ELSEVIER

Journal of Clinical Epidemiology 66 (2013) 633–638

Journal of Clinical Epidemiology

Systematic reviews supporting practice guideline recommendations lack protection against bias

Juan P. Brito[a,b,c], Apostolos Tsapas[d], Marcio L. Griebeler[a], Zhen Wang[c], Gabriela J. Prutsky[b,c], Juan Pablo Domecq[b,c], M. Hassan Murad[c,e,f], Victor M. Montori[a,b,c,*]

[a]*Division of Endocrinology, Diabetes, Metabolism, and Nutrition, Mayo Clinic, 200 First Street SW, Rochester, MN, USA*
[b]*Unidad de Conocimiento y Evidencia (CONEVID), Universidad Peruana Cayetano Heredia, Honorio Delgado Ave. 430, Lima 31, Lima, Peru*
[c]*Knowledge and Evaluation Research Unit, Mayo Clinic, 200 First Street SW, Rochester, MN, USA*
[d]*Aristotle University of Thessaloniki, 49 str. Konstandinoupoleos, 54642 Thessaloniki, Greece*
[e]*Division of Preventive, Occupational and Aerospace Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN, USA*
[f]*Division of Health Care Policy and Research, Mayo Clinic, 200 First Street SW, Rochester, MN, USA*

Accepted 13 January 2013; Published online 16 March 2013

Abstract

Objective: To evaluate the quality of systematic reviews (SRs) affecting clinical practice in endocrinology.

Study Design and Setting: We identified all SRs cited in The Endocrine Society's Clinical Practice Guidelines published between 2006 and January 2012. We evaluated the methodological and reporting quality of the SRs in duplicate using the Assessment of Multiple Systematic Reviews (AMSTAR) tool. We also noted if the guidelines recommendations that are clearly supported by SRs acknowledged their quality.

Results: During the 5-year period of study, endocrine guidelines cited 69 SRs. These SRs had a mean AMSTAR score of 6.4 (standard deviation, 2.5) of a maximum score of 11, with scores improving over time. SRs of randomized trials had higher AMSTAR scores than those of observational studies. Low-quality SRs (methodological AMSTAR score 1 or 2 of 5, $n = 24$, 35%) were cited in 24 different recommendations and were the main evidentiary support for five recommendations, of which only one acknowledged the quality of SRs.

Conclusion: Few recommendations in endocrinology are supported by SRs. The quality of SRs is suboptimal and is not acknowledged by guideline developers.

Keywords: Clinical practice guidelines; Systematic reviews; Quality; AMSTAR; Endocrinology

1. Introduction

The practice of evidence-based medicine requires consideration of the body of pertinent research evidence rather than the most salient single study. Systematic reviews (SRs) of the literature, when conducted and reported with rigor, can and should support the formulation of evidence-based recommendations [1,2].

The development of an SR requires methodological expertise, access to content expertise, and a rigorous protocol indicating carefully crafted question, search strategy, and eligibility criteria to identify the primary studies. The protocol should also indicate how reviewers will ascertain the degree of protection from error that the primary studies implemented, what data they will extract, and how these will be combined [3]. When followed, such protocols should produce highly reliable summaries of the existing body of evidence and a sense of their reliability.

For efficient practice and better application of evidence-based medicine, practitioners and guideline developers should ideally derive their evidence from SRs. However, the enthusiasm for using SRs should be tempered by the recognition that SRs, by virtue of their execution and reporting, will vary in the reliability of their results [4].

Our group has been working to improve the practice of evidence-based endocrinology [5]. Endocrinology has made important advances by focusing on deep understanding of pathophysiology but has lagged behind other fields in medicine in terms of generating highly reliable evidence about the relative safety and efficacy of alternative testing and treatment pathways. In this context, efforts to generate evidence-based practice guidelines have been challenged

Funding: This research received no grant from any funding agency in the public, commercial, or not-for-profit sectors.

* Corresponding author. Division of Endocrinology, Diabetes, Metabolism, and Nutrition, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA. Tel.: +1-507-293-0175; fax: +1-507-538-0850.

E-mail address: montori.victor@mayo.edu (V.M. Montori).

