Pesquisa Brasileira em Odontopediatria e Clínica Integrada 2024; 24:e230074
https://doi.org/10.1590/pboci.2024.040
ISSN 1519-0501 / eISSN 1983-4632

ORIGINAL ARTICLE

Methodological Quality of Systematic Reviews Addressing Orthodontic Interventions: Methodological Study

Sarah Queiroz Notaro[1], Ana Paula Hermont[1], Poliana Valdelice Cruz[1], Raiane Machado Maia[2], Walesca Melo Avila[3], Tina Poklepovic Pericic[4], Lucas Guimarães Abreu[1], Ruimin Jiao[5], Carolina Castro Martins-Pfeifer[1]

[1]Dental School, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil.
[2]Dental School, Pontifical Catholic University of Minas Gerais, Belo Horizonte, MG, Brazil.
[3]Academisch Centrum Tandheelkunde Amsterdam, Amsterdam, Neatherlands.
[4]Department of Research in Biomedicine and Health, University of Split School of Medicine, Split, Croatia.
[5]Institute of Acupuncture and Moxibustion, China Academy of Chinese Medical Science, Dongcheng, Beijing, China.

Corresponding author: Carolina Castro Martins-Pfeifer **E-mail:** carolcm@ufmg.br

Academic Editor: Alessandro Leite Cavalcanti

Received: April 20, 2023 / **Review:** July 06, 2023 / **Accepted:** October 02, 2023

How to cite: Notaro SQ, Hermont AP, Cruz PV, Maia RM, Avila WM, Pericic TP, et al. The methodological quality of systematic reviews addressing orthodontic interventions: a methodological study. Pesqui Bras Odontopediatria Clín Integr. 2024; 24:e230074. https://doi.org/10.1590/pboci.2024.040

ABSTRACT

Objective: To assess the methodological quality and characteristics of systematic reviews (SRs) of interventional studies in orthodontics and assess how the certainty of the evidence is reported using the GRADE approach. **Material and Methods:** Six electronic databases were searched, followed by a hand search of the reference lists of eligible studies (PROSPERO #CRD42020180852). The required study design was randomized and nonrandomized studies of interventions published between January 2019 and May 2020. The Assessing the Methodological Quality of Systematic Reviews (AMSTAR 2) tool was used for the quality appraisal of the included SRs. Paired reviewers independently screened the studies, extracted data, and appraised the methodological quality. **Results:** The study included 46 SRs; 19.5% had moderate to high methodological quality, and the remaining had low to critically low methodological quality. Fifty-four percent of the reviews assessed the certainty of evidence using the GRADE approach, and 34.8% followed all GRADE criteria. **Conclusion:** Most reviews had a good judgment of the AMSTAR2 items, although some critical items contributed to decreased overall quality. Half of the reviews used the GRADE approach to assess the certainty of the evidence, and this approach should be included in future systematic reviews of interventions.

Keywords: Systematic Review; Orthodontics; Malocclusion; Clinical Trial; Orthodontic Appliances.

Introduction

The volume of published systematic reviews (SR) and meta-analyses has skyrocketed during the last decades. Still, paradoxically, the number of possible unnecessary, misleading, or conflicted publications has also soared [1]. In the orthodontic field, there has also been an increase in SRs in recent years [2]. However, serious doubts about the methodological quality of some of these reviews have been raised [2]. In 2011 and 2013, Papageorgiou et al. [2,3] assessed the methodological quality of SRs in orthodontics using the A MeaSurement Tool to Assess Systematic Reviews (AMSTAR) tool. They found that the overall methodological quality was moderate [2,4]. Two years later, the quality ranged from low to moderate [3]. The most recent study on the methodological quality of SRs in orthodontics included 91 SRs of randomized controlled trials (RCTs) and used the AMSTAR tool [4,5]. Contrarily to the previous studies, this one found an increase in sound quality studies, with half of the included SRs being rated with good methodological quality.

This body of studies evaluated only SRs of RCTs and used the AMSTAR tool [2,3,5]. However, AMSTAR 2 has been developed to overcome the limitations of its predecessor [4,6]. Unlike AMSTAR, AMSTAR 2 includes the risk of bias assessment of nonrandomized clinical trials. Besides that, AMSTAR 2 simplifies the response categories, aligns the definition of research questions with the PICO (population, intervention, control group, outcome) framework, seeks justification for the review author's selection of different study designs (randomized and nonrandomized) for inclusion in systematic reviews, seeks more details on reasons for exclusion of studies from the review, determines whether the review authors had made a sufficiently detailed assessment of the risk of bias for the included studies (whether randomized or nonrandomized), determines whether risk of bias with included studies was considered adequately during statistical pooling of results (if this was performed) and determines whether risk of bias with included studies was considered adequately when interpreting and discussing the review findings [6].

