

Journal section: *Periodontology*Publication Types: *Review*

doi:10.4317/jced.60197

<https://doi.org/10.4317/jced.60197>

## Critical appraisal of systematic reviews of intervention studies in periodontology using AMSTAR 2 and ROBIS tools

Alexandre-Godinho Pereira <sup>1</sup>, Carolina-Castro Martins <sup>2</sup>, Julya-Ribeiro Campos <sup>1</sup>, Sandro-Felipe-Santos Faria <sup>1</sup>, Sarah-Queiroz Notaro <sup>2</sup>, Tina Poklepović-Peričić <sup>3</sup>, Lidiane-Cristina-Machado Costa <sup>4</sup>, Fernando-Oliveira Costa <sup>1</sup>, Luís-Otávio-Miranda Cota <sup>1</sup>

<sup>1</sup> Department of Dental Clinics, Oral Pathology, and Oral Surgery, School of Dentistry, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

<sup>2</sup> Department of Pediatric Dentistry, School of Dentistry, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

<sup>3</sup> Department of Research in Biomedicine and Health, University of Split School of Medicine, Split, Croatia

<sup>4</sup> School of Dentistry, Newton Paiva University Center, Belo Horizonte, Minas Gerais, Brazil

### Correspondence:

Periodontology Division, Department of Dental Clinics  
Oral Pathology and Oral Surgery, School of Dentistry  
Federal University of Minas Gerais  
Antônio Carlos Avenue 6627, Pampulha  
PO Box 359, Zip Code 31270-901  
Belo Horizonte, Minas Gerais – Brazil  
[luisocota@ufmg.br](mailto:luisocota@ufmg.br)

Pereira AG, Martins CC, Campos JR, Faria SFS, Notaro SQ, Poklepović-Peričić T, Costa LCM, Costa FO, Cota LOM. Critical appraisal of systematic reviews of intervention studies in periodontology using AMSTAR 2 and ROBIS tools. J Clin Exp Dent. 2023;15(8):e678-94.

Received: 26/11/2022

Accepted: 22/05/2023

Article Number: 60197 <http://www.medicinaoral.com/odo/indice.htm>  
© Medicina Oral S. L. C.I.F. B 96689336 - eISSN: 1989-5488  
eMail: [jced@jced.es](mailto:jced@jced.es)

### Indexed in:

Pubmed  
Pubmed Central® (PMC)  
Scopus  
DOI® System

### Abstract

**Background:** Systematic reviews of intervention studies are used to support treatment recommendations. The aim of this study was to assess the methodological quality and risk of bias of systematic reviews of intervention studies in the field of periodontology using AMSTAR 2 and ROBIS.

**Material and Methods:** Systematic reviews of randomized and non-randomized clinical trials, published between 2019 and 2020, were searched at MedLine, Embase, Web of Science, Scopus, Cochrane Library, LILACS with no language restrictions between October 2019 to October 2020. Additionally, grey literature and hand search was performed. Paired independent reviewers screened studies, extracted data and assessed the methodological quality and risk of bias through the AMSTAR 2 and ROBIS tools.

**Results:** One hundred twenty-seven reviews were included. According to AMSTAR 2, the methodological quality was mainly critically low (64.6%) and low (24.4%), followed by moderate (0.8%) and high (10.2%). According to ROBIS, 90.6% were at high risk of bias, followed by 7.1% low, and 2.4% unclear risk of bias. The risk of bias decreased with the increased in the impact factor of the journal.

**Conclusions:** Current systematic reviews of intervention studies in periodontics were classified as low or critically low methodological quality and high risk of bias. Both tools led to similar conclusions. Better adherence to established reporting guidelines and stricter research practices when conducting systematic reviews are needed.

**Key words:** *Bias, evidence-based dentistry, methods, periodontics, systematic review.*

## Introduction

Systematic reviews (RSs) of intervention studies are considered of high level of scientific evidence, being used to raise evidence that can support treatment recommendations and public health strategies (1). As other study designs, SRs are subject to biases that can compromise their validity and quality of evidence (2). Some tools were developed to assess the methodological quality and risk of bias of SRs, such as AMSTAR 2 (A Measurement Tool to Assess Systematic Reviews 2) (3) an updated version of ASMTAR (Risk of Bias in Systematic Reviews) (4), the Cochrane Collaboration tool for risk of bias of SRs (5).

Some overviews in the periodontal field have assessed the methodological quality of SRs through AMSTAR, showing inconstant quality (6-9). One overview assessed the methodological quality of SRs using the AMSTAR 2 and the risk of bias through ROBIS, and demonstrated very low overall quality (10). Among 23 SRs, only 3 SRs on peri-implantitis therapy had high quality according to AMSTAR 2, and only one were judged as low risk of bias according to ROBIS (10). This low overall quality raised questions about the general quality of the available evidence from RSs in periodontology.

Hence, this overview aimed to: 1) describe the characteristics of SRs in periodontology; 2) assess if the certainty of the evidence is reported in these reviews; 3) assess the methodological quality using the AMSTAR 2; 4) assess the risk of bias using the ROBIS.

## Material and Methods

This methodological survey was designed and performed following the recommendations of the Cochrane Handbook for Systematic Reviews of Interventions (11) and was reported in accordance with the PRISMA checklist (12).

### -Research question

What is the methodological quality and risk of bias of the SRs of intervention studies in periodontology published in 2019-2020?

### -Eligibility criteria

Inclusion criteria were SRs of intervention studies – randomized (RCTs) and non-randomized clinical trials (nRCTs) – with or without meta-analysis, in the field of periodontology, indexed between October 1st, 2019 to October 1st, 2020. SRs that authors classified the studies as having prospective design were included as nRCTs. According to the Risk of Bias in Non-randomized Studies of Interventions (ROBINS-I), nRCTs are cohort studies in which intervention groups were allocated during the usual course of treatment instead of randomization (13). To be consistent, all non-randomized studies, nominated by authors as clinical trials, controlled clinical trials, prospective controlled trials, non-randomized prospective studies, prospective clinical studies,

prospective controlled clinical studies and retrospective cohort studies, were classified as nRCTs.

Exclusion criteria were: (a) SRs not related to the field of periodontology, (b) narrative or scope reviews, clinical guidelines, editorials or expert opinion papers, SRs of case-control and cross-sectional studies with PECO question, case reports and case series, pilot, in vitro and/or animal studies.

### -Search in databases

An expert in SRs (CCM) designed and verified the strategies searches, and one reviewer (AGP) searched the following databases: MedLine (Pubmed), Embase (Elsevier), Web of Science, Scopus, Cochrane Library and LILACS for articles indexed between October 1st, 2019 to October 1st, 2020, with no language restrictions. This time length is enough to represent the current status of the quality of evidence in periodontology in the previous years, as the average time between the last search for a SR and its publication varies between 8 (14) to 15 months (15), and the mean time between the protocol's publication and the SR's publication is about 16 months (16).

Grey literature was searched in OpenGrey, GreyLit and Google Scholar. A hand search was performed in the references list of selected articles, and in the main journals of periodontology found in the Journal Citation Reports (JCR) in the category “Dentistry, Oral Surgery and Medicine”: Journal of Clinical Periodontology, Journal of Periodontology, Journal of Periodontal Research, International Journal of Periodontics & Restorative Dentistry, Journal of Periodontal and Implant Science, Periodontology 2000.

Additional information of search strategies, including search terms, is detailed in the supplementary material (Supplement 1) ([http://www.medicinaoral.com/medoralfree01/aop/jced\\_60197\\_s01.pdf](http://www.medicinaoral.com/medoralfree01/aop/jced_60197_s01.pdf)).

### -Studies selection

Two pairs of independent reviewers screened studies based on titles and abstracts and then full text (AGP and SFF; JRC and LCMC). The reviewers were trained with a set of 10% of studies in each phase. In cases of less than 80% of agreement, additional rounds of training were carried out until reaching the necessary standard for each step. After reviewers achieved at least 80% of agreement, they underwent the screening process with the remaining of studies. The Rayyan platform (17) was used for studies screening. In cases of disagreement, an expert reviewer was consulted (CCM).

-Data extraction and assessment of methodological quality  
Data extraction, assessment of methodological quality and risk of bias were performed through the AMSTAR 2 (3) and ROBIS (5) tools by four pairs of independent reviewers (AGP and SFF; AGP and LCMC; JRC and SQN; CCM and TPP), using Excel spreadsheet editor. Reviewers were trained by two reviewers (AGP and CCM), the second one with broad experience in syste-

matic reviews methodology. Again, the reviewers underwent as many rounds of training as necessary, until reaching 80% of agreement. All disagreements were solved by discussion and consensus. If consensus was not achieved, the principal investigator made the final decision.

