Journal of Dentistry 80 (2019) 63–68

Contents lists available at ScienceDirect

Journal of Dentistry

journal homepage: www.elsevier.com/locate/jdent

Risk of bias over time in updates of Cochrane oral health reviews

Clovis Mariano Faggion Jr.[a,*], Luisiana Aranda[b], Nikolaos Pandis[c], Marco Antonio Alarcón[b], Karla Tatiana Diaz[d,e]

[a] Department of Periodontology and Operative Dentistry, Faculty of Dentistry, University of Münster, Münster, Germany
[b] Academic Department of Clinical Stomatology, Section of Periodontology and Implants, Cayetano Heredia Peruvian University, Lima, Peru
[c] Department of Orthodontics and Dentofacial Orthopedics, Dental School/Medical Faculty, University of Bern, Switzerland
[d] Academic Department of Clinical Stomatology, Section of Integral Oral Implantology, Cayetano Heredia Peruvian University, Lima, Peru
[e] School of Public Health, Faculty of Medicine, Universidad de Chile, Santiago, Chile

ARTICLE INFO

Keywords:
Bias
Cochrane tool
Systematic reviews
Methods
Methodological study
Randomized controlled trial

ABSTRACT

Objectives: To assess the changes in the risk of bias (RoB) across different versions of the same Cochrane systematic review, and to identify characteristics of systematic reviews which may be associated with different RoB scores by means of regression analysis.

Methods: We examined changes in RoB ratings in domains of randomized controlled trials (RCTs) and controlled trials (CTs) included in original Cochrane systematic reviews and their updates published in oral health. First, we checked the number of domains assessed for RoB in the different versions of the systematic review. Then, we computed the percentage of different ratings of RoB (low, high and unclear) in these systematic review versions. All data selection, extraction and analysis were conducted independently and in duplicate by two assessors. Time trends were reported in the form of line graphs. We also assessed systematic review characteristics as predictors of RoB scores by means of regression analysis.

Results: A total of 173 reviews consisting of the original reviews and their updates were examined. The proportion of different ratings of RoB was kept stable over the different versions. However, in more recent versions, the proportion of unclear RoB slightly increased, and the proportion of high RoB decreased. Cochrane risk of bias domains were a significant RoB score predictor (Likelihood ratio test p-value < 0.001).

Conclusions: Methodological improvements in RCTs and CTs included in Cochrane reviews are needed. This comprehensive information on the RoB trend may help oral health researchers improving the methodology related to specific domains.

Clinical significance: Methodological improvements are necessary for primary studies included in Cochrane reviews in oral health. The increase of domains rated as unclear RoB is of concern and suggests that strategies should be developed to improve the level of communication between trialists and systematic reviewers.

1. Introduction

Cochrane publishes systematic reviews with high methodological rigor, and they are generally with higher quality than journal-based reviews [1–3]. Authors of Cochrane systematic reviews are asked to regularly update their reviews to provide the most updated evidence by adding newer eligible trials [4]. To assess whether the quality of evidence included in the systematic reviews improved over time, it is necessary to evaluate the risk of bias (RoB) of primary studies.

RoB is related to the internal validity of the trial, and it reflects our confidence in the trial results. The Cochrane risk of bias tool includes the following domains: sequence generation, allocation concealment, blinding of participants and personnel, other potential threats to validity (other sources of bias), blinding of outcome assessment (blinding of the examiner), incomplete outcome data, and selective outcome reporting [5]. Three RoB classification levels are used: high, low and unclear (when the rater assumes there is not enough information to assign a definite rating to a specific domain). By evaluating different versions of the same systematic review, we may identify areas that improved over time and areas that require further attention. Therefore, it would be of importance to evaluate whether RoB levels have improved overtime. To our knowledge, such an evaluation has not been performed so far.

