BIOCELL
2019 43(4): 233-251

Tech Science Press

Retrieval, reporting and methodological characteristics for systematic reviews/meta-analyses of animal models: a meta-epidemiological study

Shuzhen SHI[1,2]; Ming LIU[1,2]; Wenjuan MA[1,2]; Ya GAO[1,2]; Long GE[3]; Xiping SHEN[3]; Jiarui WU[4]; Junhua ZHANG[5,*]; Jinhui TIAN[1,2,*]

[1] Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China
[2] Key Laboratory of Evidence-Based Medicine and Knowledge Translation of Gansu Province, Lanzhou, China
[3] School of Public Health, Lanzhou University, Lanzhou, China
[4] Department of Clinical Chinese Pharmacy, School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing, China
[5] Evidence-Based Medicine Center, Tianjin University of Traditional Chinese Medicine, Tianjin, China

Key words: Animal models, AMSTAR 2, Meta-epidemiological Study, PRESS, PRISMA, Systematic review/Meta-analysis

Abstract: The study aimed to analyze the reporting and methodological quality of systematic reviews (SRs)/meta-analyses (MAs) of animal models to provide references for later studies and avoid the waste of medical resources. EMBASE and MEDLINE databases were searched from inception to November 2017, with no language restriction. Two reviewers selected inclusion dependently and extracted the basic characteristics. Review Manager 5.3, stata 12.0, and SPSS 21 software were used to conduct analyses. A total of 46 SRs/MAs were included. The results showed that the English databases with high retrieval frequency are PubMed/MEDLINE, EMBASE, and Web of Science. 67.31% (31/46) of the articles reported the search strategy in the full text or the appendix. 65.22% (30/46) reported the literature screening flow diagram, and only 19.57% (9/46) reported the number of works of literature retrieved in each database. 60.87% (28/46) illustrated supplement retrieval. Through 2 subgroup analyses, it was found that there were no significant differences in the quality of reports of PRISMA items. But referring to the methodological quality or reporting of PRESS items, SCI was better than that of non-SCI, while there seemed a source of funding to have no significant impact on the methodological quality or the items of PRESS. The results of PRESS, AMSTAR 2, and PRISMA were correlated, and the correlation between PRISMA and AMSTAR 2 was strong. These results demonstrated that search strategies of animal model SRs/MAs are still not enough comprehensive, report specification and methodological quality still need to be ameliorated. To show users the scientificity and rigor of the study, future research should focus on these various guidelines like PRESS, PRISMA, and AMSTAR 2 checklists that have been issued, it can help to increase the value of research and improve the utilization of medical resources.

Introduction

Animal models are not the perfect tools for the full understanding of human development and behavior, but they can be an important place to start (Tania *et al.*, 2014), its main purpose in enhancing our understanding of physiologic and pathologic processes (Mueller *et al.*, 2014). Appropriate animal models play an important role for pathogenesis analysis in many areas of diseases, for example, prevention (Sena *et al.*, 2018), treatment (Nunes *et al.*, 2018), diagnosis of prognostic markers (Pathuri *et al.*, 2016), drug screening and evaluation (Beedie *et al.*, 2017), and vaccine development (Cardona and Williams, 2017). It also could provide convincing measures of safety and early-stage toxicity where these are of generic concern (McGonigle and Ruggeri, 2014).

Systematic, which is the defining feature of the review itself, and results in the systematic review being considered the highest level of evidence available to guide and inform practice (Baker and Weeks, 2014). Systematic reviews or meta-analyses (SRs/MAs) critically appraise and formally synthesize the best existing evidence to provide a conclusion statement that answers specific clinical questions (Harris *et al.*, 2014). As SR is becoming a developing methodology, a previous study (Moher *et al.*, 2007) by only searching

*Address correspondence to: Jinhui Tian, tjh996@163.com; Junhua Zhang, zjhtcm@foxmail.com