General data were extracted from the articles, and the list of the extracted data is available in the supplementary material (Supplement 2) ([http://www.medicinaoral.com/medoralfree01/aop/jced\\_60197\\_s02.pdf](http://www.medicinaoral.com/medoralfree01/aop/jced_60197_s02.pdf)). We uploaded the SRs protocols from the registration platform to compare with the published review, and extracted the JCR impact factor and the h-5 index of the journals from the JCR and Google Scholar Metrics, respectively.

Disagreements during this step were resolved between the pair of reviewers. If disagreement persisted, the principal investigator was responsible for reaching a final consensus. Two reviews in Mandarin were translated using a translation tool.

-Statistical analysis

Data was entered in IBM SPSS Statistics for Windows version 25 (Armonk, NY: IBM Corp.) for descriptive analyses. We calculated the relative and absolute frequencies for categorical variables, and mean, standard deviation and minimum/maximum values were provided for continuous variables. Analyses were performed considering all SRs and stratified by: SRs with RCTs and nRCTs, SRs with RCTs only and the impact factor of the journals (<3, ≥3 <6, ≥6).

**Results**

-Literature search

One hundred twenty-seven SRs were included. Figure 1 shows the screening process. A list of excluded studies with reasons for exclusion is available on (Supplement 3) ([http://www.medicinaoral.com/medoralfree01/aop/jced\\_60197\\_s03.pdf](http://www.medicinaoral.com/medoralfree01/aop/jced_60197_s03.pdf)) and a list of references of the included RSs is available in (Supplement 4) ([http://www.medicinaoral.com/medoralfree01/aop/jced\\_60197\\_s04.pdf](http://www.medicinaoral.com/medoralfree01/aop/jced_60197_s04.pdf)).

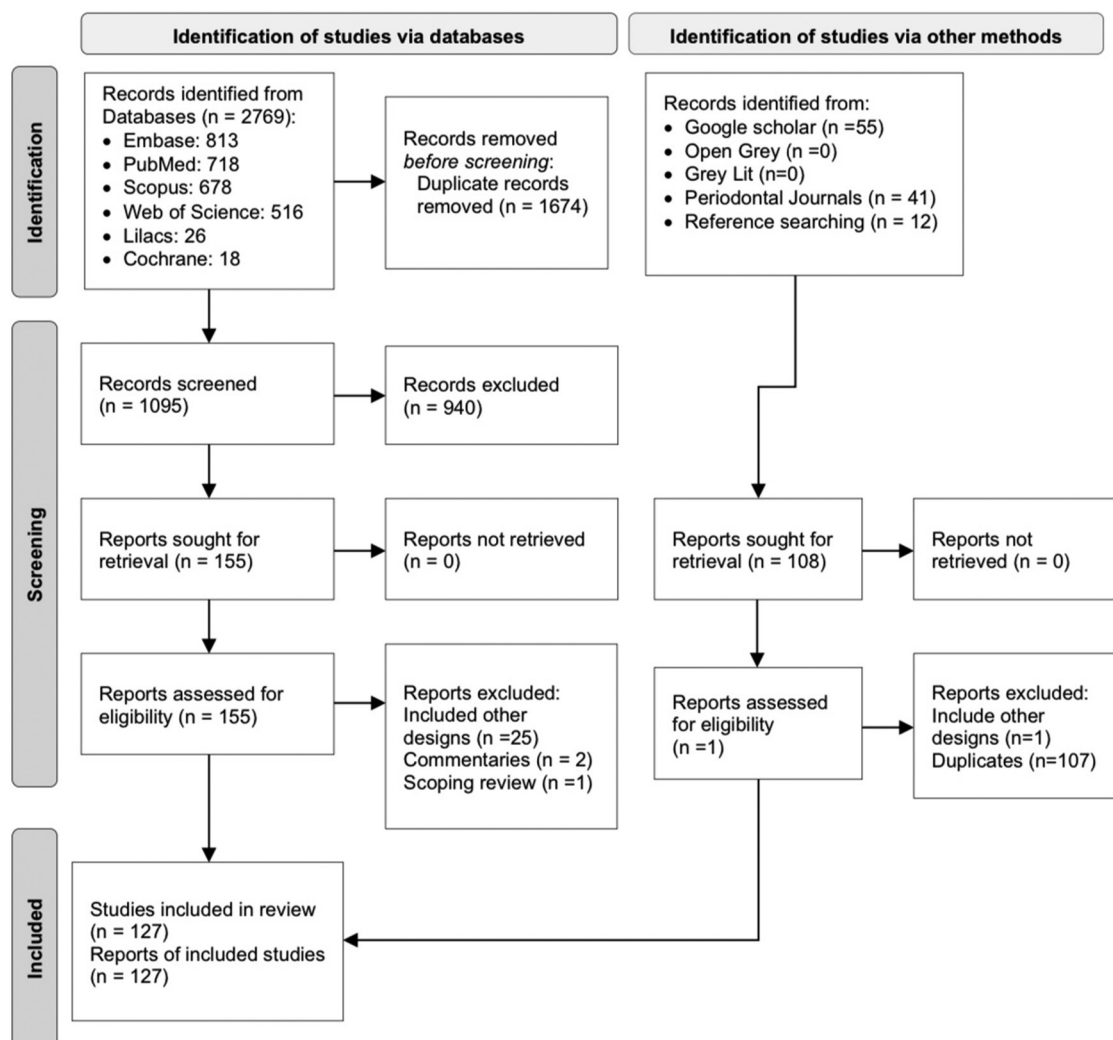


Fig. 1: PRISMA flowchart.

#### -Characteristics of the studies

Studies characteristics are presented in Tables S1-S4 (Supplement 5) ([http://www.medicinaoral.com/medoral-free01/aop/jced\\_60197\\_s05.pdf](http://www.medicinaoral.com/medoral-free01/aop/jced_60197_s05.pdf)). The main language of publication was English (n= 124; 97.6%). The most common journals where the SRs were published were: Journal of Clinical Periodontology (n=17; 13.4%), Clinical Oral Investigations (n=10; 7.9%) and BMC Oral Health (n=9; 7.1%). Only 34.6% (n=44) of SRs were open access, and 66.1% (n=84) were available without restriction. Ninety-three (73.2%) journals had JCR impact factor with mean of 4.4±2.5. One-hundred three (96.8%) journals had h5 index, with an average of 46.1±29.3 (1.6 – 9.3).

#### -Origin and authorship

The number of authors ranged from 1 to 24 (mean 5.3 ± 2.6). Twenty-seven per cent of SRs had authors from different continents sharing the authorship (n=34). Few SRs included an epidemiologist (n=6, 4.7%), a librarian (n=9, 7.1%) and a biostatistician (n=9, 7.1%) in the research team.

#### -Main topics of study

The main intervention topics addressed by SRs were related to: antiseptics (n=23; 18.1%), soft tissue regeneration (n=20; 15.7%), periodontal photomedicine (n=18; 14.2%), antimicrobials (n=14; 10.9%) and guided bone regeneration (n=13; 10.2%).

The main periodontal conditions treated were: periodontitis (n=37; 29.1%), gingivitis and plaque/oral hygiene (n=33; 26.0%), gingival recessions (n=14; 11.0%), peri-implant diseases (n=13; 10.2%) and periodontal bone defects (n=12; 9.4%).

#### -Funding and conflict of interest

The main source of funding was: own financing (n=48; 37.8%), government/university (n=45; 35.4%) and industry (n=4; 3.1%). Twenty-seven SRs (21.3%) did not report funding sources. One-hundred eight SRs (85.0%) reported not having conflict of interest.

#### -Protocol, register and PRISMA

More than a third of the SRs (n=48; 37.8%) did not mention a study protocol. Among the 79 SRs that reported a protocol, 71 (89.9%) and 8 (10.1%) had registered and non-registered protocols, respectively. The most common registration platform was the International Prospective Register of Systematic Reviews (PROSPERO) (n=61; 85.9%).