The objectives of this report were twofold: (a) to assess the changes

* Corresponding author at: Faculty of Dentistry, University of Münster, Waldeyerstraße 30, 48149, Münster, Germany.
E-mail address: Clovis.Faggion@ukmuenster.de (C.M. Faggion Jr).

https://doi.org/10.1016/j.jdent.2018.10.004
Received 17 August 2018; Received in revised form 8 October 2018; Accepted 15 October 2018
0300-5712/

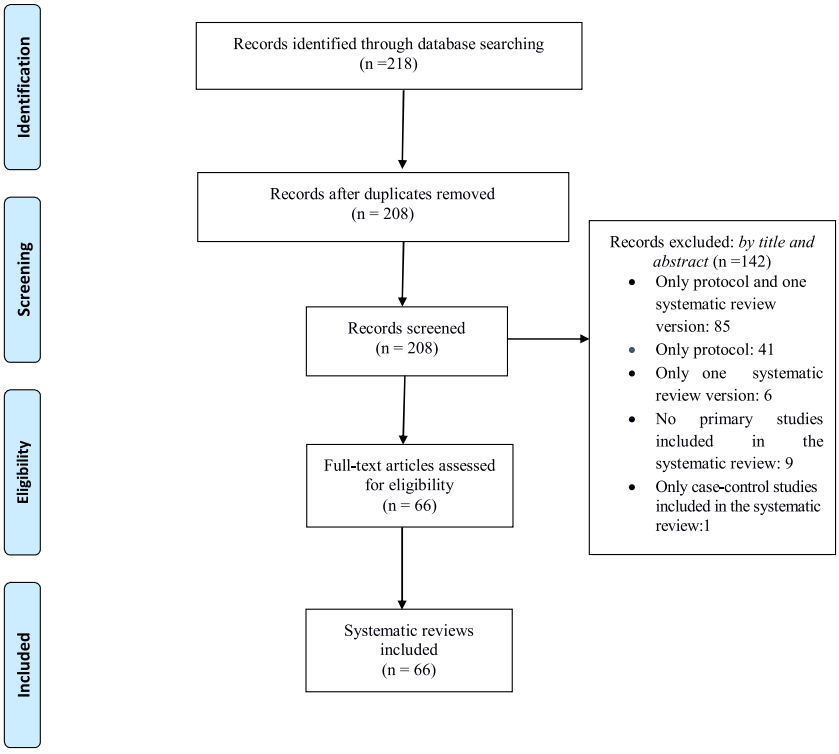

Fig. 1. PRISMA flowchart of the screening process.

in RoB in different versions of the same Cochrane systematic review published over time in oral health, and (b) to identify characteristics of systematic reviews which may be associated with different RoB scores by means of regression analysis.

2. Materials and methods

2.1. Eligibility criteria

We included oral health systematic reviews of randomized controlled trials (RCTs) or controlled trials (CTs) published by Cochrane with at least two versions (one update). Systematic reviews in a single version (not updated), empty reviews and systematic review protocols were excluded from the analysis.

2.2. Literature search and selection

Two independent reviewers (KD, LA) searched for Cochrane systematic reviews in the Cochrane database for systematic reviews (section oral heath [(http://oralhealth.cochrane.org/oral-health-evidence]) published since database inception up to January 4, 2018. The same reviewers selected and defined the final sample of systematic reviews by checking the number of versions in each topic searched. For example, in the case of a systematic review on therapies for peri-implantitis, authors selected all published versions of the review. At this stage, systematic reviews not meeting the pre-defined eligibility criteria were excluded. The search and selection procedures were performed independently. Authors discussed any disagreements in the literature search output and determined the initial sample of systematic reviews by consensus.

2.3. Data extraction

Two authors (KD, LA) extracted, independently and in duplicate, data on the characteristics of systematic reviews directly into a standardised Microsoft Excel worksheet. The two authors discussed any disagreements. In case of lack of consensus, a third author (CMF) resolved the disagreement. The following characteristics were extracted from the reviews: a) country (first author); b) number of authors; c) meta-analysis performed (yes/no); d) dental field; e) type of study (RCTs or CTs); type of RCT (parallel, split-mouth or cross-over); f) number of citations in Google Scholar (retrieved on January 4, 2018); and g) number of primary studies included in the review.