That being said, it is also essential to evaluate reviews of nonrandomized studies of interventions (NRSIs), which previous studies did not include [2,3,5]. NRSI is a vital study design as it can be used as a replacement for RCT for several reasons, e.g., they can raise awareness of significant evidence for long-term outcomes [7], which is the case of some orthodontic outcomes that take time to be in place. Moreover, there has been an increase in the use of the Grading of Recommendations, Assessment, Development, and Evaluation approach (GRADE) in SRs lately [8]. One study found that 79% of evidence originating from SRs in orthodontics is of low to very low certainty [9]. However, studies still need to evaluate whether the certainty of the evidence is appropriately addressed in orthodontic research. Hence, a survey of the quality of current SRs of interventions in orthodontics is necessary either to identify flaws or to indicate the need for improvements in scientific research.

This methodological study aimed to assess the methodological quality and characteristics of SRs of interventional studies (RCTs and NRSIs) in the orthodontic field and to explore how the certainty of the evidence has been reported.

Material and Methods

Protocol and Registration

This study was reported according to the new Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) [10]. The protocol was registered *a priori* in the International Prospective Register of Systematic Reviews (PROSPERO #CRD42020180852).

Eligibility Criteria

We included SRs of RCTs and NRSIs in the orthodontic field on humans with or without meta-analysis published from January 01, 2019, to May 30, 2020. We included SRs with patients of any age, sex, and health status being submitted to an orthodontic procedure or with an orthodontic outcome. There was no limitation regarding the language of publication. This time length was chosen once it represents the current status of the literature in the previous years, as the average time between the last search and the review publication is about 8 to 15 months and approximately 16 months between the publication of the protocol and the publication of the systematic review [11-13].

We considered NRSIs when authors named them as prospective "cohorts' or "case-control studies," and the intervention groups were allocated during the usual treatment (not randomized), according to the definitions of the ROBINS-I [14]. For a proper definition according to ROBINS-I and the GRADE approach, we called these designs NRSIs [8,15].

Exclusion criteria were: SR of observational studies assessing exposure (PECO question), studies not related to orthodontic treatments, non-orthodontic outcomes, studies without a clearly defined PICO question, ultimately, not qualifying for an SR, scoping reviews, overviews, methodological reviews, narrative reviews, primary studies, case reports/series, letters/editorials of SRs as well as *in vitro* and animal studies.

Information Source and Search Strategy

Six electronic databases were searched by one author (SQN) without any language restriction: MEDLINE, Embase (both through Ovid), Cochrane Database of Systematic Reviews, Scopus, Web of Science, and Latin American and Caribbean Health Sciences Literature (Lilacs) through the Virtual Health Library (Bireme). An additional table file shows detailed search strategies for each database (supplementary material).

The reference lists for all eligible articles were hand-searched for any further related studies. Finally, the titles and abstracts were imported to EndNote software version X9.3.1 (Philadelphia, PA: Clarivate Analytics).

Study Selection

Two pairs of independent reviewers (SQN/RMM, WMA/PVC) screened titles and abstracts following the eligibility criteria using the Rayyan QCRI website. The same reviewers obtained and independently screened the full texts of selected studies. Disagreements were resolved by consensus or by consulting a senior reviewer (CCMP). An additional table shows the list of excluded studies with reasons for exclusion (supplementary material).

Data Collection and Data Items

Independent paired reviewers extracted data using a spreadsheet built on Excel software (SQN/APH, WMA/PVC). To ensure consistency across reviewers, the principal investigator conducted two pilots and training rounds before the extraction. The following data were extracted from each SR: number of authors, the continent where the authors are based, year of publication, areas in the field of orthodontics, design of included studies, being a Cochrane or a non-Cochrane review, funding source, reported conflict of interest, presence of meta-analysis, number of primary studies, presence of epidemiologist, biostatistician, and librarian on the research team, the tool used for risk of bias assessment, and assessment of the certainty of evidence using

GRADE approach [16]. Two experienced paired reviewers extracted data of the certainty of the evidence (GRADE approach) (CCMP/TPP). In addition, one author (RJ) extracted data from Chinese studies.

Disagreements were resolved by consensus and by consulting a third senior reviewer (CMMP). The corresponding authors of the included studies were contacted once through e-mail to clarify details when required.

Methodological Quality

Independent and trained paired reviewers (SQN/APH, WMA/PVC) assessed the methodological quality of the included reviews using AMSTAR 2 tool [6]. Disagreements were discussed and solved by consensus. If an agreement was not achieved, a third reviewer took the final decision (CCMP).

