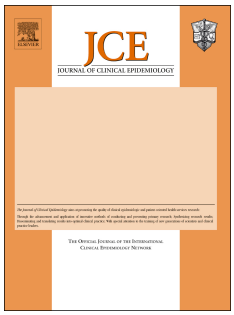


Methodological and reporting quality in non-Cochrane systematic review updates could be improved: a comparative study

Ya Gao, Yitong Cai, Kelu Yang, Ming Liu, Shuzhen Shi, Ji Chen, Yue Sun, Fujian Song, Junhua Zhang, Jinhui Tian

PII: S0895-4356(19)30788-7

DOI: https://doi.org/10.1016/j.jclinepi.2019.11.012

Reference: JCE 10018

To appear in: *Journal of Clinical Epidemiology*

Received Date: 24 August 2019

Revised Date: 12 November 2019

Accepted Date: 18 November 2019

Please cite this article as: Gao Y, Cai Y, Yang K, Liu M, Shi S, Chen J, Sun Y, Song F, Zhang J, Tian J, Methodological and reporting quality in non-Cochrane systematic review updates could be improved: a comparative study, *Journal of Clinical Epidemiology* (2019), doi: https://doi.org/10.1016/j.jclinepi.2019.11.012.

Methodological and reporting quality in non-Cochrane systematic review updates could be improved: a comparative study

Ya Gao[1†], Yitong Cai[2†], Kelu Yang[2], Ming Liu[1], Shuzhen Shi[1], Ji Chen[2], Yue Sun[2], Fujian Song[3], Junhua Zhang[4]*, Jinhui Tian[1,2,5]*

[1]Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

[3]Evidence-Based Nursing Center, School of Nursing, Lanzhou University, Lanzhou, China

[3]Public Health and Health Services Research, Norwich Medical School, University of East Anglia, Norwich, UK

[4] Evidence-Based Medicine Center, Tianjin University of Traditional Chinese Medicine, Tianjin, China

[5]Key Laboratory of Evidence-Based Medicine and Knowledge Translation of Gansu Province, Lanzhou, China

[†]Ya Gao and Yitong Cai contributed equally to this work.

*** Correspondence:** Jinhui Tian. Evidence-Based Medicine Centre, School of Basic Medical Sciences, Lanzhou University; Evidence-Based Nursing Center, School of Nursing, Lanzhou University, No.199, Donggang West Road, Lanzhou City, 730000,

Gansu Province, China. (Email: tianjh@lzu.edu.cn)

Junhua Zhang. Evidence-Based Medicine Center, Tianjin University of Traditional Chinese Medicine, No. 312 Anshanxi Street, Nankai District, Tianjin 300193, China. (Email: zjhtcm@foxmail.com)

Abstract

Objectives: To compare the methodological and reporting quality of updated systematic reviews (SRs) and original SRs.

Study Design and Setting: We included 30 pairs of non-Cochrane updated and original SRs, identified from a search of PubMed and Embase.com. We used AMSTAR 2 to assess methodological quality and PRISMA for reporting quality. Stratified analyses were conducted to compare the differences between updated SRs and original SRs and explore factors that might affect the degree of quality change.

Results: Of the 60 non-Cochrane SRs, only 2 (3.3%) were of low quality, the remaining 58 (96.7%) were of critical low quality. There were no statistically significant differences in methodological quality between the updated SRs and original SRs, although the compliance rates of 8 items of updated SRs were higher than that of original SRs. Updated SRs showed an improvement on 15 PRISMA items, but no items with statistically significant differences. The differences in fully reported AMSTAR 2 and PRISMA items between original SRs and updated SRs were also not statistically significant after adjusting for multiple review characteristics.

Conclusions: The methodological and reporting quality of updated SRs were not improved compared with original SRs, although the quality could be further improved for both updated SRs and original SRs.

Keywords: Systematic reviews; Updating; AMSTAR 2; PRISMA; Methodological quality; Reporting quality

Running title：Methodological and reporting quality of original and updated SRs

What is new?

Key findings

- The methodological and reporting quality of non-Cochrane updated SRs were not improved compared with the original SRs.

What this adds to what was known?

- This study investigated the methodological and reporting quality of the non-Cochrane updated SRs and original SRs in both overall fully reported items and individual item of AMSTAR 2 and PRISMA checklists and assessed certain factors that may affect the extent of the methodological and reporting quality changes during the update process.

What is the implication and what should change now?

- The identified deficiencies should be paid more attention, especially for the updated SRs. Researchers, journal editors, and peer reviewers should ensure that the methodological and reporting guidelines are strictly followed before publication. Further research should focus on developing a methodological or a reporting quality tool that is specifically applicable to updated SRs.

1. Introduction

Systematic reviews (SRs) are fundamental tools for generating reliable medical information [1, 2], which provides a comprehensive synthesis of a large amount of evidence to help clinicians keep up with the pace of medical literature, explains the differences between studies on the same issue, formulates clinical policies, combines best evidence with clinical practice, and suggests directions for new research [3-6]. Recent estimates suggest that more than 8,000 new SRs were published in Medline annually, which is equivalent to a threefold increase over the past decade [7, 8]. However, SRs are most useful when they are kept up to date [9-11]. Therefore, as SRs continue to increase, more and more SRs are updated to include new evidence, which can relatively reduce publication bias and increase the credibility of the results of SRs [10, 12, 13]. However, empirical studies have shown that there are a large number of unnecessary, misleading, and conflicting SRs and meta-analyses, in part because of inappropriate methodological design, conduct, or reporting [14, 15]. SRs of incomplete reports or flawed implementation methods can lead to biased recommendations and may distort decisions [16, 17], limiting the role of SR in decision-makers [18].

The methodological quality and reporting quality are considered as the two main aspects of the quality of an SR, which aims to assess, and hopefully improve, the design, conduct, and reporting of SRs [19]. For example, the Assessment of Multiple Systematic Reviews 2 (AMSTAR 2) tool is aimed to improve the methodological quality of SRs as well as the Preferred Reporting Items of Systematic reviews and

Meta-Analyses (PRISMA) statement focused on the advance of reporting quality [20, 21]. Previous studies have evaluated the quality of SRs in different fields using AMSTAR and PRISMA tools and they all showed that SRs have some weaknesses and the quality needs to be further improved [2, 8, 15, 22-24]. Besides, the study also has compared the quality of updated and original Cochrane SRs [25], which to some extent were considered to have better methodological rigor and more frequent updates than peer-reviewed paper journals [23, 26]. But there was still a lack of empirical evidence to evaluate the impact of updating on the quality of non-Cochrane reviews, and whether certain factors affect the improvement of quality during the update process has not been evaluated nor studied.

This study is a sister paper of our project, another paper will explore whether the updated SRs exhibit outcomes change and whether the updated SRs improve the precision of outcomes. The primary objective of the present study was to assess methodological and reporting quality of included SRs. The secondary objective was to compare the differences between the updated SRs and original SRs in the quality. The third objective was to determine whether certain characteristics (e.g., updated team, intervention, included trials) affect the degree of quality change during the update process.

2. Methods

2.1. Criteria for considering reviews for inclusion

Systematic reviews of interventions with or without meta-analysis that met the following criteria were included: (1) was a review article and explicitly described methods of study selection, and explicitly reported the methods of evidence synthesis [22]; (2) original SRs and updated SRs of randomized controlled trials (RCTs) or quasi-randomized controlled trials that evaluated clinical effects of health care interventions; and (3) all the original SRs and updated SRs were published in English language.

Studies including the following were excluded: (1) SRs that included RCTs and nonrandomized studies or only included nonrandomized studies; (2) SRs did not focus on health care interventions such as etiology, diagnosis, and prognosis; (3) SRs did not clearly state "update" in the titles or articles; (4) the original SR or updated SR is a Cochrane review; (5) the second analysis of previous SR; (6) overviews of SRs, methodological reviews, umbrella overviews, scoping or rapid reviews, review protocols, abstracts, conference proceedings, and letters to editors.

