ELSEVIER

Journal of Clinical Epidemiology ■ (2013) ■

Journal of Clinical Epidemiology

BRIEF REPORT

Blinded interpretation of study results can feasibly and effectively diminish interpretation bias☆

Teppo L.N. Järvinen[a,b], Raine Sihvonen[c], Mohit Bhandari[d,e], Sheila Sprague[d,e], Antti Malmivaara[f], Mika Paavola[a], Holger J. Schünemann[d,g], Gordon H. Guyatt[d,g,*]

[a]*Department of Orthopaedics and Traumatology, Helsinki University Central Hospital, Töölö Hospital, Topeliuksenkatu 5, 00260 Helsinki, Finland*
[b]*Department of Clinical Sciences, University of Helsinki, Tukholmankatu 8 B, PL 20 Biomedicum, Helsinki, Finland*
[c]*Department of Orthopaedics and Traumatology, Hatanpää City Hospital, Hatanpäänkatu 24, 33900 Tampere, Finland*
[d]*Department of Clinical Epidemiology & Biostatistics, McMaster University, 1280 Main St. West, Hamilton, Ontario L8S4L8, Canada*
[e]*Division of Orthopaedic Surgery, Department of Surgery, McMaster University, 1280 Main St. West, Hamilton, Ontario L8S4L8, Canada*
[f]*Centre for Health and Social Economics, National Institute for Health and Welfare, Mannerheimintie 166 I, 00271 Helsinki, Finland*
[g]*Department of Medicine, McMaster University, 1280 Main St. West, Hamilton, Ontario L8S4L8, Canada*

Accepted 25 November 2013; Published online xxxx

Abstract

Objective: Controversial and misleading interpretation of data from randomized trials is common. How to avoid misleading interpretation has received little attention. Herein, we describe two applications of an approach that involves blinded interpretation of the results by study investigators.

Study Design and Settings: The approach involves developing two interpretations of the results on the basis of a blinded review of the primary outcome data (experimental treatment A compared with control treatment B). One interpretation assumes that A is the experimental intervention and another assumes that A is the control. After agreeing that there will be no further changes, the investigators record their decisions and sign the resulting document. The randomization code is then broken, the correct interpretation chosen, and the manuscript finalized. Review of the document by an external authority before finalization can provide another safeguard against interpretation bias.

Results: We found the blinded preparation of a summary of data interpretation described in this article practical, efficient, and useful.

Conclusions: Blinded data interpretation may decrease the frequency of misleading data interpretation. Widespread adoption of blinded data interpretation would be greatly facilitated were it added to the minimum set of recommendations outlining proper conduct of randomized controlled trials (eg, the Consolidated Standards of Reporting Trials statement).

Keywords: Bias; Data interpretations; Double-blind method; Drug evaluation/methods; Randomized controlled trials as topic/methods; Research design

1. Introduction

Interpretation of data, a vitally important part of conducting research [1], is never totally objective and is therefore vulnerable to prior convictions, wishful thinking, and conflict of interest—in particular, the influence of commercial funding [2]. Presentations of results can be so profoundly misleading that the clinical message is the reverse of what should be conveyed [1,3–5]. One could argue that the best way to detect and correct such bias would be through peer-review process. The frequency of biased interpretation in the medical literature suggests, however, that many reviewers have the same sorts of biases as do the original researchers. Although guides for detecting bias and guides for consumers of research faced with misleading interpretations are available [4,6], it is often impossible to detect that the data analysis was flawed.

Author contributions: G.H.G. and T.L.N.J. conceptualized the study and led the writing of the manuscript; all authors contributed to the study. All authors have approved the final version of the manuscript. G.H.G. is the guarantor.

Competing interest statement: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare no conflicts of interest related to this article.

Funding/Support and role of the sponsor: T.L.N.J.'s work is supported by the Sigrid Juselius Foundation and the Academy of Finland. The funding sources had no role in preparation, review, or approval of the manuscript. The authors' work was independent of the funders.

* Corresponding author. Department of Clinical Epidemiology & Biostatistics, McMaster University, 1200 Main Street West, Hamilton, Ontario, Canada L8N 3Z5. Tel.: +1-905-525-9140; fax: +1-905-524-3841.

E-mail address: guyatt@mcmaster.ca (G.H. Guyatt).

0895-4356/$ - see front matter
http://dx.doi.org/10.1016/j.jclinepi.2013.11.011

What is new?