#### -Searches and eligibility restrictions

A mean of 3.6 (±1.5) databases were searched, and 1.2 (±1.5) databases for grey literature. A manual search on the articles' references was performed by 70.9% (n=90) of SRs, and 41.7% (n=53) of SRs manually searched in journals' interest area. The most common databases searched were: MedLine (n=125; 98.4%), Cochrane Library (n=97; 76.4%), Embase (n=70; 55.1%), Web of Science (n=40; 31.5%), Scopus (n=38; 29.9%), Lilacs (n=16; 12.6%) and CINAHL (n=10; 7.9%).

#### -Number of studies and the presence of meta-analysis

The mean number of studies included per SR was 15.9±15.7 (range of 2–91), being 15.6±15.7 (1–91) RCTs and 0.4±1.3 (0–12) nRCTs. The total number of participants ranged from 67 to 13,426 individuals and the mean number of analysed outcomes was 4.6±2.9 (1–19). Most SRs (n=96; 75.6%) presented meta-analysis, 13 (10.2%) presented network meta-analysis, and 8 (6.3%) had a meta-regression. On average, 14.3±16.3 (2–105) studies were included in the meta-analyses and 3.2±1.9 (1–11) outcomes were meta-analysed SR.

Assessment of risk of bias and certainty of the evidence  
The main risk of bias tool used for RCTs was the Cochrane tool (n=103; 81.1%) and ROBINS-I for nRCTs (n=17; 23.5%). Thirty-two (25.2%) SRs assessed the certainty of the evidence using the GRADE approach, and among them, 12 (37.5%) assessed the certainty of the evidence following the GRADE approach guide, 7 (21.9%) partially followed it, and 13 (40.6%) deviated from the GRADE approach guide.

#### -AMSTAR 2

The overall AMSTAR 2 methodological quality of all SRs was classified as critically low (n=82, 64.6%), low (n=31, 24.4%), moderate (n=1, 0.8%) and high (n=13, 10.2%) (Table 1, 1 cont.).

The items 1 (components of PICO), 5 and 6 (study selection and data extraction in duplicate), 9 (satisfactory assessment of risk of bias), 11 (appropriate methods for meta-analysis), 14 (satisfactory discussion of heterogeneity) and 16 (report of sources of conflict of interest) received positive answers in more than 70% of SRs. The items with the highest percentage of overall negative responses were: 3 (reason for selection of certain study designs; 87.4%), 10 (report of funding sources for the included studies; 67.7%) and 4 (careful search of the literature; 65.4%). Five items considered critical according to AMSTAR 2 had large percentage of negative assessments: 2 (presence of protocol and justification for its modifications; 37.8%), 4 (careful literature search; 65.4%), 7 (list of excluded articles with justifications; 41.7%), 13 (consideration of the risk of bias in individual studies; 40.9%) and 15 (investigation and discussion of the impact of publication bias; 40.2%).

It is important to note that when analyses were performed considering the impact factor of the journal, the overall methodological quality was classified as high in ~30% of SRs in journals with impact factor ≥6. A high percentage of positive answers were also observed in the higher impact factor journals (Table 2-2 cont.-1). No expressive differences were observed when evaluating SRs according to the design of the included studies (Table 1, 1 cont.).

#### -ROBIS

The overall ROBIS evaluations considered 113 (90.6%) SRs to be at high risk of bias, 11 (7.1%) at low risk and 3 (2.4%) at unclear risk of bias (Table 3-3 cont.-2).

**Table 1:** Methodological quality assessment through AMSTAR 2 according to the type of studies included in the systematic reviews.

AMSTAR items	Systematic reviews (SRs)		
	All SRs (n = 127)	SRs with RCTs and nRCTs (n = 21)	SRs with RCTs only (n = 106)
1. Did the research questions and inclusion criteria for the review include the components of PICO?			
Yes	119 (93.7)	20 (95.2)	99 (93.4)
No	8 (6.3)	1 (4.8)	7 (6.6)
2. Did the report of the review contain an explicit statement that the review methods were established prior to conduct of the review and did the report justify any significant deviations from the protocol?			
Yes	46 (36.2)	10 (47.6)	36 (34.0)
Partial yes	33 (26.0)	5 (23.8)	28 (26.4)
No	48 (37.8)	6 (28.6)	42 (39.6)
3. Did the review authors explain their selection of the study designs for inclusion in the review?			
Yes	16 (12.6)	2 (9.5)	14 (13.2)
No	111 (87.4)	19 (90.5)	92 (86.8)
4. Did the review authors use a comprehensive literature search strategy?			
Yes	27 (21.3)	2 (9.5)	15 (14.2)
Partial yes	17 (13.4)	9 (42.9)	18 (17.0)
No	83 (65.4)	10 (47.6)	73 (68.9)
5. Did the review authors perform study selection in duplicate?			
Yes	105 (82.7)	18 (85.7)	87 (82.1)
No	22 (17.3)	3 (14.3)	19 (17.9)
6. Did the review authors perform data extraction in duplicate?			
Yes	95 (74.8)	16 (76.2)	79 (74.5)
No	32 (25.2)	5 (23.8)	27 (25.)
7. Did the review authors provide a list of excluded studies and justify the exclusions?			
Yes	74 (58.3)	13 (61.9)	61 (58.5)
No	53 (41.7)	8 (38.1)	45 (42.5)
8. Did the review authors describe the included studies in adequate detail?			
Yes	78 (61.4)	15 (71.4)	63 (59.4)
Partial yes	32 (25.2)	4 (19.0)	28 (26.4)
No	17 (13.4)	2 (9.5)	15 (14.2)
9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?			
Yes	115 (90.6)	16 (76.2)	99 (93.4)
Partial yes	4 (3.1)	1 (4.8)	3 (2.8)
No	8 (6.3)	4 (19.0)	4 (3.8)
10. Did the review authors report on the sources of funding for the studies included in the review?			
Yes	39 (30.7)	6 (28.6)	33 (31.1)
No	88 (69.3)	15 (71.4)	73 (68.9)
11. If meta-analysis was justified did the review authors use appropriate methods for statistical combination of results? (Only complete this item if meta-analysis of other data synthesis techniques were reported)			
Yes	94 (74.0)	15 (71.4)	79 (74.5)
No	2 (1.6)	0 (0.0)	2 (1.9)
No meta-analysis conducted	31 (24.4)	6 (28.6)	25 (23.6)

**Table 1 cont.:** Methodological quality assessment through AMSTAR 2 according to the type of studies included in the systematic reviews.

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	28 (22.0)	3 (14.3)	25 (23.6)
No	68 (53.5)	6 (28.6)	56 (52.8)
No meta-analysis conducted	31 (24.4)	12 (57.1)	25 (23.6)
13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?			
Yes	75 (59.1)	12 (57.1)	43 (40.6)
No	52 (40.9)	9 (42.9)	63 (59.4)
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?			
Yes	98 (77.2)	17 (81.0)	81 (76.4)
Partial yes	0 (0.0)	0 (0.0)	1 (0.9)
No	28 (22.0)	4 (19.0)	24 (22.6)
15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?			
Yes	45 (35.4)	5 (23.8)	40 (37.7)
No	51 (40.2)	10 (47.6)	41 (38.7)
No meta-analysis conducted	31 (24.4)	6 (28.6)	25 (23.6)
16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?			
Yes	116 (91.3)	19 (90.5)	97 (91.5)
No	11 (8.7)	2 (9.5)	9 (8.5)
Overall methodological quality			
High	13 (10.2%)	3 (14.3%)	10 (9.4%)
Moderate	1 (0.8%)	1 (4.8%)	0 (0.0%)
Low	31 (24.4%)	5 (23.8%)	26 (24.5%)
Critically low	82 (64.6%)	12 (57.1%)	70 (66.0%)

RCTs = randomized clinical trials; nRCTs = non-randomized clinical trials.