MEDLINE showed that the annual publication rate of SRs on MEDLINE is more than 8000, equivalent to 22 SRs per day. In the past, however, there has been little consistency in and poor quality of reporting of searches, so it is often difficult to evaluate the actual searches that were conducted (Sampson *et al.*, 2008). Therefore, it important to assess the quality of published SRs/MAs before their results are implemented into clinical or public health practice (Ge *et al.*, 2016; Li *et al.*, 2015). In this article, we will analyze the reporting quality, methodological quality, and the relationships of the SRs/MAs of the disease intervention animal model and provide a comprehensive and reliable basis for the application of the research results of similar articles. And we hope through our study to illustrate the current quality of methodology and reporting for SRs of animal models, suggesting that future researchers in the production of SRs should make full use of the data of the original studies included so that its value can be maximized and avoid the waste of the medical resources.

Materials and Methods

Eligibility criteria

We included SRs/MAs of the animal models with interventions about diseases, which is only related animal models or SRs/MAs, or about physiological and pathological mechanism, or SRs/MAs that about comparison of various animal models of the same disease, or conference abstracts, or the articles with unavailable full text were excluded.

Search strategy

EMBASE and MEDLINE databases were searched via EMBASE.com platform from inception to November 2017. ('meta analysis'/exp or 'meta analysis (topic)'/exp or 'systematic review'/exp or 'systematic review (topic)'/exp or 'meta analysis': ti, ab or 'meta analyses': ti, ab or 'meta-analysis': ti, ab or 'meta-analyses': ti, ab or 'systematic review': ti, ab or 'systematic reviews': ti, ab or metaanalysis: ti, ab or metaanalyses: ti, ab) and "(animal model)/lim" were used as search strategy. No time and language restrictions. Two reviewers developed independently.

Identification and selection

Literature search records were imported into EndNote X7 software. When the duplicates were removed, two independent reviewers examined the title and abstract of retrieved records to identify relevant SRs/MAs, and then the same two authors independently examined full text according to the eligibility criteria. Any disagreement was resolved by the discussion between the two reviewers, or through arbitration by a third party.

Data extraction and management

Two reviewers established a spreadsheet using Microsoft Excel 2010, piloted and refined this form using three initial studies. The following data were extracted from each study: (i) title; (ii) the name of the journal; (iii) the number of author(s); (iv) year of publication; (v) methodological information, including retrieval related content and the data analysis; (vi) result; (vii) categories of disease; (viii) categories of journals; (ix) funding sources.

Entries of Peer Review of Electronic Search Strategies (PRESS) (Mcgowan *et al.*, 2010) were used to evaluate the details of the retrieval. System assessment and meta-analysis preferred reporting items (PRISMA) are the latest preferred tools to standardize or evaluate the quality of SR reports (Moher *et al.*, 2009). PRISMA statement (Swartz, 2011) is a total of 27 items that report the quality of SR in terms of titles, abstracts, introduction, methods, results, discussion, and funding. AMSTAR 2 is used to detect methodological quality; it was based on the related interpretation for AMSTAR 2 (Shea *et al.*, 2017). We performed "Yes, Partial, No" three stratifications for each entry to describe the degree of compliance, at the same time, the "Unreported" and "Not applicable" were added respectively for the entries in PRESS and AMSTAR 2. Two authors were asked to use PRISMA, AMSTAR 2 and PRESS for the evaluation of reporting and methodological quality of the three articles. And the intraclass correlation coefficient (ICC), which is commonly used to estimate the similarity between quantitative measures obtained from different sources (Aly *et al.*, 2014), will determine the reviewers' consistency in checklist evaluation. When ICC reached 0.90, they can evaluate all the included literature.

Review Manager 5.3 software was used to compare and analyze the extracted information. The factors of subgroup analysis were presented as following: publication type (Science Citation Index (SCI) *vs.* non-SCI papers), funding sources (funding and no funding papers). The odds ratio (OR) value and 95% CI was chosen as the summary statistic for subgroup comparisons, by using a fixed-effect model. The items in each checklist used stata12.0 software for the proportion of occurrence and confidence interval. In order to analyze whether the results of the above three checklists were correlated, SPSS 21 was used to conduct linear analyses. We treated all items as "Yes" or "No", it meant as long as they were not "Yes", they would be "No". Each "Yes" was 1, "No" was 0, and calculated the total score of each entry in the literature under each checklist. Finally, in order to balance the bias caused by the difference in the number of checklist entries, we calculated the average score for each entry in each checklist. The score was finally used as data for linear analyses.