2.2. Electronic searches

We searched the PubMed and Embase.com using the phrases “systematic review”, “meta-analysis”, “indirect comparison”, “indirect treatment”, “mixed treatment comparison”, “multiple treatment comparison”, and “update” to identify updated SRs, across all years up to March 3, 2019. We did not apply any date

restriction. The search strategy of Embase is presented in Appendix Word 1.

2.3. Selection of reviews

We imported the retrieved records into EndNote X8 (Thomson Reuters (Scientific) LLC Philadelphia, PA, US) for management. Two overview authors (Y.G. and Y.T.C.) independently screened the results of the electronic search by title and abstract. The full-text versions of the possibly relevant reviews were obtained for further assessment to determine the final inclusion according to the eligibility criteria. If a system review had multiple updated versions, we would include the latest one. We resolved disagreements through discussions with a third reviewer (J.H.Z, or J.H.T.).

2.4. Data collection

One reviewer (Y.G., Y.T.C., K.L.Y., or M.L.) extracted data from the included SRs using a pre-defined data extraction form and a second reviewer checked the extracted data for accuracy and completeness. We resolved any discrepancies by consensus. A third reviewer (J.H.Z, or J.H.T.) was requested for discussion if the agreement could not be reached. The data extraction form included the following details: first author, journal name, publication year, country of authors, whether SRs involved co-first author, whether SRs had co-correspondence author, whether statistician, epidemiologist, or methodologist (based on the author's current academic

position) was involved, whether SRs had a priori protocol, the number of RCTs included, funding source (nonprofit, for-profit, unfunded, or not reported), topic of interest, methodological quality, and reporting quality.

2.5. Assessment of methodological and reporting quality

The methodological quality of the included SRs was evaluated according to the Assessment of Multiple Systematic Reviews 2 (AMSTAR 2) tool, which is a revised version of the original AMSTAR instrument, allowing for the evaluation of SRs based on random and non-randomized studies [27, 28]. The AMSTAR 2 contains 16 items, among which seven are critical domains. The overall confidence of the results of the review was rated into four levels: high, moderate, low, and critically low [27]. Each item was judged as “Yes” (item fully addressed), “No” (item not addressed), or “Partial Yes”(item not fully addressed).

The reporting quality was assessed using the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement, which a checklist of 27 items aimed to improve the completeness and transparency of reporting of SRs [8, 29]. We responded each item to “Yes’’ (total compliance), “Partial’’ (partial compliance), ‘‘No’’ (noncompliance), or ‘‘Can not answer’’ (limited information) [22, 30]. The quality of SRs was assessed by one reviewer (Y.G., Y.T.C., K.L.Y., S.Z.S, J.C., or Y.S.) and verified by another. Disagreements were resolved by consensus or third-party adjudication if consensus cannot be reached (J.H.Z, or J.H.T.).

2.6. Data management and synthesis

We compared the general characteristics between updated SRs and original SRs. Frequency and percentage were used for categorical variables, and median and interquartile range were used for continuous variables. Chi-squared test or Fisher exact test (if a contingency table contained a cell with five or fewer events) was used to evaluate the differences in categorical data and two-sample Wilcoxon rank-sum test or Student t-test was used to assess the differences in continuous data [8, 22]. The analyses were conducted in IBM SPSS Statistics v. 24.0 (Armonk, NY: IBM Corp.), and the statistical level of significance was set at $P < 0.05$.

The compliance rate of each item for the methodological and reporting quality was calculated with the number acquired “Yes” and the total number of included original SRs or updated SRs. Then, we performed a Chi-squared test and calculated the odds ratio (OR) with 95% confidence intervals (95% CIs) and *P* value to compare the compliance of each item between updated SRs and original SRs. We created bubble plots with Microsoft Excel 2016 (Microsoft Corp, Redmond, WA, www.microsoft.com) to present the compliance rate and OR. Considering bubble plots according to the compliance rate, the bubble size represented the number of compliance rate, the X-axis represented the AMSTAR 2 or PRISMA items, the Y-axis represented the compliance rate of each item. As for bubble plots incorporating OR, the bubble size displayed the number of OR, the X-axis represented the AMSTAR 2

or PRISMA items, the Y-axis indicated the OR of updated SRs compared to original SRs for each item.

We also computed the number of acquired “Yes” of AMSTAR 2 items and PRISMA items for each SR. The mean difference (MD) with 95%CI was calculated to compare the mean fully reported items between the original SRs and updated SRs. Either bivariate or multiple variable linear regression analysis was used to explore the potential factors that affected the fully reported AMSTAR 2 items and PRISMA items of original SRs and updated SRs.

Furthermore, we divided the included SRs into updates by the same team or by different teams. In the current study, we defined SR to be updated by the same team if the first author or corresponding author did not change by comparing the updated SR with the original SR. We categorized SRs with the same interventions or changed interventions by comparing the involved interventions of updated SR with the original SR. We also identified the updated SRs included the trials that included in the original SRs, or did not include the trials included in the original SRs, categorized the updated SRs fully included trials of original SRs or partially included trials of original SRs, and classified updated and original SRs with a priori protocol or without a priori protocol. We then conducted the Chi-squared test and calculated OR with 95% CIs to compare the compliance between updated SRs and original SRs for AMSTAR 2 and PRISMA items considering each subgroup, separately. We also calculated the ratio of odds ratio (ROR) and *P* value to compare the extent of the change in compliance of each item between SRs to be updated by the same team and different teams, between

SRs with the same interventions and changed interventions, between SRs included the trials and did not include the trials included in the original SRs, between SRs fully included trials of original SRs and partially included trials of original SRs, and between SRs with a protocol and without a protocol.

3. Results

3.1. Screening Results

The initial search returned 4997 review records, after titles, abstracts, and full-text screening according to the eligible criteria, we included 30 non-Cochrane updated SRs and 30 non-Cochrane original SRs. Details of the search screening process are presented in Figure 1. The full lists of included SRs can be found in Appendix Word 2.

3.2. General characteristics of included SRs

The main characteristics of the included systematic reviews are summarized in Table 1. The original SRs were published in 28 different journals and had a median journal impact factor of 5.206 (IQR: 2.857 to 17.870). The updated SRs were published in 27 different journals and had a median journal impact factor of 3.750 (IQR: 2.520 to 6.046). The included SRs were published between 1994 and 2018.

Compared with original SRs, the updated SRs were more likely to be published in journals with lower impact factors and tended to be published between 2011 and 2018. Only three SRs with the co-first author, one SR involved co-correspondence author, ten SRs had statistical, epidemiological, or methodological authors, and there were no statistically significant differences between the updated SRs and the original SRs. Eight updated SRs and seven original SRs had a priori protocol. 43.3% of the SRs were published by authors from two or more countries, more than half of SRs were conducted in Europe, and 25.0% were completed in North America. Only 30.0% of the SRs were funded and most of them were funded by the nonprofit sponsor, but there were still 61.7% of SRs did not report funding sources. The included SRs investigated several interventions in different clinical conditions. Neoplasms (10, 16.7%), followed by Diseases of the circulatory system (8, 13.3%), were the most studied ones, and there were no statistically significant differences in disease conditions investigated.