**Table 2:** Methodological quality assessment through AMSTAR 2 according to the impact factor of the journals.

AMSTAR items	Impact factor of the journal			
	Unranked (n = 34)	<3 (n = 37)	≥3 <6 (n = 32)	≥6 (n = 24)
1. Did the research questions and inclusion criteria for the review include the components of PICO?				
Yes	28 (82.4%)	37 (100.0%)	30 (93.8%)	24 (100.0%)
No	6 (17.6%)	0 (0.0%)	2 (6.3%)	0 (0.0%)
2. Did the report of the review contain an explicit statement that the review methods were established prior to conduct of the review and did the report justify any significant deviations from the protocol?				
Yes	4 (11.8%)	14 (37.8%)	11 (34.4%)	17 (70.8%)
Partial yes	9 (26.5%)	9 (24.3%)	9 (28.1%)	6 (25.0%)
No	21 (61.8%)	14 (37.8%)	12 (37.5%)	1 (4.2%)
3. Did the review authors explain their selection of the study designs for inclusion in the review?				
Yes	2 (5.9%)	5 (13.5%)	3 (9.4%)	6 (25.0%)
No	32 (94.1%)	32 (86.5%)	29 (90.6%)	18 (75.0%)
4. Did the review authors use a comprehensive literature search strategy?				
Yes	1 (2.9%)	4 (10.8%)	4 (12.5%)	8 (33.3%)

**Table 2 cont.:** Methodological quality assessment through AMSTAR 2 according to the impact factor of the journals.

Partial yes	6 (17.6%)	6 (16.2%)	12 (37.5%)	3 (12.5%)
No	27 (79.4%)	27 (73.0%)	16 (50.0%)	13 (54.2%)
5. Did the review authors perform study selection in duplicate?				
Yes	25 (73.5%)	29 (78.4%)	28 (87.5%)	23 (95.8%)
No	9 (26.5%)	8 (21.6%)	4 (12.5%)	1 (4.2%)
6. Did the review authors perform data extraction in duplicate?				
Yes	19 (55.9%)	30 (81.1%)	24 (75.0%)	22 (91.7%)
No	15 (44.1%)	7 (18.9%)	8 (25.0%)	2 (8.3%)
7. Did the review authors provide a list of excluded studies and justify the exclusions?				
Yes	11 (32.4%)	21 (56.8%)	20 (62.5%)	22 (91.7%)
No	23 (67.6%)	16 (43.2%)	12 (37.5%)	2 (8.3%)
8. Did the review authors describe the included studies in adequate detail?				
Yes	18 (52.9%)	20 (54.1%)	18 (56.3%)	22 (91.7%)
Partial yes	10 (24.9%)	10 (27.0%)	10 (31.3%)	2 (8.3%)
No	6 (17.6%)	7 (18.9%)	4 (12.5%)	0 (0.0%)
9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?				
Yes	27 (79.4%)	35 (94.6%)	30 (93.8%)	23 (95.8%)
Partial yes	1 (2.9%)	0 (0.0%)	2 (6.3%)	1 (4.2%)
No	6 (17.6%)	2 (5.4%)	0 (0.0%)	0 (0.0%)
10. Did the review authors report on the sources of funding for the studies included in the review?				
Yes	6 (17.6%)	12 (32.4%)	5 (15.6%)	16 (66.7%)
No	28 (82.4%)	25 (67.6%)	27 (84.4%)	8 (33.3%)
11. If meta-analysis was justified did the review authors use appropriate methods for statistical combination of results? (Only complete this item if meta-analysis of other data synthesis techniques were reported)				
Yes	18 (52.9%)	29 (78.4%)	25 (78.1%)	22 (91.7%)
No	0 (0.0%)	1 (2.7%)	1 (3.1%)	0 (0.0%)
No meta-analysis conducted	16 (47.1%)	7 (18.9%)	6 (18.8%)	2 (8.3%)
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	3 (8.8%)	8 (21.6%)	5 (15.6%)	12 (50.0%)
No	15 (44.1%)	22 (59.5%)	21 (65.6%)	10 (41.7%)
No meta-analysis conducted	16 (47.1%)	7 (18.9%)	6 (18.8%)	2 (8.3%)
13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?				
Yes	12 (35.3%)	20 (54.1%)	21 (65.6%)	22 (91.7%)
No	22 (64.7%)	17 (45.9%)	11 (34.4%)	2 (8.3%)
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?				
Yes	20 (58.8%)	27 (73.0%)	28 (87.5%)	23 (95.8%)
Partial yes	1 (2.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
No	13 (38.2%)	10 (27.0%)	4 (12.5%)	1 (4.2%)
15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?				
Yes	5 (14.7%)	13 (35.1%)	12 (37.5%)	15 (62.5%)
No	13 (38.2%)	17 (45.9%)	14 (43.8%)	7 (29.2%)
No meta-analysis conducted	16 (47.1%)	7 (18.9%)	6 (18.8%)	2 (8.3%)
16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?				
Yes	29 (85.3%)	36 (97.3%)	28 (87.5%)	23 (95.8%)

**Table 2 cont.-1:** Methodological quality assessment through AMSTAR 2 according to the impact factor of the journals.

No	5 (14.7%)	1 (2.7%)	4 (12.5%)	1 (4.2%)
Overall methodological quality				
High	0 (0.0%)	3 (8.1%)	3 (9.4%)	7 (29.2%)
Moderate	0 (0.0%)	0 (0.0%)	1 (3.1%)	0 (0.0%)
Low	5 (14.7%)	5 (13.5%)	10 (31.3%)	11 (45.8%)
Critically low	29 (85.3%)	29 (78.4%)	18 (56.3%)	6 (25.0%)

**Table 3:** Risk of bias assessment through ROBIS according to the type of studies included in the systematic reviews.

ROBIS items	Systematic reviews		
	All SRs (n = 127)	SRs with RCTs and nRCTs (n = 21)	SRs with RCTs only (n = 106)
Q1.1 Did the review adhere to pre-defined objectives and eligibility criteria?			
Yes	46 (36.2)	10 (47.6)	36 (34.0)
Probably yes	17 (13.4)	2 (9.5)	15 (14.2)
Probably no	53 (41.9)	7 (33.3)	46 (43.4)
No	11 (8.7)	2 (9.5)	9 (8.5)
Q1.2 Were the eligibility criteria appropriate for the review question?			
Yes	91 (71.7)	16 (76.2)	75 (70.8)
Probably yes	24 (18.9)	3 (14.3)	21 (19.8)
Probably no	10 (7.9)	2 (9.5)	8 (7.5)
No	2 (1.6)	0 (0.0)	2 (1.9)
Q1.3 Were eligibility criteria unambiguous?			
Yes	105 (82.7)	20 (95.2)	85 (90.2)
Probably yes	13 (10.2)	1 (4.8)	12 (11.3)
Probably no	7 (5.5)	0 (0.0)	7 (6.6)
No	2 (1.6)	0 (0.0)	2 (1.9)
Q1.4 Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?			
Yes	86 (67.7)	17 (81.0)	69 (65.1)
Probably yes	15 (11.8)	2 (9.5)	13 (12.3)
Probably no	18 (14.2)	1 (4.8)	17 (16.0)
No	7 (5.5)	1 (4.8)	6 (5.7)
Not informed	1 (0.8)	0 (0.0)	1 (0.9)
Q1.5 Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?			
Yes	59 (46.5)	15 (71.4)	44 (41.5)
Probably yes	4 (3.1)	1 (4.8)	3 (2.8)
Probably no	32 (24.4)	2 (9.5)	30 (28.3)
No	31 (25.2)	3 (14.3)	28 (26.4)
Not informed	1 (0.8)	0 (0.0)	1 (0.9)
Concerns regarding specification of study eligibility criteria			
High	95 (74.8)	10 (47.6)	85 (80.2)
Low	32 (25.2)	11 (52.4)	21 (19.8)
Unclear	0 (0.0)	0 (0.0)	0 (0.0)
Q2.1 Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?			
Yes	70 (55.1)	9 (42.9)	61 (57.5)
No	56 (44.1)	12 (57.1)	44 (41.5)

**Table 3 cont.:** Risk of bias assessment through ROBIS according to the type of studies included in the systematic reviews.