AMSTAR 2 has 16 items and four criteria, each judged as "Yes" (no critical weakness), "Partial Yes" (non-critical weakness), "No" (critical flaw), and "No meta-analysis conducted." AMSTAR 2 considers seven critical domains (2, 4, 7, 9, 11, 13, and 15) that were considered according to the recommendation [6]. Finally, we categorized the quality of each SR as being "high" when there was no or one non-critical weakness on one of the critical domains; "moderate" when there was more than one non-critical weakness; "low" when there was one critical flaw with or without non-critical weaknesses, or "critically low" when there was more than one critical flaw with or without non-critical weaknesses. For clarity in the table, we grouped "yes" and "partially yes" in one group.

Summary Measurements and Synthesis of Results

Data were imported into the software IBM SPSS Statistics for Windows Version 25 (Armonk, NY: IBM Corp.). We descriptively reported absolute and relative frequencies for categorical variables and calculated means and standard deviations for numeric variables (e.g., number of authors and included studies).

Results

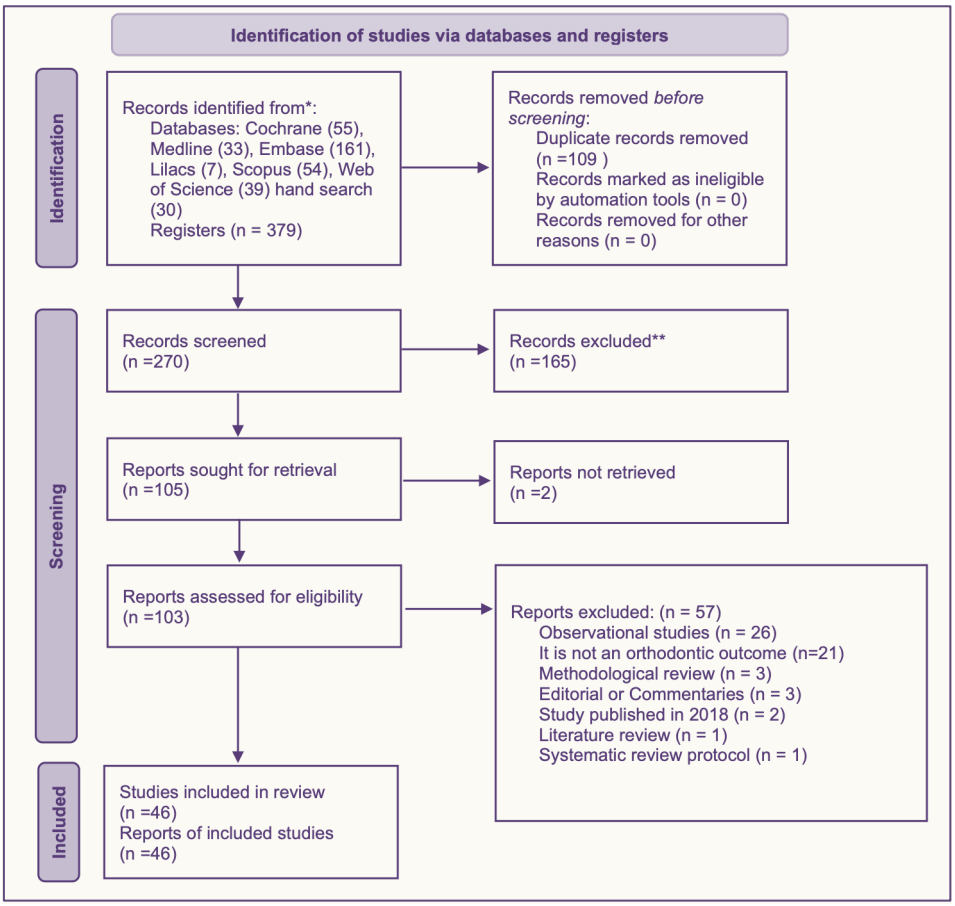
Study Selection

A total of 379 records were identified through electronic and manual searches. After duplicate removal, we screened the title and abstracts of 270 studies. Afterward, 105 full texts were included for screening. Of these, 57 studies were excluded according to the eligibility criteria. In addition, we contacted the authors to send us two studies that needed to be accessed. However, with no response from the authors, the studies were excluded from the analysis. Thus, 46 were deemed eligible for data extraction (Figure 1). The supplementary material S2 shows the list of excluded studies and reasons for exclusion, and the supplementary material S3 shows the included studies.

Study Characteristics

Table 1 (and supplementary material S4) show the review's characteristics. Most reviews had authors from multiple countries (34.8%) and were published in English (91.3%). There was no Cochrane review. Most reviews included RCTs (76.0%) and used the Cochrane risk of bias tool or the new RoB 2.0 for assessment of the risk of bias of RCTs (71.7%) and ROBINS-I for NRSIs (28.3%). In almost all the reviews (87.0%), the authors cited PRISMA for reporting the systematic review, and nearly all registrations were on the PROSPERO database (71.7%). Two reviews (4.3%) included an epidemiologist and a biostatistician in the review team. Five reviews (10.5%) included a librarian. More than half of the reviews had meta-analysis (58.7%) and assessed the

certainty of evidence through the GRADE approach (54.3%), and 34.8% followed all criteria of the GRADE approach. Assorted topics in orthodontics were investigated (supplementary material S4).



