Evaluation of the quality of systematic reviews critiqued by clinical assistants in-training

SADJ April 2021, Vol. 76 No. 3 p147 - p152
SB Khan¹, Q Isaacs²

ABSTRACT

Objective
Conduct an overview of systematic reviews (SRs) reviewed by clinical assistants (CAs) in-training.

Method
SRs relating to clinical procedures and theoretical concepts, critiqued by CAs were included. Review authors independently screened the results of the requested SRs and evaluated these using the AMSTAR-1 checklist and AMSTAR-2 tool. Differences regarding study outcomes were resolved by consultation.

Results
Articles (N=37) submitted to the researcher included 35 reviews published in accredited journals. Of the reviews, only 18 were SRs as stated in their titles and these were of mixed designs and quality; and 17 were either non-structured and biased literature or critical reviews. SR topics reviewed in-training varied; and included temporomandibular disorders, implants and implant-supported prosthesis. AMSTAR-1 scores were favourable; scores were low for most SRs using AMSTAR-2, including those with randomized controlled trials only, with the exception of one review that had no randomized controlled trials but fulfilled the critical domain criteria.

Conclusion
Students' misconceptions regarding what constitutes good SRs which are translatable into clinical practice are emphasized, impacting their learning. CAs lack of appraisal skills related to SRs which may influence clinical practices are highlighted.

INTRODUCTION

According to the hierarchy of evidence, systematic reviews (SR) are considered the most credible study type as it encompasses a rigorous methodology and critical appraisal of all primary research studies.¹,⁵ The best SRs are those which include only randomized controlled trials (RCTs) related to a specific research question.² Moreover, results from such rigorous research may be considered implementable and thus allows clinicians to modify, improve or even change their clinical practices.¹,⁶ Policy-makers consider and rely on SRs for achieving good quality evidence related to healthcare and thus use these to change policies.¹,⁶ Yet, in many instances translation of research clinically rarely occurs, despite the enormous costs associated with conducting rigorous clinical research.⁷

Furthermore, beside the structure of an acceptable SR being followed meticulously when being conducted, the reporting of the SR using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and Meta-analysis of Observational studies in Epidemiology (MOOSE) checklists, amongst others, should be ensured.⁸ Thus, by conducting a SR according to the proposed methodology, all researchers involved would ultimately be responsible for raising and maintaining the standard of this particular research design.

It has been observed in recent years that completing a SR has become the objective of established or even novel researchers and they would readily be involved with such a study. However, in their haste to complete such a study, many errors which may or may not be obvious are included whilst undertaking it. It has been observed that researchers have a tendency to omit or under report the steps taken to conduct the research including failure to record the subsequent study outcomes.⁹,¹⁰

The problems are thus two-fold: when conducting the SR using the suggested format/s, errors may be included, and when writing the SR for publication, important features may be omitted.²,⁹,¹⁰ Thus, when accepting a SR for publication, journals should maintain that researchers complete and submit a PRISMA and/or MOOSE checklists with their manuscript.⁹ More importantly, researchers and clinicians including CAs should be trained to recognize that
errors may exist within published SRs. All these contribute to whether the final results or outcomes of the SR may be used to change clinical practice, that is, whether the research findings and evidence are implementable in clinical practice.6,10-14

Tools are available to appraise these rigorous secondary research studies, such as a measurement tool to assess systematic reviews (AMSTAR-1), an 11-question checklist used to score SRs that only include randomized controlled trials (RCTs).15 The responses to the checklist has 4 options (Yes/No/Cannot answer/Not applicable) and a score of 1 is indicated for each ‘Yes’ response (Table 1).15 The ratings are grouped according to the scores obtained into high (score of 8-11), medium (score of 4-7) and low (0-3 scores) with the responses following a rigorous explanation and interpretation of what constitutes a ‘Yes’ answer.15