Two authors (KD, LA) evaluated independently the RoB of the trials included in the systematic reviews with two descriptive approaches: 1) by checking the number of domains assessed for RoB in the different versions of the systematic review, and 2) by checking the proportion (in percentage) of different levels of RoB (low, high and unclear) in the different versions of the systematic review. Due to the changes in the RoB over the years not all current domains were initially available. Therefore, we recorded the proportion of each Cochrane domain reported and evaluated by systematic review authors across the different versions. Finally, the two authors also recorded the total number of clinical trials included in the different versions of the same systematic review.

2.4. Reviewers's training

One author (CMF) piloted the standardized form in systematic reviews not included in the present sample. Then, two authors (KD, LA) performed several rounds of training by extracting the data from nine systematic reviews not included in the sample. Questions and ambiguities were discussed and resolved until a consensus was reached.

2.5. Data analysis

Time trends were presented in the form of line graphs. Trends of changes on the level of RoB were presented as percentages of the total RoB rating (low, high and unclear). Given the large number of included trials per version and review, a mean score was calculated per domain for each included trial across systematic review versions. Due to the hierarchical nature of the data, a series of random effects linear models

Table 1
Characteristics of systematic reviews included.

Number of versions (N = 173)	
2	40
3	14
4	9
5	3
Number of primary studies[a]	
Increased	44
Decreased	7
Same number	15
Continent of first author	
Asia	11
Europe	155
North America	3
South America	4
Number of authors	
Median (IQR)	5 (4–6)
Median version 1 (IQR)	4 (4–5)
Median version 2 (IQR)	4 (4–6)
Median version 3 (IQR)	5 (3–5)
Median version 4 (IQR)	6 (5–8)
Median version 5 (IQR)	4 (3–9)
Study design	
RCT	124
RCT + CT	7
RCT + Quasi RCT	32
Unclear	4
Not applicable	6
Type of RCT	
Parallel	96
Split-mouth	5
Parallel + split-mouth	45
Parallel + crossover	13
Parallel + split-mouth + crossover	4
Unclear	4
Not applicable	6
Meta-analysis	
Yes	81
No	86
Not applicable	6
Dental topic (N = 173)	
Halitosis	7
Craniofacial anomaly	26
Dental anxiety	3
Dental caries	26
Maintenance	3
Oral and maxillofacial surgery	9
Oral cancer	19
Oral hygiene	2
Oral medicine	13
Oral mucositis	5
Oral pain	4
Periodontal disease	17
Dental implants	33
Endodontics	6

IQR: Interquartile range.
RCT: randomized clinical trial.
CT: controlled trial.
[a] Changes in the number of primary studies included in the initial and last systematic review versions.

were fitted where the outcome was the mean score and version and domain as the predictor plus the following variables one at a time: continent of first author, number of included studies, number of authors and inclusion of meta-analysis. The variables that were significant at alpha = 0.20 were included in the final model. All analyses were conducted using Stata 14.2 (StataCorp, College Station, TX, USA).

3. Results

3.1. Systematic review selection

Two hundred eighteen potential documents were initially retrieved for assessment. After removal of duplicates, 208 documents were ready for evaluation. After selection of titles and abstracts, 141 documents were excluded. One more review was excluded after full-text assessment. Finally, 66 original systematic reviews with 107 updates for a total of 173 reviews were included and analysed (Fig. 1). The list of systematic reviews included and excluded documents (with reasons for exclusion) is reported in the supplementary material.

3.2. Data analysis results

The number of versions (original version and updates) of the same systematic review ranged from 2 to 5 (median 2, interquartile range [IQR] 2–3). The median number of studies included in the first version of the systematic reviews was 5 (IQR 1–10), 6 (IQR 2–16) in the second version, 9 (IQR 2–23) in the third version, 14 (IQR 9–30) in the fourth version, and 26 (IQR 9–131) in the fifth version. One hundred and twenty-four (62%) of the examined review versions included RCTs as primary studies. Meta-analyses were performed in 81 (47%) of the included review versions. Table 1 reports in detail the characteristics of the included systematic reviews. The proportion of RCTs included in the systematic reviews has risen from 62% in version 1 to 100% in versions 4 and 5. Parallel RCT was the most frequent type of RCT included in the different review versions (Table 2).