Table 1 The main characteristics of the included SRs

Characteristics	Original SRs	Updated SRs	*P* value
Journal impact factor: median (IQR)	5.206(2.857,17.870)	3.750(2.520,6.046)	0.011
Year of Publication			
1994 to 2000	3(10.0)	1(3.3)	0.612
2001 to 2005	8(26.7)	3(10.0)	0.095
2006 to 2010	9(30.0)	7(23.3)	0.559
2011 to 2015	10(33.3)	9(30.0)	0.781
2016 to 2018	0(0.0)	10(33.3)	0.001
With co-first author (%)	1(3.3)	2(6.7)	1.000
With co-correspondence author (%)	0(0.0)	1(3.3)	1.000
With statistician, epidemiologist, or methodologist (%)	7(23.3)	3(10.0)	0.166
Authors from 2 or more countries (%)	14(46.7)	12(40.0)	0.602
Origin region (%)			

Asia	4(13.3)	4(13.3)	1.000
Europe	16(53.3)	17(56.7)	0.795
North America	7(23.3)	8(26.7)	0.766
South America	1(3.3)	0(0.0)	1.000
Oceania	2(6.7)	1(3.3)	1.000
Number of RCTs included: median (IQR)	13.5(7.75, 29.25)	18.5(9.75, 30)	0.010
Funding sources (%)			
Nonprofit sponsor	9(30.0)	7(23.3)	0.559
For-profit sponsor	1(3.3)	1(3.3)	1.000
None	2(6.7)	3(10.0)	1.000
Not reported	18(60.0)	19(63.3)	0.791
Conditions investigated[a] (%)			
Neoplasms	5(16.7)	5(16.7)	1.000
Diseases of the circulatory system	4(13.3)	4(13.3)	1.000
Diseases of the digestive system	3(10.0)	3(10.0)	1.000
Diseases of the musculoskeletal system and connective tissue	3(10.0)	3(10.0)	1.000
Mental and behavior disorders	2(6.7)	2(6.7)	1.000
Pregnancy, childbirth, and puerperium	1(3.3)	1(3.3)	1.000
Diseases of the genitourinary system	1(3.3)	1(3.3)	1.000
Diseases of the infectious disease	1(3.3)	1(3.3)	1.000
Diseases of the nervous system	1(3.3)	1(3.3)	1.000
Diseases of the respiratory system	1(3.3)	1(3.3)	1.000
Endocrine, nutritional, and metabolic disease	1(3.3)	1(3.3)	1.000
Factors influencing health status and contact with health services	1(3.3)	1(3.3)	1.000
Any other	6(20.0)	6(20.0)	1.000

IQR, interquartile range.

[a] Conditions were categorized according to the International Classification of Diseases 11th Revision (ICD-11).

3.3. Results of methodological quality

3.3.1. Methodological quality of updated SRs and original SRs

Only two (3.3%) SRs were of low quality, fifty-eight (96.7%) were of critical low quality, and none of the SR was of high or moderate quality (Appendix Table 1). The difference in the mean fully reported AMSTAR 2 items between updated SRs and

original SRs was not statistically significant (MD = 0.03, 95%CI: -1.36 to 1.42, *P* = 0.966). Considering individual item, the compliance rates of three items were higher than 60.0% for both the updated SRs and original SRs, they were “research questions and inclusion criteria include the components of PICO”, “described the included studies in adequate detail”, and “used appropriate methods for statistical combination of results”. However, all the SRs did not explain their selection of the study designs for inclusion in the review, all the original SRs and 96.7% of the updated SRs did not report the sources of funding for studies included in the review, no more than 33.0% of the SRs provided explicit statement that the review methods were established prior to the conduct of the review and clarified the significant deviations from the protocol, provided a list of excluded studies and justified the exclusions, assessed the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence syntheses, carried out an adequate investigation of publication bias and discussed its likely impact on the results of the review, and reported potential sources of conflict of interest (Figure 2, Appendix Table 2). There were no statistically significant differences between the updated SRs and the original SRs in the AMSTAR 2 items, although the compliance rates of 8 items of updated SRs were higher than that of original SRs (Figure 3, Appendix Table 2).

3.3.2. Changes in methodological quality of SRs in different groups

Overall, no significant difference was found between SRs with the same team

and SRs with different teams (P>0.05), although the SRs updated by different teams have a greater degree of quality improvement on 8 items. Compared with SRs with different interventions, the SRs with the same intervention had a greater degree of quality improvement according to items 2, 8, 12, 14, and 15. However, the differences between them were not statistically significant (P>0.05). There were no significant differences in the degree of quality improvement between SRs included the trials that included in the original SRs and SRs did not include the trials included in the original SRs, between SRs fully included trials of original SRs and SRs partially included trials of original SRs, and between SRs with a protocol and SRs without a protocol (Appendix Table 3).

3.4. Results of reporting quality

3.4.1. Reporting quality of updated SRs and original SRs

Among the 27 PRISMA items, 6 items obtained a compliance rate higher than 70.0% for both the updated SRs and original SRs. However, only 26.7% of the updated SRs and 23.3% original SRs fully reported the protocol and registration, 40.0% of the SRs presented the data collection process, and less than 32.0% SRs assessed the risk of bias across studies in the Methods section. More than 78.0% of the SRs did not report the risk of bias across studies in the Results section, and most SRs did not clarify the limitations and funding (Figure 4, Appendix Table 4). Compared with the

original SRs, the reporting rates of the updated SRs had an improvement on 15 items, but no items were observed to have statistically significant differences (Figure 5, Appendix Table 4). As for the mean fully reported PRISMA items, no statistically significant difference was found between updated SRs and original SRs (MD = 0.10, 95%CI: -2.37 to 2.57, *P* = 0.937).

3.4.2. Changes in reporting quality of SRs in different groups

For SRs with different teams, the updated SRs had a low compliance rate in reporting of information sources (OR=0.08, 95%CI: 0.01 to 0.76, *P*=0.014), but there were no statistically significant differences in the degree of quality change between the SRs with the same team and SRs with different teams for all the 27 items. Compared with SRs with the same intervention, those with changed interventions often better reported 20 items, although the differences between them were not statistically significant. No significant differences were found in the degree of quality improvement between SRs included the trials of the original SRs and SRs did not include the trials included in the original SRs, between SRs fully included trials of original SRs and SRs partially included trials of original SRs, and between SRs with a protocol and SRs without a protocol. (Appendix Table 5).

3.5. Results of regression analyses

No significant differences were observed in fully reported AMSTAR 2 items of original SRs and updated SRs in either bivariate or multiple variable linear regression analyses after adjusting for factors of publication year, involved in statistician, epidemiologist, or methodologist, impact factor, authors from >1 country, and funding support. The fully reported PRISMA items were also not statistically significantly associated with original SRs or updated SRs in either bivariate or multiple variable linear regression analyses after adjusting for factors with imbalanced distribution between SRs (Appendix Table 6).

3.6. Compared with other studies

To compare the methodological quality of the included SRs and SRs in other fields, we selected two recently published studies [31, 32] that evaluated the methodological quality of SRs using the AMSTAR 2 tool. Compared with study conduct by Habtewold et al. [31], the updated SRs and original SRs had significantly higher compliances for items “used a comprehensive literature search strategy” and “described the included studies in adequate detail”, but had lower compliances for items “carried out an adequate investigation of publication bias and discussed its likely impact on the results of the review” and “reported any potential sources of conflict of interest”. The updated SRs and original SRs had significantly higher compliance rates on items 2, 8, and 11 but had significantly lower compliance rates on items 1, 3, 7, 12, 13, 15, and 16 compared with research by Piovani et al [32]. The

details of the comparisons were presented in Appendix Table 7.

Considering the comparisons of reporting quality, we selected two studies [33, 34], including large samples, that explored the reporting quality of SRs in the field of pain and cancer, respectively. Compared with the study of Riado et al. [33], the updated SRs and original SRs had lower compliance in reporting of eligibility criteria, study selection, data collection process, data items, study characteristics, limitations, and funding. The study of Xu et al. [34] often better reported the objectives, eligibility criteria, information sources, data collection process, data items, synthesis of results, risk of bias across studies, additional analyses, study characteristics, results of individual studies, limitations, and funding. However, all the updated SRs and original SRs often better reported the protocol and registration, search, and study selection (Appendix Table 8).

