ELSEVIER

Journal of Clinical Epidemiology 63 (2010) 289–298

Journal of Clinical Epidemiology

Systematic review data extraction: cross-sectional study showed that experience did not increase accuracy

Jennifer Horton*, Ben Vandermeer, Lisa Hartling, Lisa Tjosvold, Terry P. Klassen, Nina Buscemi

Department of Pediatrics, University of Alberta/Capital Health Evidence-Based Practice Centre, Edmonton Alberta, T6G 2J3 Canada

Accepted 20 April 2009

Abstract

Objective: This study assessed the impact of systematic review and data extraction experience on the accuracy and efficiency of data extraction in systematic reviews.

Study Design and Setting: We conducted a prospective cross-sectional study from October to December 2006. Participants were classified as having minimal, moderate, or substantial experience in systematic reviews and data extraction. Three studies on insomnia treatment were extracted. Our primary outcome was the accuracy of data extraction. Data sets of each experience level were analyzed for errors in data extraction and results of meta-analyses. Additionally, the time required for completion of data extraction was compared.

Results: Error rates were similar across the various levels of experience and ranged from 28.3% to 31.2%. Mean rates for errors of omission (11.3–16.4%) were generally lower than those for errors of inaccuracy (13.9–17.9%). There were no significant differences in error rates or accuracy of meta-analysis results between groups. Time required approached significance, with minimally experienced participants requiring the most time.

Conclusion: Overall, there were high error rates by participants at all experience levels; however, time required for extraction tended to decrease with experience. These results illustrate the need to develop strategies aimed at mastery of data extraction, rather than reliance on previous data extraction experience alone.

Keywords: Systematic review; Data interpretation; Statistical analysis; Meta-analysis; Data extraction; Data-handling error

1. Background

There is currently no recommended standard for data extraction in systematic reviews with respect to the experience level of reviewers in systematic reviews and data extraction. To our knowledge, there is no empirical evidence regarding the types and magnitude of errors accompanying data extraction conducted by reviewers with various levels of experience in systematic reviews and data extraction, the impact of these errors on the results of meta-analysis, or the efficiency in data extraction across experience levels. The lack of standards for data extraction contrasts with the efforts that have been made to uniformly and complete reporting of randomized control trials using the Quality of Reporting of Meta-analyses statement [1]. The data extraction stage of the systematic review process is understudied, yet it is one of the most complex and time-consuming stages of the systematic review process and crucial to the validity of the results and the resource expenditure for the review.

Previously published works have shown that discrepancies occur with respect to extraction of sample sizes, design, start and end dates, selection criteria, and secondary outcomes [2], and found that the rate of inaccuracies was "unacceptable" [3]. Errors were found in 20 of 34 published systematic reviews when data extraction was repeated by an experienced statistician [3]. Subsequent analysis showed that these errors did not affect the review conclusions [3]. The results of these studies have led to calls to increase the rigor of the data extraction process to reduce errors. In keeping with these results, a previous study by our group [4] compared the frequency of errors generated from single data extraction with independent verification vs. double data extraction. Rates of disagreement were high (28.0% overall) and ranged widely by variable (range, 0.0–76.7%). Double independent data extraction resulted in fewer errors but did not substantially affect the effect estimates. In contrast, Gøtzsche et al. demonstrated that the meta-analyses of previously published systematic reviews may contain a high proportion of errors. This was shown when they attempted to replicate the original author's results and found errors both in the determination of the standard deviation (10 of 27) and in the subsequent pooled estimates

* Corresponding author. Tel: 780-952-5854; fax: 780-407-6435.
E-mail address: jhorton@ualberta.ca (J. Horton).

0895-4356/10/$ – see front matter
doi: 10.1016/j.jclinepi.2009.04.007

What is new?

- Data extraction is an understudied but complex and crucial step to the production of systematic reviews.
- The impact of factors such as data extractors' systematic review experience on the accuracy and efficiency of data extraction is not known.
- Participants were categorized by experience level, and extracted data from three articles on insomnia.
- Analysis of the error rates and results of subsequent meta-analysis demonstrated high rates of error but no significant difference between experience levels.
- Further study is required to determine how to train extractors in demonstrable mastery of efficient and accurate data extraction.

(7 of 10) [5]. These results expanded on a previous analysis their research group had done in which they were unable to replicate the meta-analysis of the analgesic effect of placebos because of the errors in the original data extraction and analysis [6,7]. However, Haywood et al. [8] reported good agreement in data extraction between three reviewers with different expertise using an electronic database with careful instruction and training. There is, therefore, little consensus in this small body of literature about the frequency of errors in data extraction or their impact on systematic review results and meta-analysis. Given the paucity of evidence supporting the accuracy and efficacy of the data extraction process, there is need for more research of how to optimize the data extraction process for both accuracy and efficiency. This issue is relevant to numerous stakeholder groups that require valid data to inform policy and decision making.

The overall goal of this study was to assess the impact of systematic review and data extraction experience on the accuracy and efficiency of data extraction in systematic reviews. The specific objectives were to compare the (1) frequency of errors of inaccuracy and omission, (2) results of meta-analysis derived from data extraction, and (3) time required for data extraction.

2. Methods

2.1. Participant recruitment

The participants of this study were recruited through The Cochrane Collaboration, the Evidence-based Practice Center program of the US Agency for Healthcare Research and Quality, and relevant departments at the University of Alberta. A letter of invitation was sent to the members and students of these respective entities, which directed them to an online screening questionnaire. Individuals with prior knowledge of the systematic review process by education and/or experience were eligible for participation.

The participants were blind to the objective of the study. Participants were assigned a unique identification number to blind the investigators to their identity and the experience level to which they were allocated. Only the research assistant had access to the identification number code key during the study. Communication from the research assistant to participants regarding their participation in the study was by e-mail.

Participants signed a consent form before proceeding with data extraction and were offered a $75 CDN stipend or gift certificate for participation in the study. This study was approved by the University of Alberta Health Research Ethics Board before commencement.

2.2. Classification of systematic review and data extraction experience

The screening questionnaire collected information on the amount of time participants had been involved in systematic reviews and data extraction and the number of systematic reviews with which they had been involved. Questions not relevant to systematic review experience were also included in the questionnaire to keep participants blind to the specific purpose of the study. Questions on the geographic distribution of participants and their training in systematic reviews were also included. Questionnaire responses were used to rate participants' level of systematic review and data extraction experience according to predefined criteria that were developed for this study (Box 1). Before the start of the study, the questionnaire and rating criteria were evaluated for face validity by five experts in systematic reviews.

Some participants were not classified by this scheme (e.g., a participant having 3 years of systematic review experience, involvement in production of three systematic reviews, and extraction of 25 articles would have been ineligible for inclusion by the original classification scheme). Data for these additional participants were not included in the main analysis based on the *a priori* classification, but were included in a *post hoc* analysis of either data extraction or systematic review experience and of the continuous variable of number of years of experience.

2.3. Article source for data extraction

Our previous systematic review [9] on manifestations and management of chronic insomnia in adults formed the basis for this study; however, the current analysis was limited to the efficacy and safety of the non-benzodiazepine Zopiclone. Three studies [10–12] from the review were selected on the basis of representing all the relevant outcomes and varied in their manner of reporting.

2.4. Data extraction

Participants were directed to an online survey tool (www.surveymonkey.com) to complete data extraction.