Taking all review versions together, the proportion of low, high and unclear RoB was kept stable over the different systematic review versions (Fig. 2). In fact, in more recent versions, the proportion of unclear RoB slightly increased, and the proportion of high RoB decreased. When individual domains were analysed, some trends were noticed. For example, the proportion of sequence generation, allocation concealment, and selective outcome reporting domains with low RoB decreased in more recent systematic review versions (Fig. 3). Regarding high RoB, the domain on blinding of participants and personnel had a lower proportion of high RoB ratings in more recent systematic review versions. In contrast, the other domains had substantially greater proportion of high RoB ratings in more recent versions (Fig. 4). Sequence generation, allocation concealment, and selective outcome reporting domains had a greater proportion of unclear RoB ratings in more recent versions (Fig. S1, supplementary material). The percentage of domains not evaluated decreased in the updated versions of the reviews. Allocation concealment and blinding of outcome assessment (blinding of the examiner) were the domains evaluated the most across the different versions (Fig. S2, supplementary material).

It was not possible obtain the ratings of RoB in 1074 (9%) of the possible reported domain entries (Fig. S3, supplementary material). In 993 (8%) entries the domains were reported in the material and method sections, but no rating was provided in the results section. In 81 (1%) entries, it was not possible to obtain the RoB scores because authors reported methodological limitations to rate RoB (e.g., blinding of heterogeneous procedures), or no reason at all.

The final multivariable model included version, RoB domain, number of included studies and geographic location. Only continent and domain were significant mean score RoB predictors whereas version and number of included studies were not significant. Asia, Europe and North America had lower mean scores compared to South America indicating lower RoB after adjusting for the other variables. Version had no effect on the mean score ($p = 0.87$). Domain was a significant RoB score predictor (Likelihood ratio test p-value < 0.001) with allocation having the highest RoB compared to each one of the other domains (Table 3).

4. Discussion

This study demonstrated that the proportion of unclear RoB increased, and the proportion of high RoB decreased in more recent updates of Cochrane reviews published in oral health. Furthermore, the regression analysis demonstrated that RoB domain was a significant

Table 2
Study characteristics (type of studies and type of randomized controlled trial [RCT]) included in the different versions of the systematic reviews. Values in parentheses are percentages.

Type of study	Version 1 n = 66	Version 2 n = 66	Version 3 n = 26	Version 4 n = 12	Version 5 n = 3
RCT	41 (62)	47 (71)	21 (81)	12 (100)	3 (100)
RCT + CT	4 (6)	2 (3)	1 (4)	0	0
RCT + Quasi RCT	13 (20)	15 (23)	4 (15)	0	0
Unclear	2 (3)	2 (3)	0	0	0
Not applicable	6 (9)	0	0	0	0
Type of RCT					
RCT-parallel	38 (58)	39 (59)	14 (54)	5 (42)	0
RCT-split-mouth	2 (3)	2 (3)	1 (4)	5 (42)	0
RCT-parallel + split-mouth	13 (20)	18 (27)	7 (27)	0	2 (67)
RCT-parallel + crossover	4 (6)	4 (6)	3 (11)	1 (8)	1 (33)
RCT-parallel + split-mouth + crossover	1 (1)	1 (2)	1 (4)	1 (8)	0
Unclear	2 (3)	2 (3)	0	0	0
Not applicable	6 (9)	0	0	0	0

CT: controlled trial.

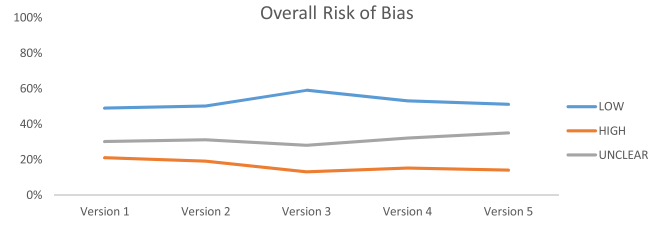

Fig. 2. Overall risk of bias (RoB) of primary studies included in the different versions of the Cochrane systematic reviews. Values are presented as percentages of RoB.