4. Discussion

4.1. Summary of findings

This study identified 30 non-Cochrane updated SRs and 30 non-Cochrane original SRs published between 1994 and 2018 and assessed the methodological quality using AMSTAR 2 tool and reporting quality in terms of the PRISMA checklist. Results indicated that the overall methodological quality and reporting quality of updated SRs were not improved compared with original SRs, but the compliance rates

of many AMSTAR 2 items and PRISMA items were slightly improved with no significant difference. The overall fully reported AMSTAR 2 items and PRISMA items between original SRs and updated SRs were also not statistically significant after adjusting for multiple review factors.

For overall methodological quality, only two SRs were of low quality and the remaining 58 were of critical low quality which indicates that the methodological quality needs to be further improved. Considering individual item of AMSTAR 2, only one item obtained compliance higher than 70.0%. Research protocols are an important feature of SR, which helps to increase the transparency of the review objectives and methods and avoids bias in outcome reporting, and the absence of a protocol may result in post-modification of methods [22, 35]. However, only about 22.0% of SRs provided an explicit statement that the review methods were established before the conduct of the review and clarified the significant deviations from the protocol. Similar to previous studies [28, 31], our study found that none of the SRs explained their selection of the study designs for inclusion in the review and only about 21.7% SRs provided a list of excluded studies and justified the exclusions. If there is a publication bias, the treatment effect may be overestimated even if the bias of risk of the included individual trial is low [36]. But there were still 73.3% of SRs did not carry out an adequate investigation of publication bias and discussed its likely impact on the results of the review. Furthermore, almost none of the SRs reported the sources of funding for the studies included in the review, most SRs did not assess the potential impact of risk of bias in individual studies on the results of the meta-analysis

or other evidence syntheses and did not report potential sources of conflict of interest. There were no statistically significant differences between the updated SRs and the original SRs in all AMSTAR 2 items, indicating that the methodological quality of updated SRs was not significantly improved. Overall, no significant differences were found between SRs with the same team and SRs with different teams, between SRs with the same intervention and SRs with changed interventions, between SRs included the trials that included in the original SRs and SRs did not include the trials included in the original SRs, and between SRs with a protocol and SRs without a protocol, which revealed that these factors did not affect the extent to which the methodological quality changes during the update process.

For reporting quality, the compliance rates of six items were higher than 70.0%, and only one item was fully reported for the updated SRs. The compliance rates of updated SRs were better than original SRs in 15 items, but no items were observed to have statistically significant differences. However, there were many common defects in these SRs, such as “report the protocol and registration”, “present the data collection process”, “assess the risk of bias across studies”, and “clarify the limitations and funding”. What is more, the reporting flaws were also found in the following items: objectives, search, study selection, data items, risk of bias in individual studies, and additional analyses. Although the release of the PRISMA statement was important to quickly improve the reporting quality of SRs [15], current research indicates that this tool does not seem to be well followed. Considering fully reported PRISMA items, the difference between original SRs and updated SRs was

not statistically significant before and after adjusting for multiple review characteristics. For each PRISMA item, there were also no significant differences between SRs with the same team and SRs with different teams, between SRs with same intervention and SRs with changed interventions, between SRs included the trials of the original SRs and SRs did not include trials of the original SRs, and between SRs with a protocol and SRs without a protocol. This means whether the updates were conducted by the same team, whether the intervention was changed, whether the trials of original SRs were included, and whether SRs had a priori protocol did not affect the degree of quality improvement during the update process.

4.2. Compared with other studies

To our knowledge, a previous study published in 2006 has evaluated the methodological quality using Overview Quality Assessment Questionnaire (OQAQ) and reporting quality using Quality of Reporting of Meta-analyses (QUOROM) statement of updated and original versions of the Cochrane SRs [25]. This study revealed that there was no overall improvement in the updated SRs in items of reporting quality and methodological quality, which is similar to our evaluation using AMSTAR 2 and PRISMA checklist. Besides, our research also compared methodological and reporting quality of the included updated SRs and original SRs with SRs in the other fields. It was found that the methodological quality of the updated SRs and original SRs was similar to SRs in the field of biomedical and public

health, as assessed by Habtewold et al [31]. However, the compliances of eleven AMSTAR 2 items of the updated SRs and original SRs were lower than that of the study of Piovani et al [32], although only seven items with significant differences. This indicated that the methodological quality of both the updated SRs and original SRs was lower than that of the SRs of the inflammatory bowel diseases. But all the SRs had the same low compliance rates for ‘‘review contain an explicit statement that the review methods were established prior to the conduct of the review”, “explanation of the selection of the study designs for inclusion”, and “sources of funding for the studies included in the review”. As for reporting quality, the compliance rates of sixteen items of the included SRs were lower than the study of Riado et al. [33], and the compliance rates of twenty items were lower than Xu et al.'s study [34], revealing that the reporting quality of the included SRs was lower than SRs in the field of pain and cancer.

4.3. Suggestions for future work

The current study showed that both the non-Cochrane updated SRs and the original SRs were of low quality according to the AMSTAR 2 and PRISMA checklists, and many items had low compliance rates. The identified deficiencies should be paid more attention, especially for the updated SRs. Researchers, journal editors, and peer reviewers should ensure that the methodological and reporting guidelines are strictly followed prior to publication [15]. Our research confirmed that the methodological

and reporting quality of the included SRs was lower than SRs in the other fields, so the quality of both the original SRs and updated SRs needs to be further improved. A lot of important information was not fully reported for updated SRs, and even some were not reported. However, many SRs are more focused on new data and analytical methods during the update process. But we should know that the scientific quality is not only based on data and results, but also on the rigor and appropriateness of methods to conduct and report the study. Because decision-makers need to examine the evidence before implementing new interventions or diagnostic techniques, the role of SRs is increasingly important in healthcare and SRs must provide reliable and valid evidence. Thus, it is necessary to develop a methodological quality tool and a reporting quality tool that are specifically applicable to the assessment of updated SRs or provide a methodological framework for updating SRs. For example, we can modify some items of the PRISMA checklist, which defines that the report should be identified as an updated SR in the title, the updated SR should report the reason for the update, clarify the studies identified from previous SR, and provide a flow diagram including the studies which identified from previous SR. We have conducted stratified analyses to explore factors that may affect the degree of quality change during the update process. However, all the selected factors did not affect the quality change. Further investigations should be performed to identify other factors related to quality change which can help to improve the quality during the update process.

4.3. Strengths and limitations

This study investigated the methodological and reporting quality of the non-Cochrane updated SRs and original SRs in both overall fully reported items and individual item of AMSTAR 2 and PRISMA checklists, which eliminates the question of potential differences in the weights of the items. Furthermore, we also conducted stratified analyses to examine certain characteristics that may affect the extent of the methodological and reporting quality changes during the update process. Findings from this study can be used to promote improvements in the quality and method of updated SR. However, our study has some limitations. First, the sample size included in this study is not large enough, although we searched two databases to incorporate all the eligible SRs and manually searched the reference lists of included SRs to obtain additional references. Since we were unable to identify all the SRs that did not explicitly state "update" in the titles or abstracts, many of them may have been ignored. Second, only SRs of RCTs published in English were enrolled, the results may not apply to SRs published in other languages and SRs of other types such as cohort studies and observational studies [22]. Third, the results of some stratified analyses on methodological quality and reporting quality may be less convincing because some subgroups contain only a small number of SRs. Fourth, since different teams may differ in the criteria for the assessment of AMSTAR 2 and PRISMA items, the credibility of the results of comparison with the quality of SRs in other areas may be weakened. Finally, we only used AMSTAR 2 and PRISMA assessed the methodological and reporting quality and did not evaluate the quality of Cochrane

systematic reviews. Further studies to explore the quality using other checklists and compare the quality of Cochrane reviews or reviews published in other databases between the updated SRs and original SRs are still needed.