However, the AMSTAR tool has now also undergone an upgrade to a more comprehensive 16-item AMSTAR-2 assessment system, which allows evaluation of SRs that include randomized and non-randomized studies.16 The AMSTAR-2 checklist includes 10 items from AMSTAR-1, with responses for all including a ‘Yes’ or ‘No’ response, and for some items a ‘Partial Yes’ response to identify partial adherence to the standard.16

The AMSTAR-2 checklist does include what is called ‘7 critical domains’, and these are highlighted as the developers believe that these may affect the validity and conclusions of a review.16 It is therefore important for each study to at least have a ‘Yes’ response for these 7 domains, so as to ensure that it is a valid and reliable review. AMSTAR-2 does not give an overall score as in AMSTAR-1, but it does give an overall rating which is an indication of the assurance of the results of the review.16 These ratings are: high=4, moderate=3, low=2 and critically low=1; and are all dependent on the ‘Yes’ responses of the critical domains. A low rating means one ‘critical flaw’ and a critically low rating implies more than one critical flaw with or without a non-critical weakness in each SR.16

AMSTAR 1 or 2 scoring of SRs are crucial in thus assessing the quality of the published SRs. Developers of both AMSTAR tools, firmly endorse that AMSTAR-2 as an assessment system allows for greater clarity, better assessment, increased ability to check for bias, and to establish greater reliability of results following a more rigorous methodology.16

As part of their training, CAs should critique, review and discuss published literature in order to broaden their knowledge base. This forms part of their assessment portfolio where they have to review a certain number of relevant publications per year related to different topics in their respective fields provided to them which are relevant to their learning.

With the escalating amount of attention given to SRs in academic settings, these are the first choice for most of the CAs when searching for relevant publications. Being involved in postgraduate teaching and training, and especially as it relates to SRs and what teachings these convey, it is important to evaluate whether these CAs understand and are able to recognize what constitutes a good SR where the results may be clinically implementable.6,11-14

Thus completing an appraisal of the SRs reviewed and presented by these CAs using validated tools, would be a good indicator of not just their understanding but also of postgraduate teachings and how it may be improved.15-18

During training and at their seminar session, CAs present the chosen articles to colleagues and here they have to contend with questions related to their choice of SR (especially the topic of interest), the guidelines used to complete article review and whether they can recommend clinical implementation of the subject matter under discussion. The work presented may also serve as a guide for clinical implementation of future patient cases presented to this forum.

The aim of this study is thus to complete an overview using validated tools of all systematic reviews critiqued by postgraduate clinical assistants as part of their training.

Objectives
1. To collate and evaluate systematic reviews (Cochrane and non-Cochrane reviews) as it relates to different concepts or topics of interest.
2. To critically appraise each systematic review using the AMSTAR-1 checklist and/ or AMSTAR-2 tool.

METHODS
Protocol development
A protocol was developed (not published) to include all aspects of an overview of SR namely: selection criteria, search strategy, selection methods using predetermined eligibility criteria, data extraction using a preformed data sheet, AMSTAR-1 and AMSTAR-2 tools to evaluate the methodological quality of each included SR, where appropriate.

Ethical clearance for the primary studies that were included in each of the SRs included had to be obtained from the respective institutions involved at that time. Written informed consent was also obtained from the participants in the primary studies according to the Declaration of Helsinki.17

Criteria for considering studies included for this overview
Inclusion criteria
Completed SRs related to interventions or concepts after critiquing by CAs.

Exclusion Criteria
Primary and secondary research studies on animals that did not meet the inclusion criteria

Types of outcome measures
Primary and secondary outcomes were pre-specified and include:
Primary outcomes
1. Steps for an acceptable systematic review included.
2. Acceptable methodological quality (high scores for each SR) according to the AMSTAR-1 and AMSTAR-2 tools.

Secondary outcomes
Procedures (clinical or laboratory), subjective or patient-or investigator-reported outcomes in patients with interventions were based on students’ choice and interests in their field from a list of topics provided during training.