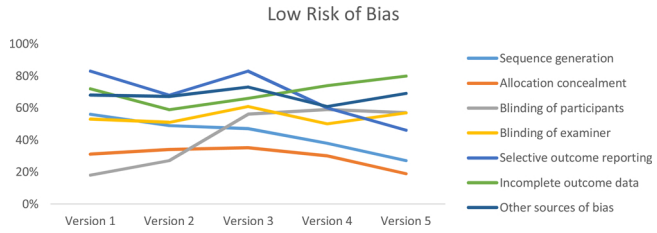

Fig. 3. Low risk of bias (RoB) of primary studies included in the different versions of the Cochrane systematic reviews. Values are presented as percentages of RoB for each individual domain.

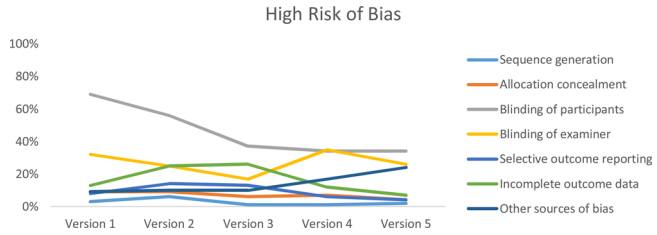

Fig. 4. High risk of bias (RoB) of primary studies included in the different versions of the Cochrane systematic reviews. Values are presented as percentages of RoB for each individual domain.

RoB score predictor suggesting that the score differs significantly between domains after adjusting for continent, number of included studies and review version.

Our findings suggest that RCTs and CTs included in Cochrane systematic reviews from the oral health field have room for methodological improvement. Only 50% of overall domains were rated as low RoB. More disappointing were the long-term results: ratings on low RoB were kept stable over the different versions of the same systematic review. Overall, the proportion of unclear RoB slightly increased possibly due to the slight decrease of domains with high RoB (Fig. 2). These findings suggest three potential explanations: 1) the quality of reporting of trials did not improve over the years; 2) the strategy of contacting authors to obtain more detailed information about their trials has not being efficient. Hence, these results may support further development of the Cochrane strategy for contacting authors [5]; and 3) improved RoB assessment of trials by systematic review authors due to increased awareness and education. It is important to report that the number of primary studies included in more recent versions increased considerably. Forty-four reviews (67% of the sample) had a median number of 6 (IQR = 3–13) more primary studies included in the last versions when compared to the first versions. Hence, the lack of improvement of RoB of trials over the years could not be explained by the inclusion of the same trials in the different versions.

Some domains presented different trends over the different versions. Interestingly, sequence generation considerably received more unclear RoB ratings in more recent versions. These are surprising results due to the easiness of application of unbiased sequence generations. This problem might also be partially explained by the lack of information on details about the strategy used to generate the sequence at random [6]. These findings suggest both that the quality of reporting at trial level, and the communication of systematic review authors with RCT researchers should be improved to reduce the number of unclear ratings.

The greater proportion of high RoB in the “other sources of bias” domain in more recent versions was also an interesting finding (Fig. 4). The trend of a greater proportion of high RoB is apparent from the third to more recent systematic review versions. One potential explanation is that authors likely became more aware in the recent years for the need to identify potential biases not included in the specific domains. This domain (other sources of bias) is particularly difficult to evaluate because there is not strict and focused guidance on how to rate this type of bias, and the Cochrane suggests several factors that can generate bias in this domain [5]. One could hypothesize that many authors would not have the adequate methodological background to evaluate and rate such a domain. Hence, systematic review authors may fail to identify these factors and tend to report low RoB. Nevertheless, a greater identification of “other sources of bias” in more recent versions may be a consequence of improved knowledge on methodological evaluation by Cochrane systematic reviewers.