*Consider, if feasible, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers); **If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.

Figure 1. PRISMA 2020 flow diagram for new systematic reviews included searches of databases and registers only.

Table 1. Frequency distribution of reviews' characteristics (N=46).

Variables	N (%)
Academic affiliation of authors	
Orthodontic department	14 (30.4)
Orthodontic department + other	22 (47.8)
Other department	10 (21.7)
Language of publication	
English	42 (91.3)
Chinese	2 (4.3)
Portuguese	1 (2.2)
Spanish	1 (2.2)
Is "systematic review," "meta-analysis," or "network meta-analysis" stated in the title?	
Yes	45 (97.8)
No	1 (2.2)
Is this a Cochrane review?	
Yes	0 (0.0)
No	46 (100.0)
Study design of included studies	

RCT	14 (30.4)
RCT + NRSI	21 (45.6)
NRSI	11 (24.0)
The tool used for assessment of the risk of bias of RCTs	
Conventional Cochrane risk of bias tool for RCTs	30 (65.2)
Cochrane risk of bias tool for randomized studies (RoB 2.0)	3 (6.5)
Authors erroneously reported they used the GRADE approach	1 (2.2)
Not reported	1 (2.2)
Not applicable	11 (23.9)
What tool is used for assessment of the risk of bias of NRSIs?	
ROBINS-I	13 (28.3)
MINORS	6 (13.0)
Newcastle-Ottawa Scale (NOS)	2 (4.3)
Conventional Cochrane risk of bias tool for RCTs	1 (2.2)
Others[a]	3 (6.5)
Not reported	1 (2.2)
Not applicable	20 (43.5)
Funding	
None	22 (47.8)
Government/university grant	7 (15.2)
Not reported	17 (37.0)
Statement on conflict of interest	
The authors declare no conflict of interests	35 (76.1)
No mention of a conflict of interests	11 (23.9)
The authors declare a potential conflict of interests	0 (0.0)
Is there a registered protocol?	
Yes	35 (76.1)
No	5 (10.9)
Unclear[b]	1 (2.2)
Not reported	5 (10.9)
If yes, where was the protocol registered?	
PROSPERO database	33 (71.7)
Open Science Framework (OSF)	1 (2.2)
Published in the local university database	1 (2.2)
Unclear[b]	1 (2.2)
Not registered	10 (21.7)
Do the authors cite PRISMA for reporting the systematic review?	
Yes	40 (87.0)
Partial yes[c]	3 (6.5)
No	3 (6.5)
Presence of meta-analysis?	
Yes	27 (58.7)
No	19 (41.3)
Is there a meta-regression?	
No	44 (95.7)
Yes	2 (4.3)
Is there an epidemiologist involved?	
Yes	2 (4.3)
Not reported	44 (95.7)
Is there a biostatistician involved?	
Yes	2 (4.3)
Not reported	44 (95.7)
Is there a librarian involved?	
Yes	5 (10.9)
Not reported	41 (89.1)
Did the authors assess the certainty of evidence through the GRADE approach?	
Yes	25 (54.3)
No	20 (43.5)
Unclear[d]	1 (2.2)
Did the authors follow all criteria of the GRADE approach to assess the certainty of evidence?	

Yes[e]	16 (34.8)
Partial yes[f]	6 (13.0)
No[g]	3 (6.5)
Unclear	1 (2.2)
Not applicable[h]	20 (43.5)
The certainty of the evidence was evaluated per:	
Outcome	22 (47.8)
Study	3 (6.5)
Unclear	1 (2.2)
Not applicable[h]	20 (43.5)
Is there a summary of the findings table?	
Yes	24 (52.2)
No	2 (4.3)
Not applicable	20 (43.5)

[a]Down and Black checklist, Lagravere et al., 2005, National Heart, Lung and Blood Institute (NHLBI) score, Methodologic Scoring System from Methods Guide for Effectiveness and Comparative Effectiveness Reviews, published in 2008, Adjusted predetermined criteria of Bondemark Lagrave et al., 2005, Methodologic Scoring System from Methods Guide for Effectiveness and Comparative Effectiveness Reviews, published in 2008; [b]The authors reported that review was registered, but protocol was not found; [c]The authors provided the flowchart PRISMA, but did not report following PRISMA checklist on Methods; [d]The GRADE approach is reported in Methods, but not found in Results; [e]The authors assessed correctly the certainty evidence or they assessed correctly, but the certainty was not assessed for all outcomes; [f]Misjudgment of the GRADE domains; no explanation of why down rating the evidence; [g]The authors assessed the certainty of evidence per study; [h]No certainty of evidence assessed.