Not informed	1 (0.8)	0 (0.0)	1 (0.9)
Q2.2 Were methods additional to database searching used to identify relevant reports?			
Yes	53 (41.7)	10 (47.6)	43 (40.6)
Probably yes	4 (3.1)	0 (0.0)	4 (3.8)
Probably no	0 (0.0)	0 (0.0)	0 (0.0)
No	69 (54.3)	11 (52.4)	58 (54.7)
Not informed	1 (0.8)	0 (0.0)	1 (0.9)
Q2.3 Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?			
Yes	46 (36.2)	8 (38.1)	38 (35.8)
Probably yes	22 (17.3)	3 (14.3)	19 (17.9)
Probably no	40 (31.5)	9 (42.9)	31 (29.2)
No	13 (10.2)	1 (4.8)	12 (11.3)
Not informed	6 (4.7)	0 (0.0)	6 (5.6)
Q2.4 Were restrictions based on date, publication format, or language appropriate?			
Yes	47 (37.0)	14 (66.7)	33 (31.1)
Probably yes	3 (2.4)	0 (0.0)	3 (2.8)
Probably no	37 (29.1)	3 (14.3)	34 (32.1)
No	38 (29.9)	4 (19.0)	34 (32.1)
Not informed	2 (1.6)	0 (0.0)	2 (1.9)
Q2.5 Were efforts made to minimize error in selection of studies?			
Yes	88 (69.3)	17 (81.0)	71 (67.0)
Probably yes	14 (11.0)	1 (4.8)	13 (12.3)
Probably no	0 (0.0)	0 (0.0)	0 (0.0)
No	2 (1.6)	1 (4.8)	1 (0.9)
Not informed	23 (18.1)	2 (9.5)	21 (19.8)
Concerns regarding specification of study eligibility criteria			
High	103 (81.1)	18 (85.7)	85 (80.2)
Low	23 (18.1)	2 (9.5)	21 (19.8)
Unclear	1 (0.8)	1 (4.8)	0 (0.0)
Q3.1 Were efforts made to minimize error in data collection?			
Yes	74 (58.3)	15 (71.4)	59 (55.7)
Probably yes	21 (16.5)	1 (4.8)	20 (18.9)
Probably no	2 (1.6)	1 (4.8)	1 (0.9)
No	0 (0.0)	0 (0.0)	0 (0.0)
Not informed	30 (23.6)	4 (19.0)	26 (24.5)
Q3.2 Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?			
Yes	79 (62.2)	15 (71.4)	64 (60.4)
Probably yes	12 (9.4)	2 (9.5)	10 (9.4)
Probably no	21 (16.5)	3 (14.3)	18 (17.0)
No	15 (11.8)	1 (4.8)	14 (13.2)
Q3.3 Were all relevant study results collected for use in the synthesis?			
Yes	69 (54.3)	11 (52.4)	58 (54.7)
Probably yes	35 (27.6)	6 (28.6)	29 (27.4)
Probably no	14 (11.0)	4 (19.0)	10 (9.4)
No	8 (6.3)	0 (0.0)	8 (7.5)
Not informed	1 (0.8)	0 (0.0)	1 (0.9)
Q3.4 Was risk of bias (or methodological quality) formally assessed using appropriate criteria?			
Yes	98 (77.2)	11 (52.4)	87 (82.1)

**Table 3 cont.-1:** Risk of bias assessment through ROBIS according to the type of studies included in the systematic reviews.

Probably yes	19 (15.0)	5 (23.8)	14 (13.2)
Probably no	6 (4.7)	4 (19.0)	2 (1.9)
No	4 (3.1)	1 (4.8)	3 (2.8)
Q3.5 Were efforts made to minimize error in risk of bias assessment?			
Yes	69 (54.3)	17 (81.0)	52 (49.1)
Probably yes	8 (6.3)	0 (0.0)	8 (7.5)
Probably no	2 (1.6)	1 (4.8)	1 (0.9)
No	0 (0.0)	0 (0.0)	0 (0.0)
Not informed	48 (37.8)	3 (14.3)	45 (42.5)
Concerns regarding methods used to collect data and appraise studies			
High	46 (36.2)	9 (42.9)	34 (34.9)
Low	44 (34.6)	9 (42.9)	36 (34.0)
Unclear	37 (36.2)	3 (14.3)	33 (31.1)
Q4.1 Did the synthesis include all studies that it should?			
Yes	61 (48.0)	10 (47.6)	51 (48.1)
Probably yes	56 (44.1)	9 (42.9)	47 (44.3)
Probably no	9 (7.1)	2 (9.5)	7 (6.6)
No	0 (0.0)	0 (0.0)	0 (0.0)
Not informed	1 (0.8)	0 (0.0)	1 (0.9)
Q4.2 Were all pre-defined analyses reported or departures explained?			
Yes	56 (44.1)	12 (57.1)	44 (41.5)
Probably yes	6 (4.7)	1 (4.8)	5 (4.7)
Probably no	3 (2.4)	0 (0.0)	3 (2.8)
No	11 (8.7)	2 (9.5)	9 (8.5)
Not informed	51 (40.2)	6 (28.6)	45 (42.5)
Q4.3 Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?			
Yes	94 (74.0)	17 (81.0)	77 (72.6)
Probably yes	30 (23.6)	4 (19.0)	26 (24.5)
Probably no	0 (0.0)	0 (0.0)	0 (0.0)
No	3 (2.4)	0 (0.0)	3 (2.8)
Q4.4 Was between-study variation (heterogeneity) minimal or addressed in the synthesis?			
Yes	74 (58.3)	11 (52.4)	63 (59.4)
Probably yes	32 (25.2)	7 (33.3)	25 (23.6)
Probably no	12 (9.4)	1 (4.8)	11 (10.4)
No	9 (7.1)	2 (9.5)	7 (6.6)
Q4.5 Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?			
Yes	19 (15.0)	3 (14.3)	16 (15.1)
Probably yes	13 (10.2)	3 (14.3)	10 (9.4)
Probably no	45 (35.4)	6 (28.6)	39 (36.8)
No	50 (39.4)	9 (42.9)	41 (38.7)
Q4.6 Were biases in primary studies minimal or addressed in the synthesis?			
Yes	17 (13.4)	3 (14.3)	14 (13.2)
Probably yes	31 (24.4)	1 (4.8)	30 (28.3)
Probably no	40 (31.5)	7 (33.3)	33 (31.1)
No	39 (30.7)	10 (47.6)	29 (27.4)
Concerns regarding specification of study eligibility criteria			
High	110 (86.6)	20 (95.2)	90 (84.9)

**Table 3 cont.-2:** Risk of bias assessment through ROBIS according to the type of studies included in the systematic reviews.

Low	12 (9.4)	1 (4.8)	11 (10.4)
Unclear	5 (3.9)	0 (0.0)	5 (4.7)
Overall risk of bias			
High	115 (90.6%)	18 (85.7%)	97 (91.5%)
Low	9 (7.1%)	3 (14.3%)	6 (5.7%)
Unclear	3 (2.4%)	0 (0.0%)	3 (2.8%)

RCTs = randomized clinical trials; nRCTs = non-randomized clinical trials.

Seventy-five per cent (n=95) SRs were judged with high risk on domain 1 (study eligibility criteria), and 25.2% (n=32) were judged as low risk of bias. The main issues were unjustified absence or deviation from the protocol (Q1.1; n=64; 50.6%) and unjustifiable restrictions in the eligibility criteria (Q1.5; n=63; 49.6%).

In domain 2 (identification and selection of studies), 81.1% of SRs (n=103) were at high risk and 18.1% (n=23) low risk of bias. The main issues were related to deficiencies in the literature searching in the main databases (Q2.1; n=56; 44.1%) and in the complementary searches (Q2.2; n=69; 54.2%).

The domain 3 (data collection and study appraisal), 36.2% of RSTs (n=46) were at high risk of bias, 34.6% (n=44) low risk of bias and 29.1% (n=37) were at unclear risk of bias. The main issues were related to the lack of details of the included studies (Q3.2; n=36; 28.3%) and lack of the inclusion of relevant study results (Q3.3; n=22; 17.3 %). It is noteworthy that this was the domain with the highest percentage of undefined risk of

bias, mainly due to use of an inappropriate risk of bias tool (Q3.4; n=48; 37.8%) and lack of independent reviewers to extract data (Q3.1; n=30; 23.6%).

The domain 4 (synthesis and findings) accounted with the highest overall risk of bias among the 4 domains: 110 (86.6%) SRs had high risk, while 12 (9.4%) SRs had low risk and 5 (3.9%) had unclear risk of bias. The main issues were the lack additional analysis or synthesis to test robustness of the results (Q4.5; n=95; 74.8%) and lack of assessment of the high risk of bias studies in the synthesis of results (Q4.6; n=79; 62.2%).

No expressive differences were observed when SRs were evaluated according to the design of the included studies (Tables 3). However, the risk of bias decreased with the increase of the impact factor of the journal (Table 4). Detailed ROBIS assessments such as concerns regarding study eligibility criteria, methods used to collect data and appraise studies and the synthesis and findings also decreased with the increase of the impact factor of the journal (Table 4- 4 cont.-2).