Blinding of participants and personnel had a clear trend of a greater proportion of low RoB toward more recent versions. This domain is composed by two parts: patients and personnel (i.e., operators and staff members directly related to the procedures) [5]. In many occasions it is challenging to have a low RoB in this domain due to technical difficulties [7]. For example, in an RCT comparing a surgical against a non-

Table 3
Results from regression analysis for the effect on the mean risk of bias (RoB) score of continent, number of included studies, number of versions and RoB domain.

Predictor		β-coefficient	95% Confidence Intervals	p-value
Continent	South America	reference		
	Asia	−.65	−1.13, − .160	0.01
	Europe	−.49	−.93, −.05	0.03
	North America	−.54	−1.06 −.02	0.04
Number of included studies (per unit)		0.002	−.001, .005	0.17
Version (per unit)		−0.004	−.06 .05	0.87
Domain	Allocation concealment	reference		
	Blinding of the examiner	−.45	−.57 −.38	< 0.001
	Blinding of participants	−.41	−.57 −.25	< 0.001
	Incomplete outcome data	−.49	−.62 −.36	< 0.001
	Other sources of bias	−.45	−.59 −.31	< 0.001
	Selective outcome reporting	−.59	−.74 −.45	< 0.001
	Sequence generation	−.27	−.40 −.14	< 0.001

surgical procedure there is an obvious limitation on making the type of therapies blinded to patients and professionals performing the procedures. Hence, in such cases a high RoB assessment would be likely. A potential explanation for the greater proportion of low RoB in more recent versions was the inclusion of interventions in which the application of blinding was feasible. It is important to report the lack of separation of performance and detection bias in older versions of the handbook [8], which included blinding of participants, personnel and outcome assessors in a single domain.

Selective outcome reporting had considerably lower proportion of low RoB in more recent versions. This was mainly due to the greater proportion of domains judged at unclear RoB. It can be postulated that this could be the result of increased awareness of this type of bias and the increasing numbers of pre-registered trials [9]. Cochrane suggests the detailed analysis of the material and methods reported in the published article to identify potential selective outcome reporting [5]. Not detailed registered research protocols or lack of detailed information from RCT authors may have generated the great proportion of unclear RoB.

Incomplete outcome data had a slight but steady increasing proportion of low RoB since version 2. These results are difficult to explain, because this trend should potentially go to high RoB. The rationale is that the longer the duration of an RCT, the higher will be the probability of loss to follow-up [10]. Substantial loss of patients and consequent missing data would likely generate concerns for attrition bias [5], and reviewers would tend to assign a high RoB rating. One potential explanation for low RoB scores in this sample would be the inclusion of new RCTs with short follow-ups with a lower probability of patient withdrawal and missing data.

The Cochrane approach for evaluating RoB changed over the years. The format based on seven areas or domains was published from 2008 [11]. In previous versions [12], some domains were not present such as the “other sources of bias” domain, and there was also not well-stablished evaluation on judging the RoB in the domains. The trend on the evaluation of domain can be observed in Fig. S2 (Supplementary material) where the “other sources of bias” domain has almost 80% of entries not evaluated in the first version of the reviews included. Nevertheless, for all domains we can notice that the percentage of entries not evaluated by reviewers dropped in more recent versions, mainly versions four and five.

A great proportion of domains rated at unclear RoB is still present in more recent versions, mainly the sequence generation, allocation concealment, and selective outcome reporting domains. These results suggest that RCT authors should adhere more carefully to the Consolidated Standards of Reporting Trials (CONSORT) guidelines for reporting randomized studies [13]. Recent initiatives have improved reporting dramatically in a dental specialty journal [14,15]. Furthermore, strategies for contacting authors of the trials should be tailored to clarify issues, mainly those related to the three domains reported above.

The Cochrane tool was recently updated (named RoB 2.0 tool) with important changes, mainly in the organization of the domains and in the way RoB assessment is carried out [16], however this tool is not yet included in the current Cochrane handbook [5]. In the background information and detailed guidance for using the RoB 2.0 tool [17], the authors of the revised tool explain that they avoided the use of type of biases names published in the previous versions to make the RoB evaluation clearer. The objective of RoB 2.0 is to cover all potential factors that may lead to RoB [18] and avoid the limitations of the older versions such as omitting a RoB domain because it is not applicable to that specific trial condition [18]. It would be of interest in the future to compare the RoB 2.0 tool with the previous version to better understand the psychometric properties (e.g, construct validity and inter-rater reliability [19]) of the revised tool.