The reviews included two to eight authors (a mean of 4.6 per study) and two to 27 primary studies. The total number of patients ranged from 122 to 2078, and two reviews did not report the number of patients. The reviews assessed one to 29 outcomes. Twenty-seven reviews had meta-analysis (varying from one to 42 forest plots - mean of 7.9 forest plots per study) (Table 2).

Table 2. Reviews' characteristics according to continuous variables.

Variable (Number of Systematic Reviews)	Mean ± SD	Minimum–Maximum
Number of authors (N= 46)	4.6 ± 1.5	2 - 8
Number of primary studies included in the systematic review (N= 43)	9.8 ± 6.0	2- 27
Number of RCTs per systematic review (N= 35)	5.8 ± 4.6	1- 24
Number of NRSIs per systematic review (N= 24)	6.1 ± 4.9	1 – 19
Total number of patients (N=45)	485.1± 397.4	122 – 2078
Number of outcomes (narrative + meta-analysis) (N= 46)	7.9 ± 7.5	1 – 29
Studies included in the meta-analysis (N= 27)	7.3 ± 4.9	2 – 25
Number of outcomes in meta-analysis (N= 27)	5.6 ± 4.9	1 – 24
Number of forest plots (N= 25)	7.9 ± 8.8	1 – 42
Number of funnel plots (N= 5)	2.4 ± 2.1	0 – 5
What is the number of outcomes in the summary of findings table? (N= 23)	4.9 ± 3.3	0 – 11

SD: Standard Deviation.

Methodological Quality of Included Reviews

Table 3 shows the methodological quality of the systematic reviews. About 19.5% of the reviews had high to moderate methodological quality.

Table 3. Overall methodological quality judged by AMSTAR 2.

AMSTAR 2 (n=46)	N (%)
High	3 (6.5)
Moderate	6 (13.0)
Low	13 (28.3)
Critically low	24 (52.2)

Table 4 describes the methodological quality according to AMSTAR 2 criteria for all the included SRs. The most common strength points of the reviews were: the appropriate inclusion of PICO components in the

research question and inclusion criteria (item 1, 100.0%), comprehensive literature search (item 4, 97.8%), duplicate and independent screening of studies (item 5, 99.2%) and data extraction (item 6, 80.4%), an adequate description of studies' characteristics (item 8, 100.0%), use of a satisfactory tool to access the risk of bias (item 9, 87.0%), appropriate assessment of the risk of bias among the studies included in the meta-analysis (item 12, 26.1%) and while discussing the results (item 13, 76.1%), and reporting conflict or interests and funding of the review (item 16, 80.4%).

On the other side, there were some weaknesses, such as the absence of an explicit statement of the rationale for any significant deviations from the protocol (item 3, 87.0%), the absence of a list of excluded studies, and the rationale for exclusion (item 7, 63.0%), and absence of reporting the source of funding for the studies included in the SR (item 10, 100%).

Table 4. Frequency distribution of AMSTAR 2 domains among reviews.

AMSTAR	N (%)
1. Did the research questions and inclusion criteria for the review include the components of PICO?	
Yes/ Partial Yes	46 (100.0)
No	0 (0.0)
2. Did the review report contain an explicit statement that the review methods were established prior to the review, and did the report justify any significant deviations from the protocol?	
Yes/ Partial Yes	26 (56.5)
No	20 (43.5)
3. Did the review authors explain their selection of the study designs for inclusion in the review?	
Yes/ Partial Yes	6 (13.0)
No	40 (87.0)
4. Did the review authors use a comprehensive literature search strategy?	
Yes/ Partial Yes	45 (97.8)
No	1 (2.2)
5. Did the review authors perform study selection in duplicate?	
Yes/ Partial Yes	41 (99.2)
No	5 (10.9)
6. Did the review perform data extraction in duplicate?	
Yes/ Partial Yes	37 (80.4)
No	9 (19.6)
7. Did the review authors provide a list of excluded studies and justify the exclusions?	
Yes/ Partial Yes	17 (37.0)
No	29 (63.0)
8. Did the review authors describe the included studies in adequate detail?	
Yes/ Partial Yes	46 (100.0)
No	0 (0.0)
9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies included in the review?	
Yes/ Partial Yes	40 (87.0)
No	6 (13.0)
10. Did the review authors report on the funding sources for the studies included in the review?	
Yes/ Partial Yes	0 (0.0)
No	46 (100.0)
11. If meta-analysis was justified, did the review authors use appropriate methods for statistical combination of results? (Only complete this item if a meta-analysis of other data synthesis techniques were reported)	
Yes/ Partial Yes	27 (58.7)
No	0 (0.0)
No meta-analysis conducted	19 (41.3)
12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	
Yes/ Partial Yes	12 (26.1)
No	15 (32.6)
No meta-analysis conducted	19 (41.3)

13. Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review?	
Yes/ Partial Yes	35 (76.1)
No	11 (23.9)
14. Did the review authors provide a satisfactory explanation for and discussion of any heterogeneity observed in the results of the review?	
Yes/ Partial Yes	36 (78.3)
No	10 (21.7)
15. If they performed quantitative synthesis, did the review authors carry out an adequate investigation of publication bias (slight study bias) and discuss its likely impact on the results of the review?	
Yes/ Partial Yes	26 (56.5)
No	1 (2.2)
No meta-analysis conducted	19 (41.3)
16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	
Yes/ Partial Yes	37 (80.4)
No	9 (19.6)

RoB: Risk of Bias.

Discussion

About one-fifth of the reviews had high to moderate methodological quality, although a significant number is still low and critically low. However, most reviews attended the yes/partial yes criteria for AMSTAR2 items.

Hooper et al. [5] found a higher proportion of studies with high-quality papers (50%) than our study. The discrepancy can be due to the different tools used. While Hooper et al. [5] used AMSTAR, we used AMSTAR 2. The divergences across studies may indicate that constant scrutiny of the quality of SRs is highly recommended.

Due to the heavy and outspoken criticism of AMSTAR, Shea et al. [6] revised and updated this instrument [6]. An assessment of the risk of bias in nonrandomized intervention studies was included in AMSTAR 2. This evaluation is vital given the diversity of study designs that might be included in such reviews and the factors that may have biased their results [6]. Moreover, compared with the original tool, AMSTAR 2 has been more closely aligned with the PICO question and other premises of an SR, such as a detailed justification of the studies' design and more information about the excluded studies [6].

Previous studies have demonstrated that most reviews were conducted in European institutions [2,3,5]. In our research, the SRs assessed had authorship attributed to researchers from multiple countries, and the most prolific continent was Asia (21.7%). It shows that the number of SRs developed in countries such as China, South Korea, and Japan has increased. In addition, the number of collaborations between/among teams of authors from different countries has also risen sharply in the last few years. Furthermore, more than 50% of the reviews were published in orthodontic journals per preceding studies [2,3].

AMSTAR 2 does not indicate which instruments related to the risk of bias should have been used to assess the risk of bias in included studies. Instead, AMSTAR 2 states that this decision should be made by the authors of the study, who are evaluating the SRs [6]. The most commonly used instruments for assessing the risk of bias in RCTs are the Cochrane risk of bias tool [17], RoB 2.0 [18], and Joanna Briggs Critical Appraisal tool for RCTs [19]. For NRSIs, ROBINS-I and MINORS have been mainly employed [14,20]. Therefore, most SRs used appropriate tools according to the study design.

In general, the majority of the SRs had more strengths than weaknesses, as the majority of the AMSTAR 2 items had a high proportion of items judged as "yes" or "partially yes," which means that the studies

had fulfilled the quality criteria. However, the final AMSTAR 2 judgment was mainly low and critically low. This must be explained by some "critical items" that did not score well. E.g., almost half the SRs in our study did not report a protocol. This item (2) can be considered a critical weakness and can explain many SRs with critical low and low quality. According to other reviews, prior protocol registration seems to increase the AMSTAR score [2,3,5]. A protocol allows the researchers to clearly determine the objectives and methods of the review, mitigating the risk of bias and the chance of duplication of registries with the same clinical question [21]. Although not mandatory, a priori protocol registration contributes to improving the review quality [22]. Other critical items (4, 7, 9, 11, 13, and 15) also contributed to decreasing the overall score. That being said, we believe the general score of AMSTAR 2 must be carefully interpreted and might underestimate the overall methodological quality of the SRs.

The participation of a statistician or an epidemiologist in an SR can also improve the quality of the manuscript [2]. In our study, only a few SRs had an epidemiologist and a biostatistician among the authors, contrary to our expectations. In 2013, Papageorgiou et al. [3] had already observed a decrease in the participation of statisticians or epidemiologists in SRs, and this finding may also be suggestive of an improvement in the statistical skills of researchers with an orthodontic background [23] who will be able to aggregate data in meta-analyses, for instance. However, we also believe that the number of epidemiologists is higher than what was reported by the SRs. Many authors might have degrees in epidemiology, but this was not shown by the "authors' affiliations."