**Table 4:** Risk of bias assessment through ROBIS according to the impact factor of the journals.

ROBIS items	Impact factor of the journal			
	Unranked (n = 34)	<3 (n = 37)	≥3 <6 (n = 32)	≥6 (n = 24)
Q1.1 Did the review adhere to pre-defined objectives and eligibility criteria?				
Yes	5 (14.7%)	16 (43.2%)	11 (34.4%)	14 (58.3%)
Probably yes	3 (8.8%)	5 (13.5%)	4 (12.5%)	5 (20.8%)
Probably no	24 (70.6%)	14 (37.8%)	13 (40.6%)	2 (8.3%)
No	2 (5.9%)	2 (5.4%)	4 (12.5%)	3 (12.5%)
Q1.2 Were the eligibility criteria appropriate for the review question?				
Yes	18 (52.9%)	27 (73.0%)	27 (84.4%)	19 (79.2%)
Probably yes	7 (20.6%)	10 (27.0%)	3 (9.4%)	4 (16.7%)
Probably no	8 (23.5%)	0 (0.0%)	1 (3.1%)	1 (4.2%)
No	1 (2.9%)	0 (0.0%)	1 (3.1%)	0 (0.0%)
Q1.3 Were eligibility criteria unambiguous?				
Yes	28 (82.4%)	28 (75.7%)	27 (84.4%)	22 (91.7%)
Probably yes	4 (11.8%)	7 (18.9%)	1 (3.1%)	1 (4.2%)
Probably no	2 (5.9%)	2 (5.4%)	3 (9.4%)	0 (0.0%)
No	0 (0.0%)	0 (0.0%)	1 (3.1%)	1 (4.2%)

**Table 4 cont.:** Risk of bias assessment through ROBIS according to the impact factor of the journals.

Q1.4 Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?				
Yes	23 (67.6%)	22 (59.5%)	21 (65.5%)	20 (83.3%)
Probably yes	4 (11.8%)	4 (10.8%)	4 (12.5%)	3 (12.5%)
Probably no	6 (17.6%)	6 (16.2%)	6 (18.8%)	0 (0.0%)
No	1 (2.9%)	4 (10.8%)	1 (3.1%)	1 (4.2%)
Not informed	0 (0.0%)	1 (2.7%)	0 (0.0%)	0 (0.0%)
Q1.5 Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?				
Yes	9 (26.5%)	15 (40.5%)	20 (62.5%)	15 (62.5%)
Probably yes	12 (35.3%)	3 (8.1%)	1 (3.1%)	0 (0.0%)
Probably no	1 (2.9%)	9 (24.3%)	9 (28.1%)	2 (8.3%)
No	12 (35.3%)	10 (27.0%)	2 (6.3%)	7 (29.2%)
Not informed	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Concerns regarding specification of study eligibility criteria				
High	30 (88.2%)	30 (81.1%)	23 (71.9%)	12 (50.0%)
Low	4 (11.8%)	7 (18.9%)	9 (28.1%)	12 (50.0%)
Unclear	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Q2.1 Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?				
Yes	18 (52.9%)	17 (45.9%)	19 (59.4%)	16 (66.7%)
No	15 (44.1%)	20 (54.1%)	13 (40.6%)	8 (33.3%)
Not informed	1 (2.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Q2.2 Were methods additional to database searching used to identify relevant reports?				
Yes	6 (17.6%)	14 (37.8%)	16 (50.0%)	17 (70.8%)
Probably yes	1 (2.9%)	2 (5.4%)	1 (3.1%)	0 (0.0%)
Probably no	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
No	26 (76.5%)	21 (56.8%)	15 (46.9%)	7 (29.2%)
Not informed	1 (2.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Q2.3 Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?				
Yes	3 (8.8%)	12 (32.4%)	11 (34.4%)	20 (83.3%)
Probably yes	7 (20.6%)	4 (10.8%)	9 (28.1%)	2 (8.3%)
Probably no	14 (41.2%)	14 (37.8%)	11 (34.4%)	1 (4.2%)
No	6 (17.6%)	7 (18.9%)	0 (0.0%)	0 (0.0%)
Not informed	4 (11.8%)	0 (0.0%)	1 (3.1%)	1 (4.2%)
Q2.4 Were restrictions based on date, publication format, or language appropriate?				
Yes	8 (23.5%)	11 (29.7%)	15 (46.9%)	13 (54.2%)
Probably yes	0 (0.0%)	3 (8.1%)	0 (0.0%)	0 (0.0%)
Probably no	12 (35.3%)	13 (35.1%)	8 (25.0%)	4 (16.7%)
No	14 (41.2%)	9 (24.3%)	1 (3.1%)	7 (29.2%)
Not informed	0 (0.0%)	1 (2.7%)	8 (25.0%)	0 (0.0%)
Q2.5 Were efforts made to minimize error in selection of studies?				
Yes	21 (61.8%)	23 (62.2%)	24 (75.0%)	20 (83.3%)
Probably yes	2 (5.9%)	4 (10.8%)	7 (21.9%)	1 (4.2%)
Probably no	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
No	1 (2.9%)	1 (2.9%)	0 (0.0%)	0 (0.0%)
Not informed	10 (29.4%)	9 (24.3%)	1 (3.1%)	3 (12.5%)
Concerns regarding specification of study eligibility criteria				
High	32 (94.1%)	34 (91.9%)	23 (71.9%)	14 (58.3%)

**Table 4 cont.-1:** Risk of bias assessment through ROBIS according to the impact factor of the journals.

Low	2 (5.9%)	3 (8.1%)	9 (28.1%)	9 (37.5%)
Unclear	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (4.2%)
Q3.1 Were efforts made to minimize error in data collection?				
Yes	14 (41.2%)	21 (56.8%)	18 (56.3%)	21 (87.5%)
Probably yes	5 (14.7%)	9 (24.3%)	6 (18.8%)	1 (4.2%)
Probably no	1 (2.9%)	1 (2.7%)	0 (0.0%)	0 (0.0%)
No	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Not informed	14 (41.2%)	6 (16.2%)	8 (25.0%)	2 (8.3%)
Q3.2 Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?				
Yes	18 (52.9%)	21 (56.8%)	18 (56.3%)	22 (91.7%)
Probably yes	5 (14.7%)	3 (8.1%)	4 (12.5%)	0 (0.0%)
Probably no	4 (11.8%)	9 (24.3%)	6 (18.8%)	2 (8.3%)
No	7 (20.6%)	4 (10.8%)	4 (12.5%)	0 (0.0%)
Q3.3 Were all relevant study results collected for use in the synthesis?				
Yes	12 (35.3%)	18 (48.6%)	20 (62.5%)	19 (79.2%)
Probably yes	11 (32.4%)	13 (35.1%)	7 (21.9%)	4 (16.7%)
Probably no	7 (20.6%)	4 (10.8%)	2 (6.3%)	1 (4.2%)
No	4 (11.8%)	1 (2.7%)	3 (9.4%)	0 (0.0%)
Not informed	0 (0.0%)	1 (2.7%)	0 (0.0%)	0 (0.0%)
Q3.4 Was risk of bias (or methodological quality) formally assessed using appropriate criteria?				
Yes	20 (58.8%)	29 (78.4%)	26 (81.3%)	23 (95.8%)
Probably yes	7 (20.6%)	6 (16.2%)	5 (15.6%)	1 (4.2%)
Probably no	3 (8.8%)	2 (5.4%)	1 (3.1%)	0 (0.0%)
No	4 (11.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Q3.5 Were efforts made to minimize error in risk of bias assessment?				
Yes	17 (50.0%)	19 (51.4%)	19 (59.4%)	14 (58.3%)
Probably yes	0 (0.0%)	5 (13.5%)	2 (6.3%)	1 (4.2%)
Probably no	1 (2.9%)	1 (2.7%)	0 (0.0%)	0 (0.0%)
No	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Not informed	16 (47.1%)	12 (32.4%)	11 (34.4%)	9 (37.5%)
Concerns regarding methods used to collect data and appraise studies				
High	17 (50.0%)	15 (40.5%)	12 (37.5%)	2 (8.3%)
Low	7 (20.6%)	12 (32.4%)	12 (37.5%)	14 (58.3%)
Unclear	10 (29.4%)	10 (27.0%)	8 (25.0%)	8 (33.3%)
Q4.1 Did the synthesis include all studies that it should?				
Yes	15 (44.1%)	17 (45.9%)	16 (50.0%)	13 (54.2%)
Probably yes	16 (47.1%)	15 (40.5%)	15 (46.9%)	10 (41.7%)
Probably no	2 (5.9%)	5 (13.5%)	1 (3.1%)	1 (4.2%)
No	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Not informed	1 (2.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Q4.2 Were all pre-defined analyses reported or departures explained?				
Yes	6 (17.6%)	18 (48.6%)	14 (43.8%)	18 (75.0%)
Probably yes	2 (5.9%)	1 (2.7%)	1 (3.1%)	2 (8.3%)
Probably no	2 (5.9%)	0 (0.0%)	0 (0.0%)	1 (4.2%)
No	3 (8.8%)	3 (8.1%)	3 (9.4%)	2 (8.3%)
Not informed	21 (61.8%)	15 (40.5%)	14 (43.8%)	1 (4.2%)