Our study has limitations and strengths. It was only possible to evaluate trends on reporting RoB in Cochrane systematic reviews, because these reviews use a standard approach to evaluate RoB in the included trials. Therefore, the analysis does not involve systematic reviews published in other dental or medical journals. However, the sample of systematic reviews is robust and representative for the oral health field. Moreover, the comparisons were performed by the same authors who developed the different versions of the same systematic review. Therefore, there was likely not much variability in rating RoB, which is a sensitive aspect of analysing RoB with a domain-based approach such as this used by Cochrane [20]. The present work reports a comprehensive scenario about RoB in oral health that can guide further development of research methodology in this field. Our findings are based on what was reported in the SRs and any incorrect assessment in the reviews was transferred in our results. The inter-rater agreement can, nevertheless, vary across domains of the RoB tool, mainly when more subjective judgement is necessary [21]. Furthermore, Cochrane is in general against assigning scores to bias, however, we needed to summarize the ratings across versions as a means to describe trends.

Regression analysis demonstrated that only domain was a significant RoB score predictor. The number of studies was not a significant predictor, but the inclusion of more studies could have been potentially viewed as a reason for worse RoB scores. It is of interest to note that by fitting regression models separately for each domain, there was evidence that for the allocation concealment domain the number of included studies was a significant predictor ($p = 0.02$).

In conclusion this article reported trends in reporting RoB in different versions of Cochrane systematic reviews published in oral health. The findings provide important information on the RoB of individual domains over the years. These findings can guide RCT authors to focus

on the respective domains with high and unclear RoB to improve the methodology and reporting of their trials.

Acknowledgement

Authors of this study have no conflict of interest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jdent.2018.10.004.