5. Conclusions

The methodological and reporting quality of non-Cochrane updated SRs were not improved compared with the original SRs, although compliance rates were slightly improved on certain individual items. There is room for improvement of methodological and reporting quality for both original and updated SRs. The identified methodological and reporting deficiencies should be paid more attention, especially for the updated SRs. Future research should insistent on developing a methodological or a reporting quality tool that is specifically applicable to updated SRs.

Abbreviations

SR: Systematic review; AMSTAR 2: Assessment of Multiple Systematic Reviews 2; PRISMA: Preferred Reporting Items of Systematic reviews and Meta-Analyses; RCT: randomized controlled trial.

Acknowledgments

We would like to thank Dr. Long Ge, Evidence-Based Medicine Center, School of

Basic Medical Sciences, Lanzhou University, who provided input on the methods of this manuscript.

Funding

None.

Authors' contributions

YG and JT planned and designed the study. YG and JT developed search strategies. YG, YC, KY, and ML screened potential studies and extracted data from the included studies. YG, YC, KY, SS, JC, and YS performed the quality assessment. YG and JT performed the statistical analysis. JZ, FS, and JT conducted arbitration under disagreement and ensured that there were no errors. YG, YC, and JT wrote the first draft. JZ, FS, and JT revised the draft. All authors approved the final version of the manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

[1] Mulrow CD. Rationale for systematic reviews. BMJ (Clinical research ed). 1994;309:597-9.

[2] Gao Y, Ge L, Ma X, Shen X, Liu M, Tian J. Improvement needed in the network geometry and inconsistency of Cochrane network meta-analyses: a cross-sectional survey. Journal of clinical epidemiology. 2019;113:214-27.

[3] Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. Annals of internal medicine. 1997;126:376-80.

[4] Bashir R, Surian D, Dunn AG. The risk of conclusion change in systematic review updates can be estimated by learning from a database of published examples. Journal of clinical epidemiology. 2019;110:42-9.

[5] Cohen AM, Ambert K, McDonagh M. Studying the potential impact of automated document classification on scheduling a systematic review update. BMC medical informatics and decision making. 2012;12:33.

[6] Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. Lancet (London, England). 2014;383:166-75.

[7] Page MJ, Shamseer L, Altman DG, Tetzlaff J, Sampson M, Tricco AC, et al. Epidemiology and Reporting Characteristics of Systematic Reviews of Biomedical Research: A Cross-Sectional Study. PLoS medicine. 2016;13:e1002028.

[8] Tian J, Zhang J, Ge L, Yang K, Song F. The methodological and reporting quality of systematic reviews from China and the USA are similar. Journal of clinical epidemiology. 2017;85:50-8.

[9] Garner P, Hopewell S, Chandler J, MacLehose H, Schunemann HJ, Akl EA, et al. When and how to update systematic reviews: consensus and checklist. BMJ (Clinical research ed). 2016;354:i3507.

[10] Hopewell S, Clarke M, Stewart L, Tierney J. Time to publication for results of clinical trials. Cochrane database of systematic reviews. 2007:Mr000011.

[11] Murad MH, Montori VM, Ioannidis JP, Jaeschke R, Devereaux PJ, Prasad K, et al. How to read a systematic review and meta-analysis and apply the results to patient care: users' guides to the medical literature. Jama. 2014;312:171-9.

[12] Montori VM, Devereaux PJ, Adhikari NK, Burns KE, Eggert CH, Briel M, et al. Randomized trials stopped early for benefit: a systematic review. Jama. 2005;294:2203-9.

[13] Ioannidis J, Lau J. Evolution of treatment effects over time: empirical insight from recursive cumulative metaanalyses. Proceedings of the National Academy of Sciences of the United States of America. 2001;98:831-6.

[14] Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Quarterly. 2016;94:485-514.

[15] Tonin FS, Borba HH, Leonart LP, Mendes AM, Steimbach LM, Pontarolo R, et al. Methodological quality assessment of network meta-analysis of drug interventions: implications from a systematic review. International journal of epidemiology.

2019;48:620-32.

[16] Carroll K, Hemmings R. On the need for increased rigour and care in the conduct and interpretation of network meta-analyses in drug development. Pharmaceutical Statistics. 2016;15:135-42.

[17] Zhang J, Carlin BP, Neaton JD, Soon GG, Chu H. Network meta-analysis of randomized clinical trials: reporting the proper summaries. Clinical Trials. 2013;11:246-62.

[18] Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Reducing waste from incomplete or unusable reports of biomedical research. Lancet (London, England). 2014;383:267-76.

[19] Pussegoda K, Turner L, Garritty C, Mayhew A, Skidmore B, Stevens A, et al. Identifying approaches for assessing methodological and reporting quality of systematic reviews: a descriptive study. Systematic reviews. 2017;6:117.

[20] Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Journal of clinical epidemiology. 2009;62:1006-12.

[21] 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.

[22] Ge L, Tian JH, Li YN, Pan JX, Li G, Wei D, et al. Association between prospective registration and overall reporting and methodological quality of

systematic reviews: a meta-epidemiological study. Journal of clinical epidemiology. 2018;93:45-55.

[23] Li XX, Zheng Y, Chen YL, Yang KH, Zhang Z, Z. The reporting characteristics and methodological quality of Cochrane reviews about health policy research. Health policy (Amsterdam, Netherlands). 2015;119:503-10.

[24] Thulliez M, Angoulvant D, Pisella PJ, Bejan-Angoulvant T. Overview of Systematic Reviews and Meta-analyses on Systemic Adverse Events Associated With Intravitreal Anti-Vascular Endothelial Growth Factor Medication Use. JAMA ophthalmology. 2018;136:557-66.

[25] Shea B, Boers M, Grimshaw JM, Hamel C, Bouter LM. Does updating improve the methodological and reporting quality of systematic reviews? BMC medical research methodology. 2006;6:27.

[26] Shea B, Bouter LM, Grimshaw JM, Francis D, Ortiz Z, Wells GA, et al. Scope for improvement in the quality of reporting of systematic reviews. From the Cochrane Musculoskeletal Group. The Journal of rheumatology. 2006;33:9-15.

[27] 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 medical research methodology. 2007;7:10.

[28] Yan P, Yao L, Li H, Zhang M, Xun Y, Li M, et al. The methodological quality of robotic surgical meta-analyses needed to be improved: a cross-sectional study. Journal of clinical epidemiology. 2019;109:20-9.

[29] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al.

The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ (Clinical research ed). 2009;339:b2700.

[30] Li X, Wang R, Shi X, Su J, Pan Y, Tian J, et al. Reporting Characteristics and Quality of Systematic Reviews of Acupuncture Analgesia. Pain practice : the official journal of World Institute of Pain. 2017;17:1066-74.

[31] Habtewold TD, Alemu SM, Mohammed SH, Endalamaw A, Mohammed MA, Teferra AA, et al. Biomedical and public health reviews and meta-analyses in Ethiopia had poor methodological quality: overview of evidence from 1970 to 2018. Journal of clinical epidemiology. 2019;109:90-8.

[32] Piovani D, Danese S, Peyrin-Biroulet L, Nikolopoulos GK, Lytras T, Bonovas S. Environmental Risk Factors for Inflammatory Bowel Diseases: An Umbrella Review of Meta-analyses. Gastroenterology. 2019;157:647-59.

[33] Riado Minguez D, Kowalski M, Vallve Odena M, Longin Pontzen D, Jelicic Kadic A, Jeric M, et al. Methodological and Reporting Quality of Systematic Reviews Published in the Highest Ranking Journals in the Field of Pain. Anesthesia and analgesia. 2017;125:1348-54.

[34] Xu C, Liu Y, Zhang C, Kwong JSW, Zhou JG, Ge L, et al. An overview on the methodological and reporting quality of dose-response meta-analysis on cancer prevention. Journal of cancer research and clinical oncology. 2019;145:1201-11.

[35] Kelly SE, Moher D, Clifford TJ. Quality of conduct and reporting in rapid reviews: an exploration of compliance with PRISMA and AMSTAR guidelines.

