Journal of Clinical Epidemiology 77 (2016) 44–51

Journal of Clinical Epidemiology

Peer reviewers identified spin in manuscripts of nonrandomized studies assessing therapeutic interventions, but their impact on spin in abstract conclusions was limited

Clément Lazarus[a,b,c], Romana Haneef[a,b], Philippe Ravaud[a,b,c,d], Sally Hopewell[e], Douglas G. Altman[e], Isabelle Boutron[a,b,c,*]

[a]*INSERM, UMR 1153, Epidemiology and Biostatistics Sorbonne Paris Cité Center (CRESS), METHODS Team 1 place du Parvis Notre Dame 75004, Paris, France*

[b]*Paris Descartes University, Sorbonne Paris Cité, Faculté de Médecine, 1 place du Parvis Notre Dame 75004, Paris, France*

[c]*Centre d'Epidémiologie Clinique, AP-HP (Assistance Publique–Hôpitaux de Paris), 1 place du Parvis Notre Dame 75004, Paris, France*

[d]*Department of Epidemiology, Columbia University Mailman School of Public Health, 22 W 168th Street, New York, NY 10032, USA*

[e]*Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Botnar Research Centre, Windmill Road, Oxford OX3 7LD, UK*

Accepted 29 April 2016; Published online 7 May 2016

Abstract

Objectives: To describe the impact of peer reviewers on spin in reports of nonrandomized studies assessing a therapeutic intervention.

Study Design and Setting: This is a systematic review and retrospective before–after study. The sample consists of primary reports (n = 128) published in BioMed Central Medical Series journals between January 1, 2011, and December 31, 2013. The main outcome measures are the following: number and type of spin examples identified, deleted, or added by peer reviewers in the whole manuscript; number of reports with spin in abstract conclusions not detected by peer reviewers; the level of spin (i.e., no, low, moderate, and high level of spin) in the abstract conclusions before and after the peer review.

Results: For 70 (55%) submitted manuscripts, peer reviewers identified at least one example of spin. Of 123 unique examples of spin identified by peer reviewers, 82 (67%) were completely deleted by the authors. For 19 articles (15%), peer reviewers requested adding some spin, and for 11 (9%), the spin was added by the authors. Peer reviewers failed to identify spin in abstract conclusions of 97 (76%) reports.

Conclusion: Peer reviewers identified many examples of spin in submitted manuscripts. However, their influence on changing spin in the abstract conclusions was low.

Keywords: Spin; Nonrandomized studies; Peer review; Abstract; Quality reporting

1. Introduction

Spin or the distortion of study findings is a specific way of reporting, either intentional or unintentional, implying that the beneficial effect of the experimental treatment is greater than that shown by the results [1,2]. Several studies evaluating the presentation and interpretation of research findings have shown a high prevalence of spin in published reports [3–9]. This issue is problematic because the presence of spin can affect clinicians' interpretation of the study results [10]. Furthermore, the high prevalence of spin in published reports questions the role of peer reviewers in identifying and eliminating spin.

Nonrandomized studies are commonly used to evaluate the beneficial effect of interventions. They are particularly useful to draw conclusions about safety or efficacy of interventions in real-world settings and to assess rare or long-term adverse events or when randomization is not possible (e.g., some surgical procedures). However, these designs have important limitations, and the presentation and interpretation of the results could be distorted in published reports. Peer reviewers have an essential role in detecting and deleting spin from published reports. Nevertheless, their impact on spin in such articles has never been assessed.

Our study aimed to (1) describe the examples of spin identified, deleted, or added by peer reviewers in full-text

* Corresponding author. Tel.: +33-(0)1-42-34-78-33; fax: +33-(0)1-42-34-87-90.

E-mail address: isabelle.boutron@htd.aphp.fr (I. Boutron).

http://dx.doi.org/10.1016/j.jclinepi.2016.04.012

What is new?

Key findings

- Although peer reviewers identified many examples of spin in submitted manuscripts of nonrandomized studies assessing therapeutic interventions, several examples were not or were only partially deleted in the final publication. Furthermore, some examples of spin were added by peer reviewers. The overall impact of external peer review on the level of spin in abstract conclusions remained small and changes went in both directions.

What this adds to what was known?

- The impact of external peer review on the presence of spin in scientific reports has never been studied. This report addresses this issue in nonrandomized studies, which are specific and commonly used designs in therapeutic evaluation.

What is the implication and what should change now?

- Because the presence of spin may lead to distorted interpretation of study results, peer reviewers and editors should be more aware of this issue.

reports of nonrandomized studies assessing a therapeutic intervention and (2) evaluate the prevalence of spin in abstract conclusions that peer reviewers failed to identify.

We focused on nonrandomized studies because only few studies have assessed spin for such study design even though the risk of spin is important. Furthermore, we wanted to have a homogeneous sample of study design to have an accurate assessment of spin.

2. Methods

We used the strategy proposed by Hopewell et al. [11] to investigate the effects of peer review on the quality of reporting of published manuscripts. We selected a sample of articles published in BioMed Central Medical Series journals, which makes available all manuscript versions, peer reviewers' comments, and author's responses in the ''prepublication'' history section attached to each published article.

2.1. Selection of reports

The search strategy has been described elsewhere [12]. In brief, we searched MEDLINE via PubMed (search date January 21, 2014) for all articles published in 25 BioMed Central Medical Series journals between January 1, 2011, and December 31, 2013. Journals from the series that never published clinical studies and journals publishing only medicoeconomic assessments of therapeutic interventions were excluded from the search strategy. The list of selected journals and the complete search strategy are available in Appendices 1 and 2.

One researcher (C.L.) screened all titles, abstracts, and, if necessary, full-text articles of the citations retrieved and selected all articles (1) assessing a therapeutic intervention, defined as pharmacological or nonpharmacological treatments (e.g., drugs, surgery, therapeutic education, rehabilitation, paramedical care) proposed to patients to improve their health and (2) using a nonrandomized design (i.e., before–after study, prospective cohort study, historical cohort study, case–control study, or cross-sectional study). As a quality control procedure, a second trained reader (R.H.) assessed independently a random selection of 10% of the citations. Discrepancies were resolved by discussion.

For all citations selected, we retrieved the first submitted manuscript, the published article, and all reviewers' comments and authors' responses from the BioMed Central Web site.

2.2. Data extraction

For each selected article, two independent researchers (C.L. and R.H.), trained in the field of methodology, extracted data using a standardized data extraction form based on a classification of spin specific for nonrandomized studies evaluating therapeutic interventions developed previously and detailed in Box 1. They systematically read (1) all reviewer's comments and author's responses for all rounds of the peer-review process, referring when needed to the relevant sections of the manuscript; (2) the abstract and the full text of the published article; and (3) the abstract and full-text conclusion of the submitted manuscript.

First, they systematically recorded the examples of spin identified by the peer reviewers in the whole manuscript and determined whether these examples were partially or completely deleted or maintained in the published article. They also recorded whether the peer reviewers proposed to add some spin and whether the authors followed such recommendations.

Second, they systematically searched for spin in the abstract conclusions of the submitted and published reports to identify the examples of spin peer reviewers failed to detect or delete.

Finally, they classified the level of spin in the abstract conclusions of the submitted and published reports as follows: no spin, low level of spin (i.e., spin reported with uncertainty in the formulation of the conclusion and recommendations for further trials), moderate level of spin (i.e., spin reported with some uncertainty in the formulation of the conclusion or recommendations for further trials), and high level of spin (i.e., spin reported without any uncertainty or recommendations for further trials). This classification has been previously used in other works on spin [1].