These were analysed and discussed with colleagues, and the relevance of implementation of outcomes related to clinical practice were identified.

Search methods for identification of studies

Electronic searches
Systematic reviews (both Cochrane and non-Cochrane reviews) obtained via computerized searches completed by CAs (both MSc and MDS students) that are relevant to their training and learning were requested from them for the period 2013 to 2018.

For this study, development of a search strategy was therefore not necessary; neither was conducting electronic searches required by the researchers as the CAs had already sourced SRs via computerized searches. The request to all CAs was to forward all their SRs presented at their seminar sessions to the main researcher (SK). As expected, only English versions of studies were included and reported.

Selection methods
An eligibility form compiled from the inclusion criteria was used by the review authors (SK and QI) to independently screen and include potentially relevant SRs requested from the CAs. Studies that did not meet inclusion criteria were excluded and reasons for exclusion were reported.

Data extraction was completed by the researchers on all included SRs received from CAs on study designs included for each SR, participants, interventions, outcomes, and conclusions using a specially-designed pre-piloted data extraction form.

Furthermore, the researchers independently completed the AMSTAR-1 scoring and AMSTAR-2 assessments that critically assessed the methodological quality of SRs. Disagreements regarding data extraction were resolved by discussion between the researchers; and when agreement could not be reached a third colleague assisted with the final decision.

### Table 1. AMSTAR-1 - Evaluation of included systematic reviews

<table>
<thead>
<tr>
<th>AMSTAR-1 ITEM</th>
<th>Al Fadda 2018</th>
<th>Lemos et al. 2018</th>
<th>Said et al. 2016</th>
<th>Huang et al. 2015</th>
<th>Leao et al. 2018</th>
<th>Lemos et al. 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Was an ‘a priori’ design provided?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>2 Was there duplicate study selection and data extraction?</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>3 Was a comprehensive literature search performed?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>4 Was the status of publication used as inclusion criteria?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>5 Was a list of included and excluded studies provided?</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>6 Were the characteristics of included studies provided?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>7 Was scientific quality of included studies assessed and reported?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>8 Was scientific quality of studies used appropriately in formulating conclusions?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>9 Were the methods used to combine findings appropriate?</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>10 Was the likelihood of publication bias assessed?</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>11 Was the ‘conflict of interest’ included?</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Final AMSTAR-1 Score.</td>
<td>8</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

Key: Yes = 1 No = 0

### Table 2: AMSTAR-2 - Assessment: Critical domains of included systematic reviews

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 2. Registered protocol.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Item 4. Comprehensive literature search.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Item 7. Justification of excluded studies.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Item 9. Assessing risk of bias.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Item 11. Meta-Analysis using appropriate statistical methods combining results.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Item 13. Interpretation/discussion of results must include risk of bias of studies.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Key: Yes = X
Qualitative analysis
A qualitative discussion related to the primary and secondary outcomes stipulated for this overview is included after analysis of related topics/concepts and relevance of outcomes into clinical practice from the data extracted from each SR. In addition, the AMSTAR-1 checklist was completed to assess the quality of SRs that had included only RCTs and the scores were calculated using the online system where a ‘yes’ answer equaled a score of 1.15. Similarly, a qualitative discussion is included following an AMSTAR-2 rating of all SRs, those with RCTs only and those including mixed designs.

Data synthesis and management
Data synthesis and management included analysing all SRs presented at seminar sessions by CAs, and collating and reporting on the outcomes of the study and characteristics of each separately. Thus, the results include a report on the methodological quality of each included SR according to the AMSTAR-1 checklist and where applicable the AMSTAR-2 scoring system. These are summarised in the Table.

RESULTS
Summary of SRs included:
The registered CAs (N=9) submitted 37 articles following the request. Of these, 2 were primary studies, 17 were other types of reviews (including non-structured, literature or critical reviews) and 18 were SRs as stated in the titles.