Systematic reviews. 2016;5:79.

[36] Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. Journal of clinical epidemiology. 2011;64:1277-82.

Figure legends

Figure 1. The flowchart of the screening process.

Figure 2. Compliance rate of updated SRs and originals SRs in each AMSTAR 2 item

Figure 3. Comparison of the compliance rates of updated SRs and originals SRs in each AMSTAR 2 item. The X-axis represented each AMSTAR 2 item, the Y-axis represented the OR of updated SRs compared to original SRs, the bubble size displayed the number of OR, the color of the bubble indicated the OR is greater than 1, equal to 1, or less than 1. USRs, updated systematic reviews; OSRs, original systematic reviews; OR, odds ratio; AMSTAR 2, Assessment of Multiple Systematic Reviews 2.

Figure 4. Compliance rate of updated SRs and originals SRs in each PRISMA item

Figure 5. Comparison of the compliance rates of updated SRs and originals SRs in each PRISMA item. The X-axis represented each PRISMA item, the Y-axis represented the OR of updated SRs compared to original SRs, the bubble size displayed the number of OR, the color of the bubble indicated the OR is greater than 1, equal to 1, or less than 1. USRs, updated systematic reviews; OSRs, original systematic reviews; OR, odds ratio; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-analyses.

Supplementary files

Appendix 1. Search strategy of Embase.

Appendix Word 2. List of included SRs

Appendix Table 1. AMSTAR 2 assessment by items and overall methodological quality.

Appendix Table 2. Comparison of the methodological quality of updated SRs and original SRs in AMSTAR 2 items.

Appendix Table 3. Stratified analyses of changes of methodological quality assessment in AMSTAR 2 items.

Appendix Table 4. Comparison of the reporting quality of updated SRs and original SRs in PRISMA items.

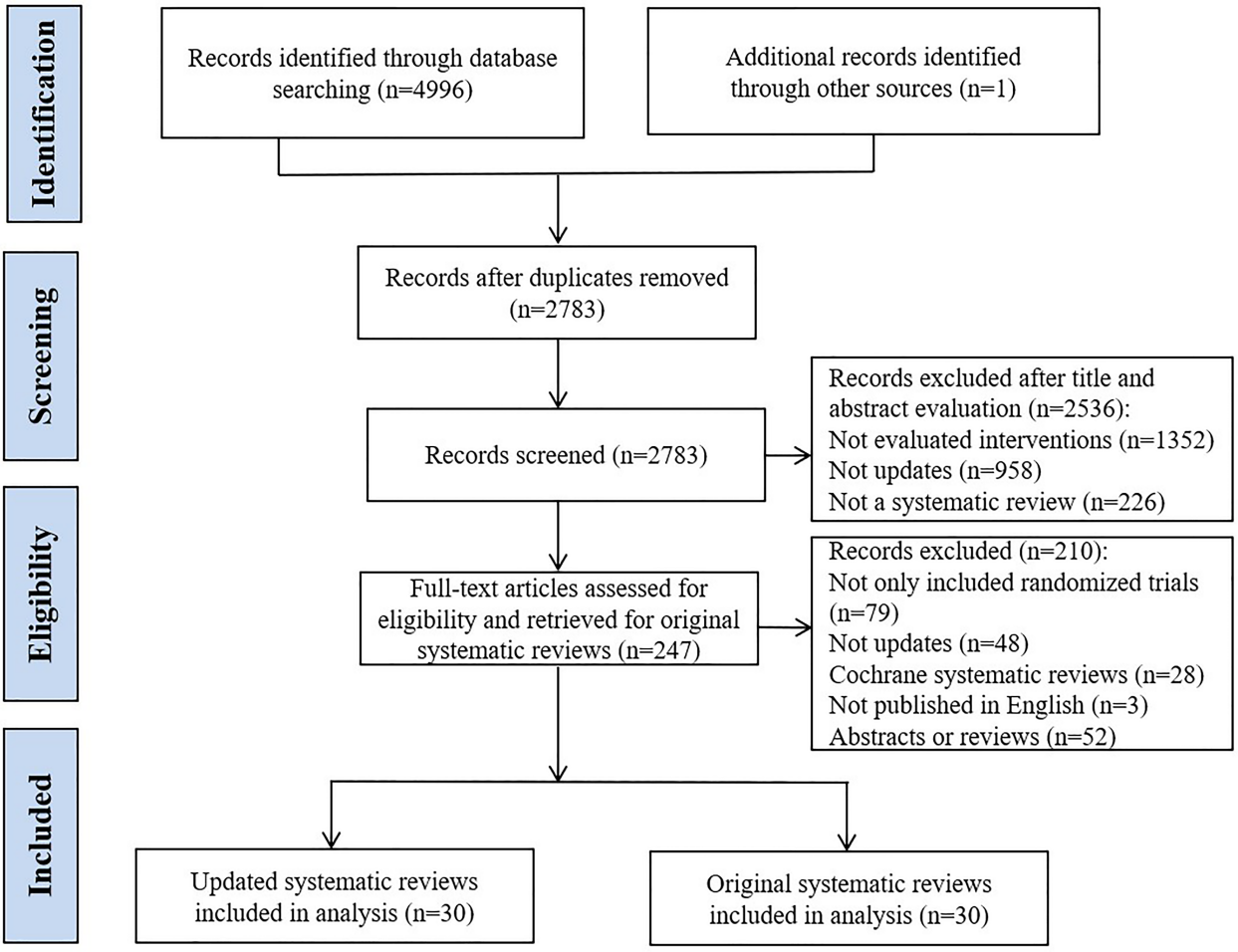
Appendix Table 5. Stratified analyses of changes of reporting quality assessment in PRISMA items.

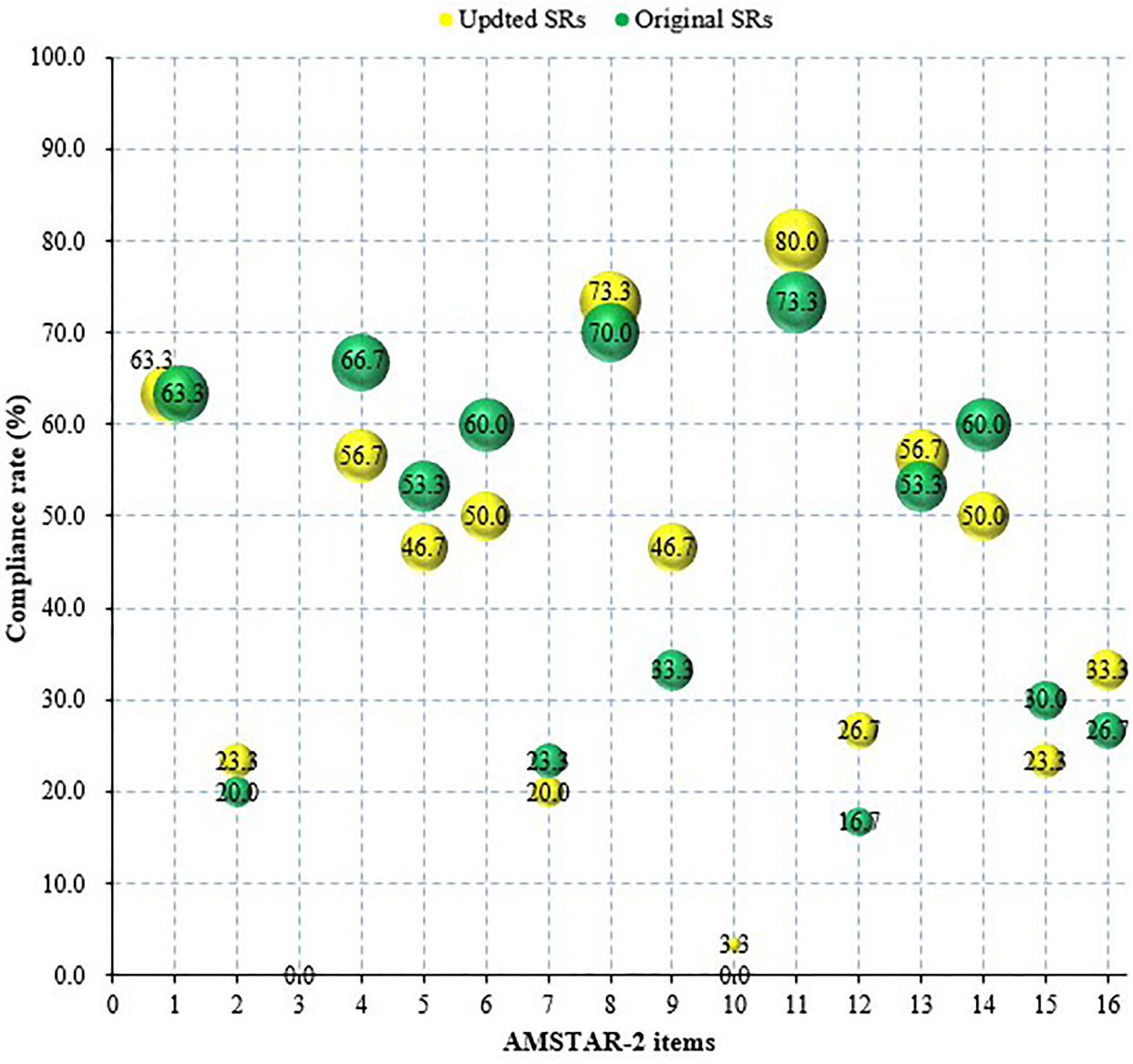
Appendix Table 6. Linear regression analysis of potential factors affecting the fully