0895-4356/$ - see front matter
http://dx.doi.org/10.1016/j.jclinepi.2013.01.008

What is new?

- A small number of recommendations in endocrinology are supported by systematic reviews (SRs).
- The quality of SRs affecting endocrinology practice is suboptimal.
- Guideline developers do not acknowledge the quality of SRs.

by the available evidence from clinical care research. SRs have served as an efficient means to summarize the research evidence and apply the results into clinical policies. Although efficient, the validity of this practice remains unexamined.

The purpose of this study was to describe the rigor of SRs cited in support of clinical practice guidelines put forth by The Endocrine Society. We were interested in the extent to which guideline authors were aware and noted the quality of these SRs. We also hypothesized that more recent reviews (given the recent proliferation of guides for the conduct and reporting of SRs [6,7]), those summarizing randomized trials, and those appearing in higher impact journals would have higher quality of methods. Having a deeper appreciation for the building blocks of guidelines and its determinants may ultimately improve the quality of guidelines and the practice of evidence-based endocrinology.

2. Methods

2.1. Data sources and eligibility

To identify SRs in endocrinology that have the greatest potential impact on contemporary clinical practice, we sought all SRs cited in The Endocrine Society's Clinical Practice Guidelines published from inception of their guideline program through January 2012. These guidelines are publicly available on The Endocrine Society's Web site (http://www.endo-society.org/guidelines/Current-Clinical-Practice-Guidelines.cfm) and were accessed for the last time on February 1, 2012. For the purpose of this study and with the intent of being as inclusive as possible, we defined eligible SRs as summaries of the literature with a description of systematic search of at least two databases and reporting a meta-analysis. We excluded diagnostic SRs and SRs of the laboratory or animal studies.

2.2. Study selection

Two reviewers screened retrieved references from eligible guidelines and obtained those likely to represent SRs in full text. These were assessed for eligibility with disagreements resolved by consensus.

2.3. Data extraction and quality assessment

For each eligible study, two reviewers independently obtained review characteristics, type of study summarized [observational study or randomized controlled trial (RCT)], characteristics of the journal of publication, including impact factor (per Journal Citation Reports, http://www.isiwebofknowledge.com) and area of focus (endocrinology or not).

We used the Assessment of Multiple Systematic Reviews (AMSTAR) tool to determine the quality of the SRs [8,9]. AMSTAR has good reliability, validity, and feasibility [8,10,11]. The tool awards a maximum of 11 points: 1 point for each of five items assessing the quality of the methods (protocol-driven reviews, duplicate study selection, comprehensive search, status of publication as inclusion criteria, and appropriate methods to combine findings) and 1 point for each of six items assessing the quality of reporting (remaining AMSTAR items). The five items for methodological quality were selected based on the high likelihood of publication bias when the item was omitted. We arbitrarily decided to label SRs with two or fewer methodological points as having low methodological quality.

Additionally, we extracted all 357 numbered recommendations included in the guidelines. For each recommendation, we searched the evidence and remark sections and noted the 46 recommendations clearly supported by SRs. Among these ones, we determined whether their quality was acknowledged by guideline developers.

2.4. Statistical analysis

The main outcome of interest was the overall AMSTAR score. The chi-square and Student *t*-tests were used to test differences in categorical and continuous variables, respectively. We developed regression models to determine the association between the dependent variable, AMSTAR score, and independent variables: publication year, journal impact factor, and design of studies included in the SR. We also used radial plots to depict the relationship between AMSTAR items and areas of study within endocrinology. We estimated the chance-adjusted interrater agreement (κ statistic) for two key decisions: whether to include an SR in this study and whether the SR had low methodological quality. We used STATA release 12 (StataCorp LP, College Station, Texas) to complete all quantitative analyses.

3. Results

3.1. The methodological quality of SRs

We identified 69 SRs (Fig. 1) with near-perfect agreement (κ, 0.91). The overall average AMSTAR score was 6.4 [standard deviation (SD), 2.5]. A comprehensive literature search and characteristics of the included studies were described in 62 (90%) SRs. Conversely, a list of included and excluded studies and assessment of publication bias were described in only 12 (17%) and 17 (25%) SRs, respectively.