**Table 4 cont.-2:** Risk of bias assessment through ROBIS according to the impact factor of the journals.

Q4.3 Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?				
Yes	22 (64.7%)	24 (64.9%)	27 (84.4%)	21 (87.5%)
Probably yes	12 (35.3%)	10 (27.0%)	5 (15.6%)	3 (12.5%)
Probably no	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
No	0 (0.0%)	3 (8.1%)	0 (0.0%)	0 (0.0%)
Q4.4 Was between-study variation (heterogeneity) minimal or addressed in the synthesis?				
Yes	17 (50.0%)	20 (54.1%)	20 (62.5%)	17 (70.8%)
Probably yes	10 (29.4%)	7 (18.9%)	8 (25.0%)	7 (29.2%)
Probably no	4 (11.8%)	5 (13.5%)	3 (9.4%)	0 (0.0%)
No	3 (8.8%)	5 (13.5%)	1 (3.1%)	0 (0.0%)
Q4.5 Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?				
Yes	0 (0.0%)	4 (10.8%)	5 (15.6%)	10 (41.7%)
Probably yes	1 (2.9%)	5 (13.5%)	4 (12.5%)	3 (12.5%)
Probably no	11 (32.4%)	11 (29.7%)	14 (43.8%)	9 (37.5%)
No	22 (64.7%)	17 (45.9%)	9 (28.1%)	2 (8.3%)
Q4.6 Were biases in primary studies minimal or addressed in the synthesis?				
Yes	1 (2.9%)	4 (10.8%)	3 (9.4%)	9 (37.5%)
Probably yes	4 (11.8%)	11 (29.7%)	10 (31.3%)	6 (25.0%)
Probably no	12 (35.3%)	10 (27.0%)	10 (31.3%)	8 (33.3%)
No	17 (50.0%)	12 (32.4%)	9 (28.1%)	1 (4.2%)
Concerns regarding the synthesis and findings				
High	34 (100.0%)	32 (86.5%)	29 (90.6%)	15 (62.5%)
Low	0 (0.0%)	3 (8.1%)	1 (3.1%)	8 (33.3%)
Unclear	0 (0.0%)	2 (5.4%)	2 (6.3%)	1 (4.2%)
Overall risk of bias				
High	34 (100.0%)	36 (97.3%)	29 (90.6%)	15 (62.5%)
Low	0 (0.0%)	1 (2.7%)	3 (9.4%)	8 (33.3%)
Unclear	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (4.2%)

**Discussion**

The majority of SRs were classified as high risk of bias according to the ROBIS that agreed with the low methodological quality of the AMSTAR 2. It seems that both tools can indicate similar results as they point out in the same direction. This is in accordance with a recent review classifying SRs in dentistry as low and critically low quality (18)

A wide variety of methodological deficiencies resulted in the classification of SRs as having high risk of bias. The absence or unjustified changes of the study protocol was the most important issue according to both tools. The prior creation and registration of a protocol is essential for ensuring the transparency of study methods and allowing adequate peer review of the proposed methodology, thus avoiding the selective reporting bias (11). The deficiency of search strategies was another important identified bias. Search strategies for SRs should be as extensive as possible, without unjustifiable restrictions, including searches in the references of selected

studies and in clinical trial registries. Additionally, complementary searches constitute an important source for the identification of potential studies. Its absence or unjustified restrictions increases the possibility of publication, language and selection biases, among others (11). Among the nine SRs that included a librarian on the research team, 77.8% had high methodological quality searches when assessed by the AMSTAR 2, in contrast to 17.8% of high-quality searches in SRs not including librarians. The inclusion of librarians, although not mandatory, is beneficial as it provides guidance at various stages of the research, such as in the processes of designing search strategies and is associated with more reproducible searches and improved methodological reporting in dental medicine SRs (19). The processes of selection, data extraction and assessment of the risk of bias, which should be ideally carried out independently by more than one reviewer, were presented incompletely in most of the SRs. Cross-checking or duplicate selection processes, data extraction and as-

assessment of risk of bias can reduce biases, as well as the potential subjectivity of one single reviewer (20).

In addition to factors associated with methodological processes, the lack of robustness of the results and excessive bias in primary studies also lead to negative classification through the ROBIS assessment. Findings from SRs, especially those with meta-analyses, must be evaluated through complementary tests to assess its robustness, such as sensitivity tests, subgroup analysis, meta-regression and funnel plots (5). Few studies have proven the robustness of their findings, and the absence of such tests can result in false positive inferences in a meta-analysed result, leading the reader to believe in ineffective treatments.

It was reported that 68% of RCTs in the field of dentistry had an unclear or high risk of bias, according to the Cochrane risk of bias domains (21). If SRs do not test the robustness by meta-analytic approaches such as sensitivity analysis and meta-regression, the overall evidence may be biased. The inclusion of non-randomized intervention studies in the SRs might be considered an indication of acceptance of less-than-adequate research designs for intervention studies leading to low methodological quality or high risk of bias classifications. Nevertheless, no expressive differences were observed when SRs were evaluated according to the design of the included studies.

The vast majority of SRs were of low and critically low quality when assessed by the AMSTAR 2 and judged as high risk of bias by the ROBIS. Overall, these two instruments led to similar conclusions in 93.7% of the assessments, although they are intended for different purposes. The first one is designed to assess the methodological quality of RS, or if the important aspects of the methods are being full filled (3). The second one can detect the risk of bias, so, although the SRs had full filled one item, it does not mean that is free of bias (5). This high agreement is probably due to the overlapping questions between these instruments (22), as well as the low general methodological quality of the SRs analysed. The main source of SRs was the collaboration among authors from different continents (26.8%) and most SRs (97.6%) were published in English. This trend demonstrates the globalization of world science with authors from different countries resulting in international partnerships, exchange of knowledge and resources between research groups and a greater visibility of scientific research (23).

Regarding the scope of the journals, 44.9% of SRs were published in general dental journals. This can be partially explained by the high percentage of studies (26%) whose interventions aimed at improving oral hygiene habits (plaque reduction and gingivitis), areas of common interest in most dental specialties. In addition, it is important to note that some periodontology journals

are no longer accepting submissions of reviews. It was recently reported that there are no significant differences between moderate/high and low/critically low methodological quality SRs in dentistry regarding publication year, continent, dental specialty and the impact factor of the journal (18). On the contrary, the increase in the impact factor of the journal decreased the risk of bias according to ROBIS in the present study.

Few RSs (7.9%) did not mention conflict of interest in the paper at all or did not mention about funding (21.3%). The presence of financial ties can be associated with positive outcomes in RCTs (24). In addition, a survey of 3,247 scientists funded by the US National Institutes of Health showed that 15.5% admitted to altering a study's design, methods, or results in response to pressure from funding sources (25). Thus, reporting potential conflicts of interest and funding sources is mandatory in scientific publications, as they aim to demonstrate the transparency and impartiality of the researchers who carry out the studies (11).

Only a quarter of SRs assessed the certainty of evidence using the GRADE approach. The assessment of the certainty of the evidence is important to help interpreting the results. As it is a more conservative approach, it can help to avoid misleading conclusions (26). Therefore, any SRs of intervention, independent of the field of science, should add the analysis of the certainty of the body of evidence in their methods (26).

Methodological and structural variability among systematic reviews have been observed and the quality of some studies is expected to vary (7). Notwithstanding the systematic and stringent approaches, not all systematic reviews are conducted and reported in the same manner and high methodological quality are uncommon according to specific checklists (7-10). Quality assessments of systematic reviews are quite recent and researchers should consider some guidelines when designing, conducting and reporting their reviews.

In the contemporary scientific scenario, it has been speculated that some issues may influence the quality, reliability and bias of currently scientific research such as the pressure for scientific publication, large volume of articles, predatory journals, quality of the peer review process, among others (27-33). It was also reported that the dental literature has been increasingly reviewed on various topics leading to SRs with questionable clinical or scientific value in terms of up-to-date information to advance knowledge (34). Overall, researches should critically reflect on these issues in order to their scientific production be aligned to core principles of evidence-based dentistry. Guidelines and quality assessment tools may be helpful to identify topics to be improved.

