ELSEVIER

Journal of Clinical Epidemiology ■ (2016) ■

Journal of Clinical Epidemiology

REVIEW ARTICLE

Diverse criteria and methods are used to compare treatment effect estimates: a scoping review

Romina Brignardello-Petersen[a,b,*], Alonso Carrasco-Labra[a], Alejandro R. Jadad[b,c,d], Bradley C. Johnston[b,e,f], George Tomlinson[b,g]

[a]*Evidence-Based Dentistry Unit, Faculty of Dentistry, University of Chile, Sergio Livingstone 943, Independencia, Santiago 8380492, Chile*
[b]*Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, 155 College Street, Toronto, Ontario M5T 3M6, Canada*
[c]*Institute for Global Health Equity and Innovation, Dalla Lana School of Public Health, University of Toronto, 155 College Street, Toronto, Ontario M5T 3M6, Canada*
[d]*Centre for Global eHealth Innovation, R Fraser Elliot Building, 190 Elizabeth Street, Toronto, Ontario M5G 2C4, Canada*
[e]*Child Health Evaluative Sciences, Hospital for Sick Children Research Institute, Peter Gilgan Centre for Research and Learning, 686 Bay Street, Room 11.9859, West Toronto, Ontario M5G 0A4, Canada*
[f]*Department of Anesthesia and Pain Medicine, Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada*
[g]*Department of Medicine, University Health Network and Mt Sinai Hospital, Toronto, Eaton North, 13th Floor, Room 238, 200 Elizabeth Street, Toronto, Ontario M5G 2C4, Canada*

Accepted 4 February 2016; Published online xxxx

Abstract

Objectives: To determine what criteria researchers use to assess whether the estimates of effect of an intervention on a dichotomous outcome are different when obtained using different study designs.

Study Design and Setting: Scoping review of the literature. We included studies of dichotomous outcomes in which authors compared the estimates of effects from different study designs. We performed searches in electronic databases and in the list of references of relevant studies. Two reviewers independently selected studies and abstracted data. We created a list of the criteria used to compare estimates of effects between study designs, described their main features, and classified them using a clinical perspective.

Results: We included 26 studies, from which we identified 24 criteria. Most of the studies focused on comparing estimates from observational studies and randomized controlled trials ($n = 19$). The most common criteria aimed to determine whether there was a difference or not ($n = 18$), provided guidance for such a judgment ($n = 16$), and were based on the point estimates ($n = 11$). We judged 14 criteria to be appropriate and classified them as either statistically related or clinically related.

Conclusion: We found that diverse criteria are used to compare effect estimates between study designs. Familiarity with these would aid in the interpretation of results from different studies regarding the same question.

Keywords: Comparison of effect estimates; Different results; Similarity criteria; Results interpretation; Dichotomous outcomes; Scoping review

1. Introduction

With increasing volumes of clinical research being published, it is common to find several reports of studies that aim to answer the same clinical question. Even though systematic reviews address this issue by summarizing all the available literature regarding a clinical question [1], it is not unusual to find more than one systematic review answering the same question.

Considering this, the authors of most primary studies and systematic reviews are encouraged to include a discussion of the similarity of their findings to previous studies that have addressed the same or a related clinical question [2–4]. Users of the scientific literature are familiar with statements claiming that the results of the study they are reading are in agreement or disagreement with the results of another study. Nevertheless, such statements are often made on the basis of informal comparisons, with no explicit rules to determine whether the results are, in fact, different. In the same way, and despite the great amount of guidance to users of the scientific literature [5–7], there seems to be no explicit guidance available for readers who wish to reach

Conflict of Interest: The authors have no conflicts of interest to report.

* Corresponding author. Tel.: 9055317013; fax: 9055243841.
E-mail address: rominabp@gmail.com (R. Brignardello-Petersen).

http://dx.doi.org/10.1016/j.jclinepi.2016.02.001
0895-4356/

What is new?

Key findings

- Authors of studies that aim to formally compare effect estimates use a wide variety of criteria, which are based on the use of statistical tests and the interpretation of point estimates and confidence intervals.
- Most of the criteria provide specific guidance to compare study results, yet some of them may be misleading depending on the clinical scenario.

What this adds to what was known?

- We propose a list and classification of the criteria that could be appropriately used as a stand alone criterion to compare two treatment effect estimates.

What is the implication and what should change now?

- Formal and informal comparisons of effect estimates from different study designs should be done using explicit criteria, which should be chosen according to the specific clinical question.

their own conclusions regarding the similarity between the results of different studies.

Acknowledging the fact that different study features may cause the estimates of effects to differ among studies answering the same clinical question, researchers have performed systematic surveys that formally compared the results obtained when using different study designs [8−10] and different aspects within a study design, such as the statistical methods to analyze the results of the study [11], or a particular strategy to minimize the risk of bias [12,13]. Unfortunately, the methods and criteria used for making these comparisons are diverse, and there is no unified approach.

Our aim was to determine the methods and criteria researchers use to assess whether the estimates of effect of an intervention on a dichotomous outcome derived from different study designs are different when compared to each other, to provide a broad picture of these criteria. We take a clinical perspective, meaning that we are interested in assessing whether a method or criterion gives a judgment on the importance to patients or caregivers of any differences that are found.

2. Methods

2.1. Searching for relevant studies

Given the broad nature of our question and the many ways in which we anticipated comparisons could be made, we saw this project as a scoping review [14−16]. We conducted an initial search using the phrase ‘‘comparing results between observational studies and randomized trials’’ in PubMed. After identifying relevant articles, we retrieved the list of MeSH terms used to index these articles. The equivalents of these terms, their synonyms, and other free-text terms, such as ‘‘comparison of effects’’ and ‘‘comparison of results,’’ were used to construct a search strategy for OVID Medline (see Appendix 1 at on the journal’s Web site at www.elsevier.com). This electronic search was limited to English-language articles published between 1990 and February 10, 2015.

One reviewer (R.B.-P.) screened the titles and abstracts of the references retrieved, and articles deemed relevant were included for full-text screening. A second reviewer (A.C.-L.) double-checked the decisions using a random sample of 10% of the references at the title and abstract screening stage and verified all the eligibility decisions at the full-text screening stage.

Because the indexing of these type of articles was inconsistent and used broad MeSH terms, we also retrieved articles to screen in full text from the following sources: (1) reference list of all the studies included, (2) the first 100 hits obtained when using the tool ‘‘See all related’’ references in PubMed for each of the included articles, and (3) the first 500 hits obtained after a free-text search, using the terms ‘‘comparison of treatment effects’’ and ‘‘comparison of effect estimates’’ in Google Scholar.

2.2. Study selection

We included systematic reviews, systematic surveys, or methodological reviews of reviews that aimed to compare estimates of the effect of an intervention on a dichotomous outcome that came from different study designs trying to answer the same clinical question. We also included reviews that reported comparisons of randomized controlled trials and nonrandomized controlled trials, as long as the objective of the authors was to make this comparison from a clinical perspective (and not to simply quantify the bias caused by lack of randomization). There were four broad types of studies that did compare effect estimates but which we excluded. We excluded reviews that made an informal comparison of effect estimates using no criteria and not presenting this comparison in the results but referring to the similarities or differences only in the discussion. We also excluded reviews that compared estimates of effect to determine the amount of bias caused by a feature within a study design (e.g., reporting of some trial or systematic review characteristics, inclusion of different languages in systematic reviews, etc.). The aim of those articles was to quantify the differences between estimates and evaluate the potential bias caused by those features, and they did not focus on whether there was a difference in effects from a clinical perspective. We excluded studies that compared the effect estimates of the same intervention among