- Although misleading interpretation of data (interpretative bias) was formally described more than 15 years ago in a seminal article by Gotzsche, currently there are few strategies for reducing the risk of interpretation bias.
- This article describes the application of an approach to execution of blinded interpretation of research data to safeguard against interpretation bias.
- The suggested procedure, best suited for the interpretation of data of randomized controlled trials, is simple, feasible, and efficient.

In this article, we describe a modification of a previously suggested approach to minimize the chance of misleading interpretation (interpretative bias) and describe its implementation.

1.1. Previous solutions

Gotzsche [3] first introduced the concept ''interpretive bias,'' although the specific term was introduced subsequently. He proposed that the authors of clinical trials should write two manuscripts, one assuming that treatment A is experimental and treatment B is control, and another article assuming the opposite (that treatment B is experimental and A is control). He suggested that both manuscripts be completed and approved by the authors before the randomization code is broken. Subsequently, Gotzsche [7] also went on to use this approach and, on three occasions, wrote two blinded manuscripts [8,9].

We implemented this approach in 2004 while in the process of preparing a manuscript that dealt with alternative approaches to eliciting patient utilities for health states [10]. The team statistician provided complete results labeled as group A and B; the rest of the research team was unaware of whether group A was exposed, or not exposed, to the marker states. One of us (H.J.S.) led us in producing many blinded draft versions, and finally two definitive manuscripts: One assuming that group A was exposed to marker states, the other that group B was exposed to marker states. We broke the code only after agreeing that there would be no further changes to the manuscripts, and submitted the appropriate manuscript. Although interesting and enlightening, we found the approach very onerous because it involved obtaining feedback from all coauthors on several revised, duplicate versions (groups A and B). In the many randomized trials our group had conducted subsequently, we have never repeated the process.

1.2. A more feasible alternative

Our next endeavor with blinded interpretation was in the reporting of the Study to Prospectively Evaluate Reamed Intramedually Nails in Patients With Tibial Fractures (SPRINT) trial [11,12], a multicenter randomized controlled trial (RCT) comparing the treatment of tibial shaft fractures with reamed or unreamed intramedullary nails. The writing committee of the trial was once again presented with an analysis of the results as treatment A and compared it with treatment B. Rather than writing two manuscripts, they discussed and came to agreement as to how they would interpret the results if treatment A proved to be reamed nailing and treatment B proved to be unreamed nailing. They recorded their decisions as ''Minutes of the Blinded Review of the Data'' document that was approved by all members of the Committee (see Appendix A at www.jclinepi.com). They then proceeded to break the randomization code, choosing the correct interpretation, and wrote the manuscript. The SPRINT Writing Committee members found this approach practical, feasible, and only marginally more time consuming than having a single interpretation.

The Finnish Degenerative Meniscal Lesion Study (FIDELITY) is a placebo−surgery controlled trial addressing the efficacy of arthroscopic partial meniscectomy (APM) in patients with degenerative meniscus lesion [13,14]. Prompted by the prior successful experience, one of the SPRINT authors (G.H.G.) proposed that the FIDELITY investigators consider using this approach in interpreting the data of the trial. As noted previously, the end result of the blinded interpretation process is a document we have called the ''Minutes of the Blinded Review of the Data'' (see Appendix B at www.jclinepi.com).

For the FIDELITY trial, the FIDELITY Writing Committee introduced two minor modifications to the procedure used in the SPRINT trial. First, they prepared a brief ''Background assumptions'' section and a succinct summary of the primary and secondary outcomes as well as key statistical analyses. These modifications were prompted by a belief that review of the theoretical basis of the trial would facilitate an objective and enlightened interpretation of the results. Second, to further increase the transparency and rigor of our blinded data interpretation, the FIDELITY Writing Committee introduced another safeguard to the process by asking an investigator not involved in any part of the FIDELITY trial (G.H.G.) to scrutinize our two interpretations (ie, to provide an ''external validation'').

This external validation (commentary) noted that for the primary outcome at 12 months, there was little issue: virtually no difference between groups, a conclusion that was secure whether A or B represented the group that received APM. The external reviewer suggested that the FIDELITY investigators may have preferred a definitive result of no benefit. Therefore, they were excessively inclined to dismiss findings at 2 months that suggested a difference in both Western Ontario Meniscal Evaluation Tool (WOMET) score