A vital point that should be mentioned is that half of the SRs used the GRADE approach to assess the certainty of the evidence, and 21.7% did not follow all the criteria developed by the GRADE approach. When reporting the certainty of the evidence, it is important to document the reasons for rating down the certainty of the evidence straightforwardly and transparently [24], and some SRs need more transparent reporting. The rationales used in these reviews could have been clearer or were not fully explained in a summary of findings (SoF) table. Moreover, one review described that the GRADE approach was used in the methods section, but the results were unavailable in the results section [25]. Finally, three reviews assessed the certainty of evidence per study, not per outcome [26-28], and one review considered the GRADE approach as a risk of bias tool [28]. This contradicts what the GRADE Working Group has stated, acknowledging that the tool is used to assess the overall body of evidence and not to appraise the risk of bias [16]. Considering that the GRADE approach is a complex tool [29], before assessing the certainty of the evidence, the authors can be trained by the online and free supporting material provided by the GRADE Working Group (https://training.cochrane.org/grade-approach).

We are optimistic about increasing the methodological quality of the SRs in the future due to the increase in collaborations among researchers from different countries and the strict requirements of the peer review process before publication. Furthermore, we hope that this result encourages and helps researchers to produce studies with higher methodological quality in the future. We also reinforce the importance of future methodological studies as a way to monitor the quality of scientific research.

Although our search was performed in several databases, including the Cochrane Database of Systematic Reviews, we found no Cochrane reviews. Therefore, the results do not apply to Cochrane reviews. As for strengths, this study is the first to assess the methodological quality and characteristics of SRs of orthodontic interventions using AMSTAR 2, the most recent critical appraisal tool for SRs in which RCTs and NRSIs of healthcare interventions are included. Furthermore, our study was the first to analyze the certainty of evidence reported by the reviews.

Conclusion

One of each of the five SRs has high standards with moderate to high quality, and the majority of the items of AMSTAR 2 were fulfilled. Half of the reviews assessed the certainty of evidence using the GRADE approach. The authors should be encouraged to improve their skills in determining the certainty of the evidence through the free GRADE online training.

Authors' Contributions

SQN	https://orcid.org/0000-0002-4810-7216	Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review and Editing and Visualization.
APH	https://orcid.org/0000-0002-0409-0926	Methodology, Investigation, and Writing - Review and Editing.
PVC	https://orcid.org/0000-0002-4218-1503	Methodology and Investigation.
RMM	https://orcid.org/0000-0002-6678-7585	Methodology and Investigation.
WMA	https://orcid.org/0000-0001-8516-6365	Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review and Editing, Visualization, Supervision, Project Administration and Funding Acquisition.
TPP	https://orcid.org/0000-0002-9686-5062	Methodology, Investigation, and Writing - Review and Editing.
LGA	https://orcid.org/0000-0003-2258-8071	Writing - Review and Editing.
RJ	https://orcid.org/0000-0002-3866-7666	Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review and Editing, Visualization, Supervision, Project Administration and Funding Acquisition.
CCMP	https://orcid.org/0000-0001-9072-3226	Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing - Review and Editing, Visualization, Supervision, Project Administration and Funding Acquisition.

All authors declare that they contributed to a critical review of intellectual content and approval of the final version to be published.

Financial Support

This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Ministry of Education, Brazil); Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Ministry of Education, Brazil); Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG APQ-00323-17) and Pró-Reitoria de Pesquisa da Universidade Federal de Minas Gerais (PRPq/UFMG – PIBIT/CNPq). The funding agencies had no role in the conception of the research and interpretation of data.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability

The data used to support the findings of this study can be made available upon request to the corresponding author.

Acknowledgments

Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG APQ-00323-17).

References

[1] Ioannidis J. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q 2016; 94(3):485-514. https://doi.org/10.1111/1468-0009.12210

[2] Papageorgiou SN, Papadopoulos MA, AE Athanasiou. Evaluation of methodology and quality characteristics of systematic reviews in orthodontics. Orthod Craniofac Res 2011; 14(3):116-137. https://doi.org/10.1111/j.1601-6343.2011.01522.x

[3] Papageorgiou SN, Papadopoulos MA, Athanasiou AE. Reporting characteristics of meta-analyses in orthodontics: Methodological assessment and statistical recommendations. Eur J Orthod 2013; 36(1):74-85. https://doi.org/10.1093/ejo/cjt008

[4] Shea BJ, Grimshaw JM, Wells GA, Boers M, Aderssson N, Hamel C, et al. Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007; 7:10. https://doi.org/10.1186/1471-2288-7-10

[5] Hooper EJ, Pandis N, Cobourne MT, Seehra J. Methodological quality and risk of bias in orthodontic systematic reviews using AMSTAR and ROBIS. Eur J Orthod 2021; 43(5):544-550. https://doi.org/10.1093/ejo/cjaa074

[6] Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017; 358:j4008. https://doi.org/10.1136/bmj.j4008

[7] Cuello-Garcia CA, Morgan RL, Brozek J, Santesso N, Verbeek J, Thayer K, et al. A scoping review and survey provides the rationale, perceptions, and preferences for the integration of randomized and nonrandomized studies in evidence syntheses and GRADE assessments. J Clin Epidemiol 2018; 98:33-40. https://doi.org/10.1016/j.jclinepi.2018.01.010