SR subject matter varied, but most related to fixed prosthesis (FPDs) or resin bonded bridges (RBBs), removable prosthesis (RPDs), complete dentures (CDs), implants and implant-supported prosthesis, inlays and onlays, maxillo-ectomy defects, retraction cord, pulp testing, tooth wear and temporo-mandibular disorders. Thus, outcomes were related to implant survival rate, prosthesis survival rate, technique accuracy and reliability, patient satisfaction, masticatory ability, treatment recommendations and (oral health related) quality of life.

For SRs including RCTs only (N=6), which form the apex of the evidence pyramid, only 3 had completed a meta-analysis.18-23 Thus, most SRs (N=12) included study designs of a mixed nature, of which 9 focused on clinical studies, 1 on a classification system, 1 on a laboratory study and 1 was a case report.24-35

AMSTAR-1 results
AMSTAR-1 includes 11 questions related to the structure and methodology of a SR (Table 1). Only 6 of the included 18 SRs were appraised using AMSTAR-1 as they fit the criteria, and had a favourable score following this evaluation.18-23 Results of the AMSTAR-1 evaluation are summarized in Table 1.

Three of the six SRs had a high score of 8, two had a medium score of 7, and one a score of 5 (Table 1). The SRs by Lemos et al., (2017 and 2018) scored 8 and Leao et al, 2018 scored 7 on AMSTAR-1 respectively, and these were the only studies who met the critical domain criteria when assessed using the AMSTAR-2 tool (Table 1, Table 2).19-20,23

AMSTAR-2 results
AMSTAR-1 was initially developed to appraise SRs including RCTs only, whereas AMSTAR-2 is an extension to include both RCTs and non-RCTs and assists in identifying high quality studies.14 Table 2 highlights the 7 critical domains and the studies that met some or all of these criteria.

When comparing the SRs included for the AMSTAR-1 evaluation, 3 of these were also included in the AMSTAR-2 assessment (Table 2).19-20,23 However, for AMSTAR-2, the 3 SRs that scored well on AMSTAR-1, did not meet all 7 criteria for the critical domains of AMSTAR-2 (Table 2).19-20,23 These are considered as low and critically low SRs with the AMSTAR-2 assessment (Table 2).

This may be attributed to the rigor required when assessing studies following these set domains. However, the SR of mixed designs by Bueno et al, 2018, that did not meet requirements to be assessed using AMSTAR-1, fulfilled the criteria for all 7 critical domains of AMSTAR-2 and was rated high.16,32

DISCUSSION
SRs by its very nature are considered as the best evidence and as such have reserved the position at the apex of the hierarchical evidence pyramid and this status should be preserved at all costs.2,4-7 However, in their haste to conduct SRs, researchers have either conducted these poorly by omitting important steps or more often than not, have inadequately reported these SRs. It may be attributed to not including, for example, PRISMA or MOOSE checklists and therefore have neglected to include important information when publishing these.36-37 With the importance of critical appraisal of studies, including SRs, to establish the standard and rigor of the research, these flaws are highlighted.

This was evident with this study as well, where suitably conducted SRs related to important concepts and interventions fared poorly after appraisal.2,5,7,15-16 This has a negative impact as the evidence produced should not be considered implementable to clinical practice.11-14 Thus, the importance of conducting good quality research and more importantly, the reporting of such research cannot be over-emphasized.