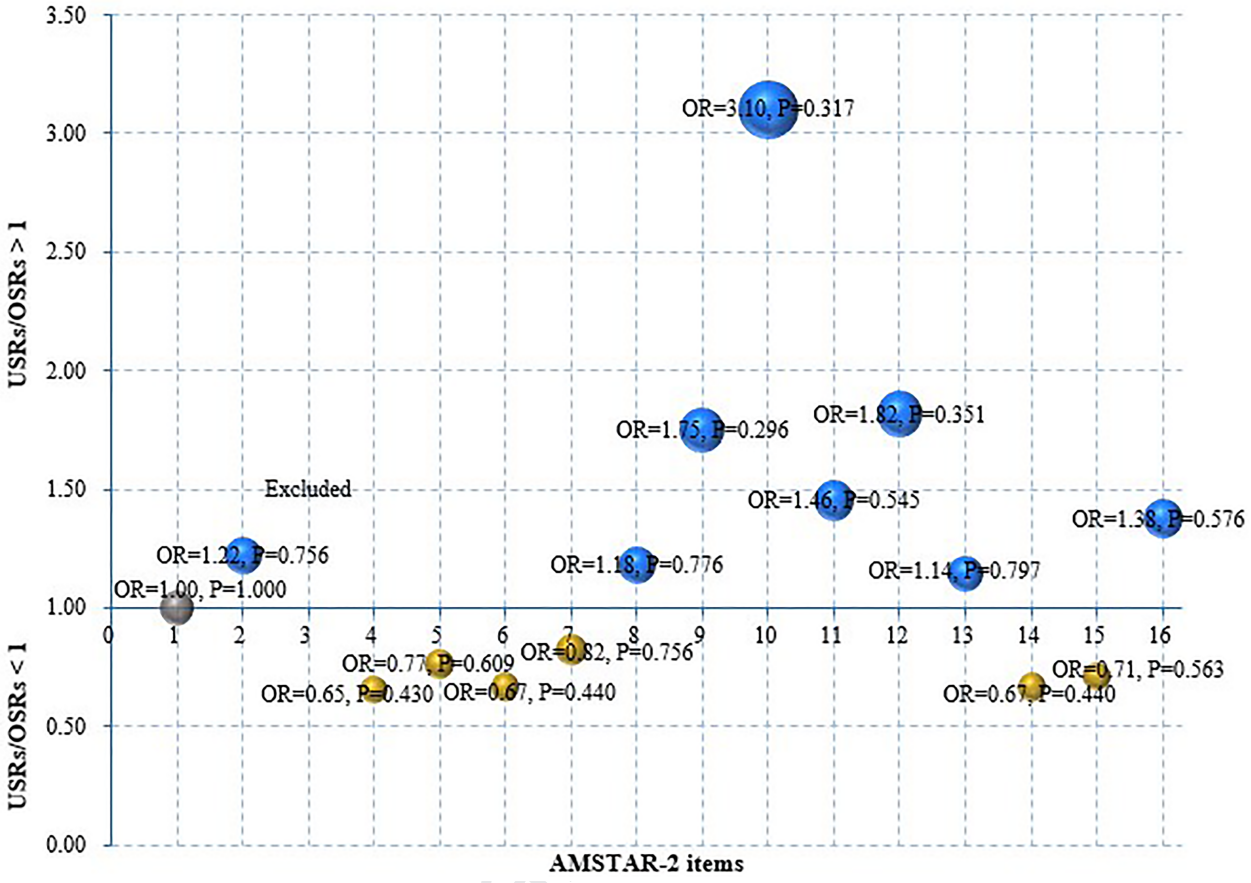
reported AMSTAR 2 and PRISMA items.

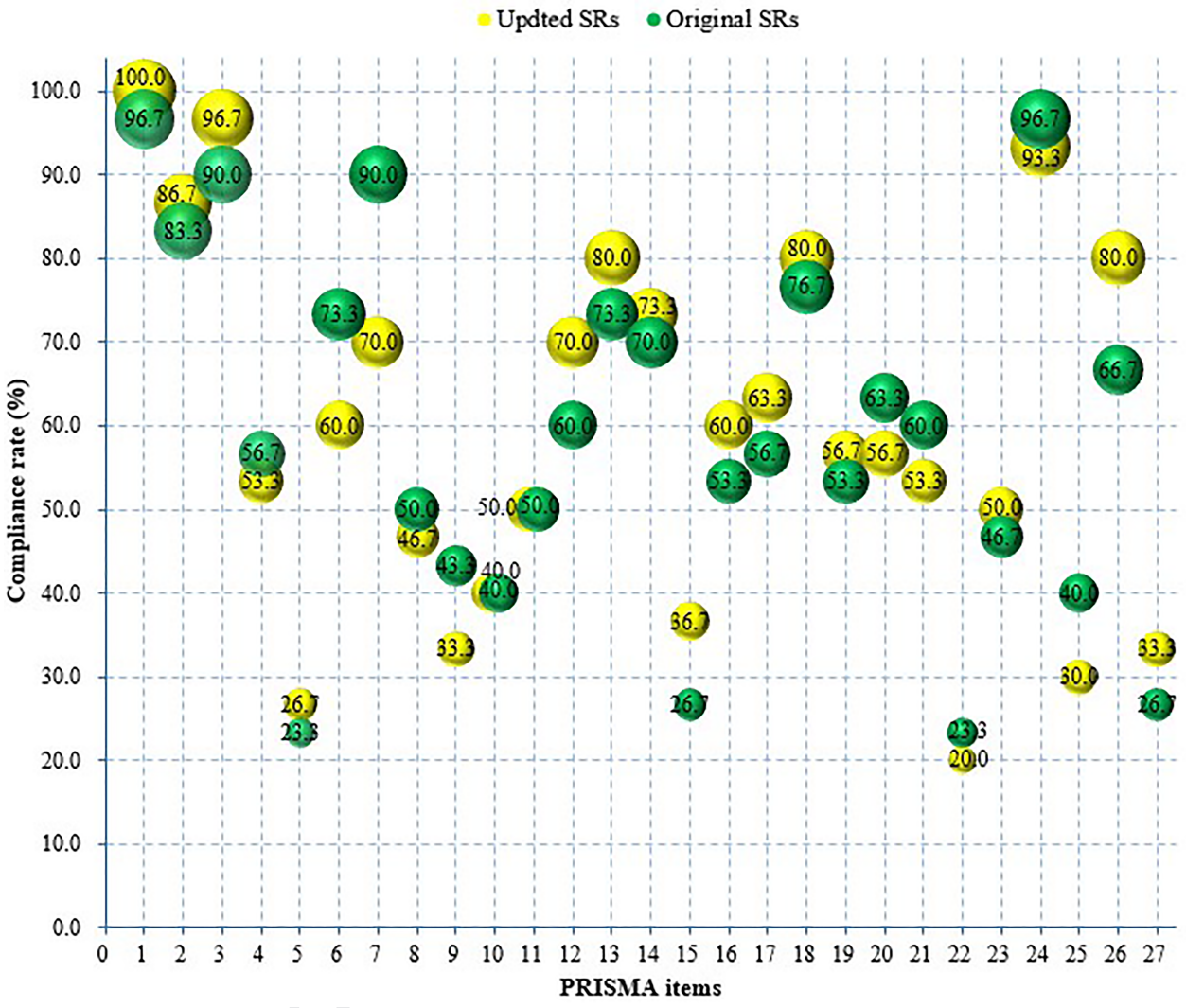
Appendix Table 7. Comparison of the methodological quality of the included SRs with SRs in other fields.