[8] Zhang Y, Akl EA, Schunemann HJ. Using systematic reviews in guideline development: The GRADE approach. Res Synth Methods 2018. https://doi.org/10.1002/jrsm.1313

[9] Koletsi D, Fleming PS, Eliades T, Pandis N. The evidence from systematic reviews and meta-analyses published in orthodontic literature. Where do we stand? Eur J Orthod 2015; 37(6):603-609. https://doi.org/10.1093/ejo/cju087

[10] Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CK, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021; 372:n160. https://doi.org/10.1136/bmj.n160

[11] Beller EM, Chen JK, Wang UL, Glasziou PP. Are systematic reviews up-to-date at the time of publication? Syst Rev 2013; 2:36. https://doi.org/10.1186/2046-4053-2-36

[12] Sampson M, Shojania KG, Garritty C, Horsley T, Ocampo M, Moher D. Systematic reviews can be produced and published faster. J Clin Epidemiol 2008; 61(6):531-536. https://doi.org/10.1016/j.jclinepi.2008.02.004

[13] Andersen MZ, Fonnes S, Andresen K, Rosenberg J. Most published meta-analyses were made available within two years of protocol registration. Eur J Int Med 2021; 44(101342). https://doi.org/10.1016/j.eujim.2021.101342

[14] Sterne JA, Hernán MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016; 355:i4919. https://doi.org/10.1136/bmj.i4919

[15] Schünemann HJ, Cuello C, Akl EA, Mustafa RA, Meerpohl JJ, Thayer K, et al. GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. J Clin Epidemiol 2019; 111:105-114. https://doi.org/10.1016/j.jclinepi.2018.01.012

[16] Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE Working Group. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008; 336(7650):924-926. https://doi.org/10.1136/bmj.39489.470347.AD

[17] Higgins JP, Altman DG, Gotzche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011; 343:d5928. https://doi.org/10.1136/bmj.d5928

[18] Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: A revised tool for assessing risk of bias in randomised trials. BMJ 2019; 366:l4898. https://doi.org/10.1136/bmj.l4898

[19] Tufanaru C, Munn Z, Aromataris E, Campbell J, Hopp L. Chapter 3: Systematic reviews of effectiveness. Aromataris E, Munn Z. JBI Manual for Evidence Synthesis. JBI; 2020.

[20] Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for nonrandomized studies (minors): Development and validation of a new instrument. ANZ J Surg 2003; 73(9):12-16. https://doi.org/10.1046/j.1445-2197.2003.02748.x

[21] Fleming PS, Seehra J, Polychronopoulou A, Fedorowicz Z, Pandis. Cochrane and non-Cochrane systematic reviews in leading orthodontic journals: A quality paradigm? Eur J Orthod 2012; 35(2):244-248. https://doi.org/10.1093/ejo/cjs016

[22] Sideri S, Papageorgiou SN, Eliades T. Registration in the international prospective register of systematic reviews (PROSPERO) of systematic review protocols was associated with increased review quality. J Clin Epidemiol 2018; 100:103-110. https://doi.org/10.1016/j.jclinepi.2018.01.003

[23] Torres D, Normando D. Biostatistics: Essential concepts for the clinician. Dental Press J Orthod 2021; 26(1):e21spe1. https://doi.org/10.1590/2177-6709.26.1.E21SPE1

[24] Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011; 64(4):383-394. https://doi.org/10.1016/j.jclinepi.2010.04.026

[25] Bin Bahar BSK, Alkhalidy SR, Kaklamanos EG, Athanasiou AE. Do orthodontic patients develop more gingival recession in anterior teeth compared to untreated individuals? A systematic review of controlled studies. Int Orthod 2020; 18(1):1-9. https://doi.org/10.1016/j.ortho.2019.08.025

[26] Khalaf K, Mando M. Removable appliances to correct anterior crossbites in the mixed dentition: A systematic review. Acta Odontol Scand 2020; 78(2):118-125. https://doi.org/10.1080/00016357.2019.1657178

[27] Khalaf K, Mando M. Effect of drugs on orthodontic tooth movement in human beings: A systematic review of randomized clinical trials. Open Dent 2019; 13(1):22-32. https://doi.org/10.2174/1874210601913010022

[28] Bondemark L, Kallunki J, Paulsson L. An updated systematic review regarding early Class II malocclusion correction. JWFO 2019; 8(3):89-94. https://doi.org/10.1016/j.ejwf.2019.06.002

[29] Langendam MW, Akl EA, Dahm P, Glasziou P, Guyatt G, Schünemann HJ. Assessing and presenting summaries of evidence in Cochrane Reviews. Syst Rev 2013; 2:81. https://doi.org/10.1186/2046-4053-2-81