References

[1] M. Petticrew, P. Wilson, K. Wright, F. Song, Quality of Cochrane reviews. Quality of Cochrane reviews is better than that of non-Cochrane reviews, BMJ 324 (2002) 545.
[2] J. Wen, Y. Ren, L. Wang, Y. Li, Y. Liu, M. Zhou, P. Liu, L. Ye, Y. Li, W. Tian, The reporting quality of meta-analyses improves: a random sampling study, J. Clin. Epidemiol. 61 (2008) 770–775.
[3] C.M. Faggion Jr., M.A. Atieh, S. Park, Search strategies in systematic reviews in periodontology and implant dentistry, J. Clin. Periodontol. 40 (2013) 883–888.
[4] P. Garner, S. Hopewell, J. Chandler, H. MacLehose, H.J. Schünemann, E.A. Akl, J. Beyene, S. Chang, R. Churchill, K. Dearness, G. Guyatt, C. Lefebvre, B. Liles, R. Marshall, L. Martínez García, C. Mavergames, M. Nasser, A. Qaseem, M. Sampson, K. Soares-Weiser, Y. Takwoingi, L. Thabane, M. Trivella, P. Tugwell, E. Welsh, E.C. Wilson, H.J. Schünemann, Panel for updating guidance for systematic reviews (PUGs). When and how to update systematic reviews: consensus and checklist, BMJ 354 (2016) i3507.
[5] J.P.T. Higgins, D.G. Altman, J.A.C. Sterne, Assessing risk of bias in included studies, in: J.P.T. Higgins, R. Churchill, J. Chandler, M.S. Cumpston (Eds.), Cochrane Handbook for Systematic Reviews of Interventions Version 5.2.0 (Updated June 2017), Cochrane, 2017 Available from www.training.cochrane.org/handbook.
[6] V.W. Berger, Sequence generation and allocation concealment, J. Evid. Dent. Pract. 16 (2016) 145–146.
[7] C.M. Faggion Jr., A. Monje, J. Wasiak, Appraisal of systematic reviews on the management of peri-implant diseases with two methodological tools, J. Clin. Periodontol. 25 (March) (2018), https://doi.org/10.1111/jcpe.12893.
[8] J.P.T. Higgins, D.G. Altman, Assessing risk of bias in included studies, Chapter 8, in: J.P.T. Higgins, S. Green (Eds.), Cochrane Handbook for Systematic Reviews of Interventions, John Wiley & Sons, Chichester (UK), 2008.
[9] C. De Angelis, J.M. Drazen, F.A. Frizelle, C. Haug, J. Hoey, R. Horton, S. Kotzin, C. Laine, A. Marusic, A.J. Overbeke, T.V. Schroeder, H.C. Sox, M.B. Van Der Weyden, International committee of medical journal editors, Clinical trial registration: a statement from the international committee of medical journal editors, Lancet 364 (2004) 911–912.
[10] K. Gustavson, T. von Soest, E. Karevold, E. Røysamb, Attrition and generalizability in longitudinal studies: findings from a 15-year population-based study and a Monte Carlo simulation study, BMC Public Health 12 (2012) 918.
[11] J.P. Higgins, D.G. Altman, P.C. Gøtzsche, P. Jüni, D. Moher, A.D. Oxman, J. Savovic, K.F. Schulz, L. Weeks, J.A. Sterne, Cochrane Bias Methods Group, Cochrane Statistical Methods Group, The Cochrane Collaboration's tool for assessing risk of bias in randomised trials, BMJ. 343 (2011) d5928.
[12] Cochrane handbook for systematic reviews of interventions 4.2.6, [updated September 2006], in: J.P.T. Higgins, S. Green (Eds.), The Cochrane Library, 4 John Wiley & Sons, Ltd., Chichester, UK, 2006.
[13] K.F. Schulz, D.G. Altman, D. Moher, CONSORT Group, CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials, J. Clin. Epidemiol. 63 (August (8)) (2010) 834–840.
[14] N. Pandis, L. Shamseer, V.G. Kokich, P.S. Fleming, D. Moher, Active implementation strategy of CONSORT adherence by a dental specialty journal improved randomized clinical trial reporting, J. Clin. Epidemiol. 67 (2014) 1044–1048.
[15] D. Koletsi, P.S. Fleming, R.G. Behrents, C.D. Lynch, N. Pandis, The use of tailored subheadings was successful in enhancing compliance with CONSORT in a dental journal, J. Dent. 67 (2017) 66–71.
[16] J.P.T. Higgins, J.A.C. Sterne, J. Savović, M.J. Page, A. Hróbjartsson, I. Boutron, B. Reeves, S. Eldridge, A revised tool for assessing risk of bias in randomized trials, in: J. Chandler, J. McKenzie, I. Boutron, V. Welch (Eds.), Cochrane Methods. Cochrane Database of Systematic Reviews (Issue 10 (Suppl 1)) (2016), https://doi.org/10.1002/14651858.CD201601.
[17] RoB 2.0 tool. https://sites.google.com/site/riskofbiastool//welcome/rob-2-0-tool. Accessed 10 May 2018.
[18] M.J. Page, I. Boutron, C. Hansen, D.G. Altman, A. Hróbjartsson, Assessing risk of bias in studies that evaluate health care interventions: recommendations in the misinformation age, J. Clin. Epidemiol. 97 (2018) 133–136.
[19] M.J. Page, J.E. McKenzie, J.P.T. Higgins, Tools for assessing risk of reporting biases in studies and syntheses of studies: a systematic review, BMJ Open 8 (2018) e019703.
[20] L. Hartling, M.P. Hamm, A. Milne, B. Vandermeer, P.L. Santaguida, M. Ansari, A. Tsertsvadze, S. Hempel, P. Shekelle, D.M. Dryden, Testing the risk of bias tool showed low reliability between individual reviewers and across consensus assessments of reviewer pairs, J. Clin. Epidemiol. 66 (2013) 973–981.
[21] L. Hartling, M. Ospina, Y. Liang, D.M. Dryden, N. Hooton, J. Krebs Seida, T.P. Klassen, Risk of bias versus quality assessment of randomised controlled trials: cross sectional study, BMJ 339 (2009) b4012.