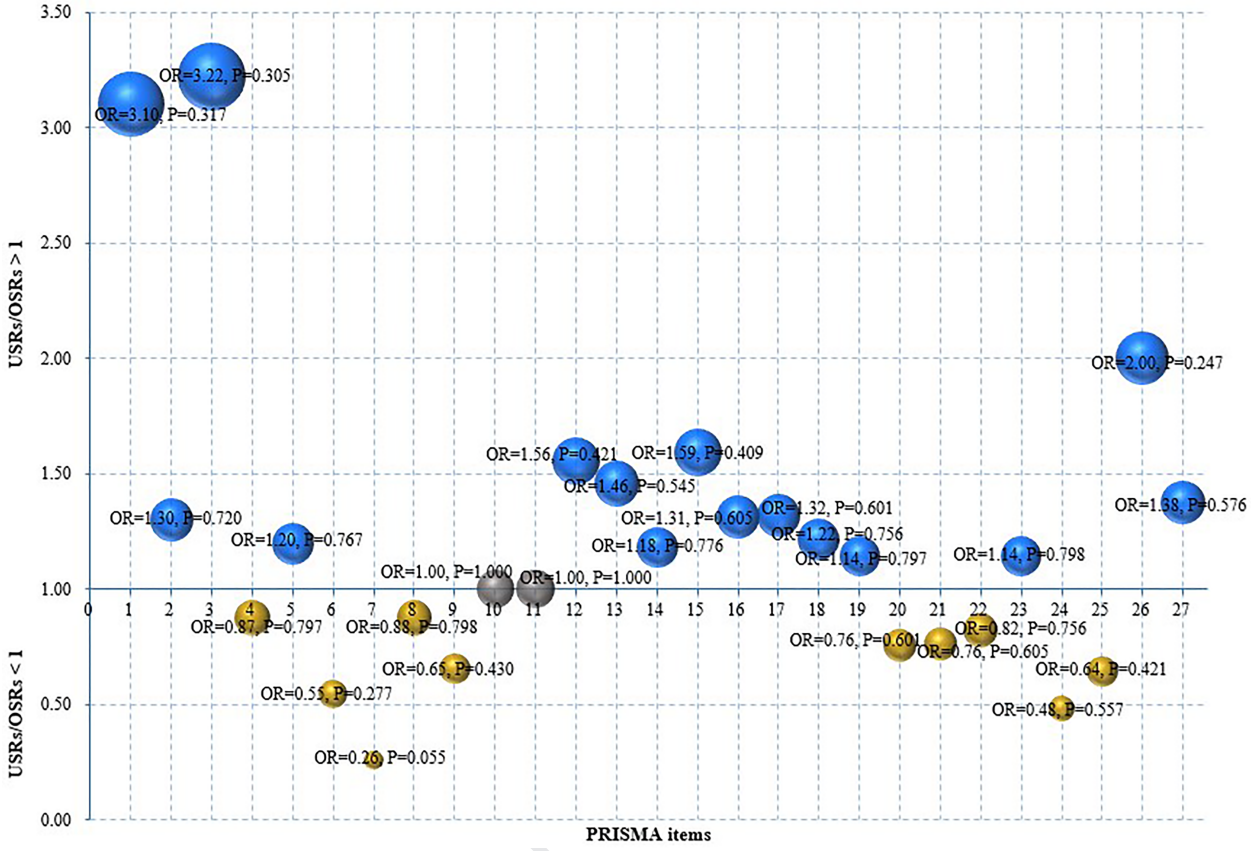
Appendix Table 8. Comparison of the reporting quality of the included SRs with SRs in other fields.


Identification
Screening
Eligibility
Included
Records identified through database searching (n=4996)
Additional records identified through other sources (n=1)
Records after duplicates removed (n=2783)
Records screened (n=2783)
Records excluded after title and abstract evaluation (n=2536):
Not evaluated interventions (n=1352)
Not updates (n=958)
Not a systematic review (n=226)
Full-text articles assessed for eligibility and retrieved for original systematic reviews (n=247)
Records excluded (n=210):
Not only included randomized trials (n=79)
Not updates (n=48)
Cochrane systematic reviews (n=28)
Not published in English (n=3)
Abstracts or reviews (n=52)
Updated systematic reviews included in analysis (n=30)
Original systematic reviews included in analysis (n=30)


Updted SRs
Original SRs
Compliance rate (%)
AMSTAR-2 items
100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
63.3
63.3
23.3
20.0
0.0
66.7
56.7
53.3
46.7
60.0
50.0
23.3
20.0
73.3
70.0
46.7
33.3
3.3
0.0
80.0
73.3
26.7
16.7
56.7
53.3
60.0
50.0
30.0
23.3
33.3
26.7


3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
USRs/OSRs > 1
USRs/OSRs < 1
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
AMSTAR-2 items
OR=3.10, P=0.317
OR=1.75, P=0.296
OR=1.82, P=0.351
Excluded
OR=1.46, P=0.545
OR=1.38, P=0.576
OR=1.22, P=0.756
OR=1.18, P=0.776
OR=1.14, P=0.797
OR=1.00, P=1.000
OR=0.82, P=0.756
OR=0.77, P=0.609
OR=0.71, P=0.563
OR=0.65, P=0.430
OR=0.67, P=0.440
OR=0.67, P=0.440


Updted SRs
Original SRs
Compliance rate (%)
PRISMA items
0.0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0
90.0
100.0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27


OR=3.22, P=0.305
OR=3.10, P=0.317
OR=2.00, P=0.247
OR=1.56, P=0.421
OR=1.59, P=0.409
OR=1.46, P=0.545
OR=1.38, P=0.576
OR=1.30, P=0.720
OR=1.31, P=0.605
OR=1.32, P=0.601
OR=1.20, P=0.767
OR=1.22, P=0.756
OR=1.18, P=0.776
OR=1.14, P=0.797
OR=1.14, P=0.798
OR=1.00, P=1.000
OR=1.00, P=1.000
OR=0.87, P=0.797
OR=0.88, P=0.798
OR=0.82, P=0.756
OR=0.76, P=0.601
OR=0.76, P=0.605
OR=0.65, P=0.430
OR=0.64, P=0.421
OR=0.55, P=0.277
OR=0.48, P=0.557
OR=0.26, P=0.055
USRs/OSRs > 1
USRs/OSRs < 1
PRISMA items