Some limitations of the present study should be discussed. It had three pairs of independent reviewers, which may have resulted in different classifications by the

peers. However, in order to establish solid classification criteria and to achieve high levels of agreement, four training and calibration sessions were conducted using the guidance documents of AMSTAR 2 (3) and ROBIS (5). A certain degree of variability in inter-examiner agreement was previously demonstrated (22, 35). This methodological review is strong as it is the first one that raised the methodological quality using the new AMSTAR 2 together with ROBIS for risk of bias. Moreover, we extracted data of several characteristics of included SRs that are detailed in the supplementary material.

## Conclusions

Most SRs of intervention studies in periodontology were classified as low methodological quality and high risk of bias. Methodological quality increased and risk of bias decreased with the increase in the impact factor of the journals. Although designed for different purposes, both AMSTAR 2 and ROBIS could lead to similar directions. Efforts should be direct to better adherence to reporting guidelines and stricter research practices when conducting SRs. AMSTAR 2 and ROBIS could help the authors to plan the protocol and the reporting of their SRs.

## References

- Cuello-Garcia CA, Morgan RL, Brozek J, Santesso N, Verbeek J, Thayer K, et al. A scoping review and survey provide 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.
- Glenny AM, Esposito M, Coulthard P, Worthington HV. The assessment of systematic reviews in dentistry. *Eur J Oral Sci.* 2003;111:85-92.
- 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 (Clinical research ed.)*. 2017;358:j4008.
- Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson 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.
- Whiting P, Savović J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol.* 2016;69:225-234.
- Faggion CM Jr, Giannakopoulos NN. Critical appraisal of systematic reviews on the effect of a history of periodontitis on dental implant loss. *J Clin Periodontol.* 2013;40:542-552.
- Elangovan S, Avila-Ortiz G, Johnson GK, Karimbux N, Allareddy V. Quality assessment of systematic reviews on periodontal regeneration in humans. *J Periodontol.* 2013;84:176-185.
- Hasuike A, Iguchi S, Suzuki D, Kawano E, Sato S. Systematic review and assessment of systematic reviews examining the effect of periodontal treatment on glycemic control in patients with diabetes. *Med Oral Patol Oral Cir Bucal.* 2017;22:e167-e176.
- Natto ZS, Hameedaldain A. Methodological Quality Assessment of Meta-analyses and Systematic Reviews of the Relationship Between Periodontal and Systemic Diseases. *J Evid Based Dent Pract.* 2019;19:131-139.
- Hasuike A, Ueno D, Nagashima H, Kubota T, Tsukune N, Watanabe N, et al. Methodological quality and risk-of-bias assessments in systematic reviews of treatments for peri-implantitis. *J Periodontol Res.* 2019;54:374-387.
- Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA. (editors). (Updated February 2021) *Cochrane Handbook for Systematic Reviews of Interventions* version 6.2. Cochrane, Retrieved from [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook).
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ (Clinical research ed.)* 2021;372:71.
- Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ (Clinical research ed.)*. 2016;355:i4919.
- Beller EM, Chen JK, Wang UL, Glasziou PP. Are systematic reviews up-to-date at the time of publication? *Systematic Reviews.* 2013;2:36.
- Sampson M, Shojania KG, Garrity C, Horsley T, Ocampo M, Moher D. Systematic reviews can be produced and published faster. *J Clin Epidemiol.* 2008;61:531-536.
- Andersen MZ, Fønnes S, Andresen K, Rosenberg J. Most published meta-analyses were made available within two years of protocol registration. *Eur J Integr Med.* 2021;44:101342.
- Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev.* 2016;5:210.
- Pauletto P, Polmann H, Réus JC, Oliveira JMD, Chaves D, Lehmkühl K, et al. Critical appraisal of systematic reviews of intervention in dentistry published between 2019-2020 using the AMSTAR 2 tool. *Evid Based Dent.* 2022 Sep 14.
- Schellinger J, Sewell K, Bloss JE, Ebron T, Forbes C. The effect of librarian involvement on the quality of systematic reviews in dental medicine. *PLoS One.* 2021;16:e0256833.
- Aromataris E, Munn Z. (Editors). *JBIM Manual for Evidence Synthesis*. Joanna Briggs Institute. 2020. Retrieved from <https://synthesis-manual.jbi.global>.
- Saltaji H, Armijo-Olivo S, Cummings GG, Amin M, Flores-Mir C. Randomized clinical trials in dentistry: Risks of bias, risks of random errors, reporting quality, and methodologic quality over the years 1955-2013. *PLoS One.* 2017;12:e0190089.
- Perry R, Whitmarsh A, Leach V, Davies P. A comparison of two assessment tools used in overviews of systematic reviews: ROBIS versus AMSTAR-2. *Syst Rev.* 2021;10:273.
- Wagner CS, Whetsell TA, Mukherjee S. International research collaboration: Novelty, conventionality, and atypicality in knowledge recombination. *Res Policy.* 2019;48:1260-1270.
- Ahn R, Woodbridge A, Abraham A, Saba S, Korenstein D, Madden E, et al. Financial ties of principal investigators and randomized controlled trial outcomes: cross sectional study. *BMJ (Clinical research ed.)*. 2017;356:i6770.
- Martinson BC, Anderson MS, de Vries R. Scientists behaving badly. *Nature.* 2005;435:737-738.
- Zhang Y, Akl EA, Schünemann HJ. Using systematic reviews in guideline development: the GRADE approach. *Res Synt Methods.* 2018:1-18.
- Neill US. Publish or perish, but at what cost? *J Clin Investig.* 2008;118:2368.
- Else H, Van Noorden R. The fight against fake-paper factories that churn out sham science. *Nature.* 2021;591:516-519.
- Bohannon J. Who's afraid of peer review? *Science (New York).* 2013;342:60-65.
- Van Noorden R. Hundreds of gibberish papers still lurk in the scientific literature. *Nature.* 2021;594:160-161.
- Fanelli D. Do pressures to publish increase scientists' bias? An empirical support from US States Data. *PLoS One.* 2010;4:e10271.
- Tiokhin L, Yan M, Morgan T. Competition for priority harms the reliability of science, but reforms can help. *Nat Hum Behav.* 2021;5:857-867.
- Sarewitz D. The pressure to publish pushes down quality. *Nature.* 2016;533:147.
- Manfredini D, Greene CS, J Ahlberg, De Laat A, Lobbezoo F, Klasser GD. Evidence-based dentistry or meta-analysis illness? *A*

commentary on current publishing trends in the field of temporomandibular disorders and bruxism. *J Oral Rehabil.* 2019;46:1-4.

35. Gates M, Gates A, Duarte G, Cary M, Becker M, Prediger B, et al. Quality and risk of bias appraisals of systematic reviews are inconsistent across reviewers and centers. *J Clin Epidemiol.* 2020;125:9-15.

#### **Protocol and register**

This study protocol was registered a priori at PROSPERO (#CRD42020215676; “Quality assessment of systematic reviews and meta-analysis of periodontal intervention studies: an overview”) and no changes were deemed necessary after the start of the study.

#### **Source of Funding**

This study was supported by grants from Conselho Nacional de Desenvolvimento Científico – CNPq, Brazil (grant #302251/2019-7), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, Brazil and Pró-Reitoria de Pesquisa da UFMG (PIBIT/PRPq), Brazil. Funding agencies had no participation in the research design and data interpretation.

#### **Data Availability Statement**

Data that support the findings of this study are available from the corresponding author upon reasonable request.

#### **Authors’ contribution**

All authors have made substantial contributions to conception and design of the study. Conceptualization: Luís Cota, Carolina Martins. Methodology: Carolina Martins. Investigation: Alexandre Pereira, Carolina Martins, Julya Campos, Sandro Faria, Sarah Notaro, Tina Poklepović-Peričić, Lidiane Costa. Formal analysis: Carolina Martins, Alexandre Pereira. Writing original draft: Alexandre Pereira, Carolina Martins, Luís Cota. Writing review and editing: Carolina Martins, Luís Cota, Fernando Costa. Project administration and supervision: Carolina Martins, Luís Cota. Funding acquisition: Luís Cota, Carolina Martins.

#### **Conflict of interest**

The authors declare no conflict of interest regarding authorship, execution and publication of the present study.