above. Duplicate evaluation of the first 60 RCTs established near-perfect agreement for all variables assessed ($\kappa \geq 0.8$).

Data Analysis
We determined the proportion of physician respondents who chose each interpretation, the frequency of the various textbook definitions, and the associated 95% confidence intervals (CIs). We also performed the Fisher exact test to assess heterogeneity in physician interpretation of single, double, and triple blinding across the 3 sites. To avoid cells with sparse data, we collapsed all but the 2 most frequently chosen options into a single category when conducting these tests. We determined the proportion of RCTs reporting the terms single, double, and triple blinding; within these studies we then determined the frequency of blinding among the groups identified above.

RESULTS
All 23 Dalhousie University cardiologists, 46 of the 52 McMaster University general internists, and 22 of the 24 University of Calgary Foothills Hospital general internists completed the survey (response rate, 92%). Of the 91 respondents, 83 provided an interpretation of triple blinding.

Of the 25 textbooks that fulfilled our entry criteria, 17, 25, and 5 provided a definition for single, double, and triple blinding, respectively. A number of textbooks provided 2 definitions for single (n=6), double (n=3), and triple (n=2) blinding.

The Table demonstrates that 75% (95% CI, 65%-83%) of physicians believed and 74% (95% CI, 52%-90%) of textbooks reported that study subjects were unaware of allocation in single-blind studies. The remaining 25% of physicians chose 9 alternative interpretations; 26% of textbooks provided 4 alternative definitions. Physicians chose 17 different interpretations of double blinding and textbooks provided 9 different definitions, the most frequent of which, blinding of both patients and their caregivers, was chosen and defined by 38% (95% CI, 28%-49%) and 43% (95% CI, 24%-63%) of physicians and textbooks, respectively. Physicians identified 15 different interpretations and textbooks provided 7 unique definitions for triple blinding. Only 18% (95% CI, 10%-28%) of the physicians and 14% (95% CI, 0%-58%) of the textbooks selected or defined the most frequent category of triple blinding.

The analyses for heterogeneity in physician interpretation of blinding terminology demonstrated similar variation in interpretation across sites, with $P=0.96$ for single blinding, $P=0.91$ for double blinding, and $P=0.13$ for triple blinding.

Of the 200 RCTs evaluated, 5 reported being single-blind and 83 double-blind. In the 5 studies reported as single-blind, 1 study made no mention of which group was blinded, 2 studies identified 1 group as blinded, 1 study identified 2 groups as blinded, and 1 study identified 3 groups as blinded. In the 83 studies reported as double-blind, 41 made no mention of which groups were blinded, 29 studies identified 1 group as blinded, 11 studies identified 2 groups as blinded, 1 study identified 3 groups as blinded, and 1 study identified 4 groups as blinded.

COMMENT
While methodologists advocate assessing the blinding status of individuals involved in an RCT when determining trial validity, clinicians depend on authors and editors to present blinding information in a transparent form. Authors often rely on the terms “single,” “double,” and “triple” blind to convey the blinding status of individuals involved in their study. Indeed, these terms, particularly double blind, have become almost a matter of convention.

Our study demonstrates important variation among physicians in their interpretation of who is unaware of allocation when authors use the terms “single,” “double,” and “triple” blind. This variability suggests that the current terminology hinders readers who are trying to accurately assess trial validity. The variation in textbook definitions demonstrates that there is no consensus among authorities as to what these terms mean. Our survey of 200 recent RCTs in high-impact journals demonstrated that, in 50% of trials, authors who use the term “double-blind” failed to make any statement about who was blinded, and specified only 1 group—likely omitting information about 1 or more other groups who were blinded—in another 35%.

©2001 American Medical Association. All rights reserved.
Given the variability we found, the current blinding terminology is failing to achieve the basic objectives of clear research communication. The likelihood that knowledge of patient allocation introduces bias makes this problem serious. The placebo effect can inflate the size of the treatment effect in studies when participants are not blinded. When unblinded, clinicians may differentially administer powerful treatments other than those under study, influence a patient’s compliance with study medication or willingness to continue in the study, and affect patient reporting of symptoms. Unblinded data collectors can distort trial results based on intensity of examination, the likelihood of repeating a test for an unexpected finding, the recording of outcomes, or differential encouragement during performance testing. Unblinded judicial assessors can bias the interpretation of marginal findings. Unblinded data analysts and authors can introduce bias through decisions on patient withdrawals, selection of outcomes to analyze or report, choice of time points demonstrating maximum or minimum effects, and a myriad of other decisions in the analysis or reporting process.

The impact of the blinding status of each of these groups remains uncertain. The only studies assessing the influence of blinding status on trial outcomes have focused on the reporting of double blinding. This research has demonstrated conflicting findings: 2 studies have suggested that RCTs without mention of double blinding are more likely to favor the experimental group than RCTs with mention of double blinding, whereas 1 study failed to confirm this association. Variability in who was actually blinded in reports of double-blind trials may account for these discrepant findings.

Our study is limited in that we surveyed physicians within the departments of medicine at only 3 academic institutions in Canada. However, given that McMaster University has made special efforts to educate its internists in critical appraisal of the medical literature, one might expect the highest likelihood of consistent interpretation in this institution. The results are remarkably similar across the 3 sites of our survey and are consistent with the varied definitions provided by textbooks, suggesting that our findings may be widely generalizable.

Our study has demonstrated enormous ambiguity in the conventional ways of describing blinding. Our results suggest that authors and journal editors should abandon the terms single, double, and triple blind, and substitute descriptions stating which of the relevant groups were unaware of allocation.
tion. This change in reporting would be consistent with the Consolidated Standards of Reporting Trials (CONSORT) statement, which calls for reporting of the binding status of the specific groups involved in RCTs. As long as journal reports of RCTs include the term “double blind,” clinicians will risk inaccurate inferences about the validity of the studies on which they base their clinical practice.

Author Contributions: Study concept and design: Devereaux, Manns, Ghali, Guyatt, Lacchetti, Montori, Bhandari. Acquisition of data: Devereaux, Manns, Ghali, Guyatt, Lacchetti, Montori, Bhandari. Analysis and interpretation of data: Devereaux, Manns, Ghali, Guyatt, Quan, Montori, Bhandari. Drafting of the manuscript: Devereaux. Critical revision of the manuscript for important intellectual content: Devereaux, Manns, Ghali, Guyatt, Quan, Lacchetti, Montori, Bhandari. Statistical expertise: Manns, Ghali, Guyatt, Quan. Administrative, technical, or material support: Manns, Ghali, Guyatt, Lacchetti. Funding/Support: Dr Devereaux is supported by a Heart and Stroke Foundation of Canada/Canadian Institutes of Health Research Award; Dr Manns is supported by a Kidney Foundation of Canada/Alberta Heritage Foundation for Medical Research Fellowship Award; Dr Ghali is supported by an Alberta Heritage Foundation for Medical Research Population Health Investigator Award.

REFERENCES
1. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. Users’ guides to the medical literature, I: how to use an article about therapy and prevention, A: are the results of the study valid? JAMA. 1993;270:2998-3001.