CA’s depend on research to implement renewed and modernized techniques and in addition, expand on their knowledge of the outcomes of these techniques. In this study, most of the concepts related to advancement of knowledge and skills with regard to dental implants and implant-supported prosthesis which created awareness amongst the CAs of the related outcome and effects.18-20,22,28 Only 6 of the SRs could be appraised using AMSTAR 1 and only 4 qualified for appraisal using AMSTAR 2.15-23 Once again, this highlights the significance of quality research designed for application in clinical practice.
Critical appraisal or AMSTAR scoring or assessment of SRs are crucial in determining the quality of the synthesis of research conducted.\textsuperscript{15-16} Even though SRs appraised by AMSTAR 1 had favourable scores, it should be noted that only one third of the selected SRs by the CAs qualified for critical appraisal using AMSTAR 1, and similarly, just under a quarter of the selected SRs by the CAs qualified for critical appraisal using AMSTAR 2.\textsuperscript{15-16,18-23} Failure to meet the criteria for AMSTAR 1 appraisal is attributable to selection of SRs by reason of included study design.\textsuperscript{15} Explanations for exclusion of appraisal of SRs using AMSTAR 2 include failure to report on a registered protocol, justification of excluded studies, unsatisfactory techniques or no risk of bias assessment, no meta-analysis conducted and no investigation of publication bias.\textsuperscript{16} These factors relate to the critical domains and ensure quality of SRs which is directly related to implementation for guiding clinical practice. Mention should be made of the SR by Bueno et al, 2018, which qualified for appraisal using AMSTAR 2 and met all the criteria stipulated for the critical domains, even though it included non-RCT studies.\textsuperscript{16,32}

The implications of this is that a rigorous mixed design SR may be used to change clinical practice as the evidence is regarded highly, and at the same time, a good RCT could have flaws which affect implementation of the acquired evidence.\textsuperscript{11-14} According to the developers, the change from the AMSTAR-1, (an 11-point scale) to the AMSTAR-2 (a 16-item tool) as an assessment system allowed greater clarity, better assessment, an increased ability to check for bias, and to establish greater reliability of results following a more rigorous methodology.\textsuperscript{15-16}

Knowledge translation (KT) refers to the assessment, review and utilization of scientific research and evidence production that may change and improve the conditions of patients, where appropriate.\textsuperscript{39-40} The KT process consists of multiple stages from design to implementation including diffusion, dissemination such as publishing and implementation of research evidence.\textsuperscript{18-20} Evidence from SRs that have a rigorous methodology, is considered implementable into clinical practice, and emphasizes the need for critical appraisal.\textsuperscript{20,28-40} CAs are directed to different topics in their respective fields to expand their knowledge base. The CAs are guided by this list of topics to select articles and prepare seminars which can improve their knowledge, skills and clinical practices. Thus, it is as important and essential to appraise all SRs published to ensure the evidence may be implementable. Accordingly, from knowledge-into-action or from evidence-to-practice, entails the translation of best evidence obtained from well conducted SRs into clinical practice.\textsuperscript{29-31}

**CONCLUSIONS**

Included studies per SR were mostly of mixed design types and quality, some of which reduced the value of the evidence. Students’ misconceptions regarding what constitutes quality SRs and the appraisal skills required may be questioned. Choice of SRs was largely based on students’ interests in the concepts or related clinical topics covered, not on the quality of the evidence needed to change or improve clinical practice.

**Clinical implications**

It is recommended that lectures and/or workshops focusing on SRs (methodology and interpretation) should be conducted with students and staff in the postgraduate program. The importance of using rigorous research or evidence to change clinical practice and thus the care provided to patients has also been emphasized.

The role of CAs in this regard is crucial in the program and also for their clinical practice after graduating. Moreover, the articles chosen to review should be a guided process, other than providing topics or areas of interest. The impact of these learning sessions on clinical practice has far-reaching effects and this must be highlighted.

**Declaration**

The authors declare no conflict of interest.

**Funding**

No funding was applied for nor received for the research project.

The National Research Foundation funded attendance at a Conference where Research was presented.

**Data availability statement**

1. The datasets generated during and/or analysed during the current study are available from the corresponding authors on reasonable request.
2. Data generated or analysed during this study are included in this published article.

**Acknowledgments**

Grateful thanks are extended to the Clinical Assistants in-training for their invaluable contributions and assistance with this research. I wish them well in their future careers.

**Ethics registration number**

BM19/3/6

**References**