Results

Search results

The total number of records identified by the searches for this study was 1676. 22 citations were duplicates, 1654 citations were included initially for further research. According to screening the title and abstract of SRs/MAs, when necessary we will refer to the full-text, 1608 were excluded for reasons: 1442 were not meeting the inclusion criteria obviously, 103 were not SRs but animal models, and 14 were related to physiological and pathological mechanism, 47 were conference abstracts, 2 articles with unavailable full text. Finally, 46 articles were included (Fig. 1).

Of the entire included publications, one was in Chinese. And another article's data was not obtained through search but got through the assistance of research centers; in all 46 articles, 40 were SCI articles, only 20 articles reported funding sources. Among the disease types, the most studied are neurological diseases such as

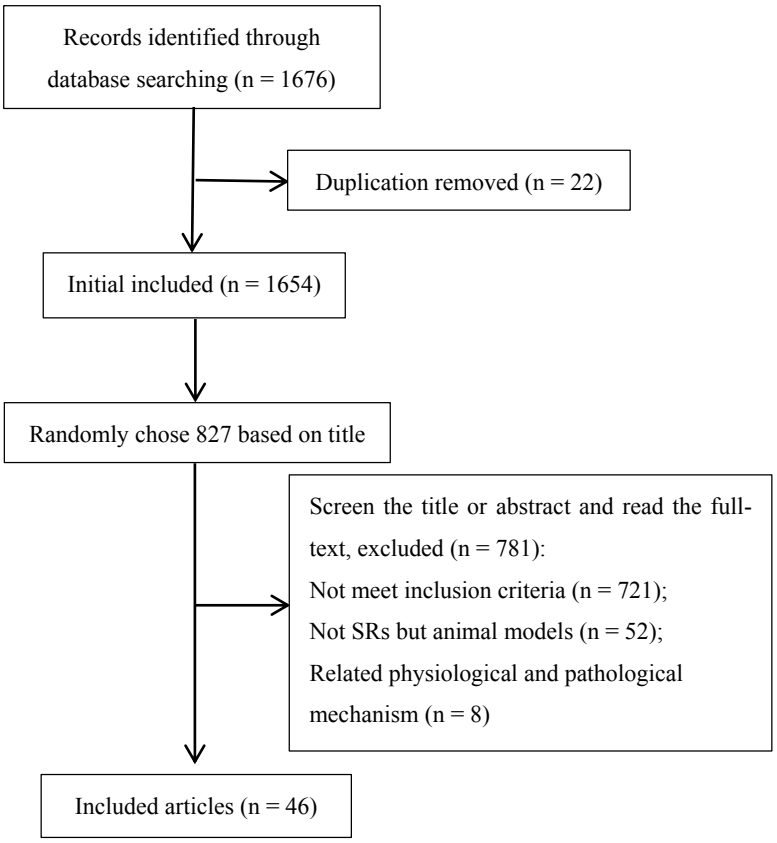


FIGURE 1. Selection of studies included in the paper.

cerebral hemorrhage, cerebral ischemia, spinal cord injury, etc. (general characteristics are shown in Appendix 1-1).

Analysis of retrieval section

General information about the retrieved databases

The median of searched databases of all included articles was 3. The median of searched databases of all Science Citation Index (SCI) articles (range is 0 to 7) and all non-SCI articles (range is 2 to 7) is 3; there was no statistical difference in the number of searched databases (Tab. 1, Appendix 2-1). All funding articles (range is 0 to 7) and all no funding articles (range is 1 to 7) median was also 3; no statistical difference was found between them (Tab. 1, Appendix 2-2).

The most frequent databases in all 46 articles were PubMed/Medline (95.65%), EMBASE (69.57%), and Web of Science (34.78%). In SCI or non-SCI articles, the ranks were

TABLE 1

The database quantity analysis of included literature

Number of databases	Publication type			Funding source		
	SCI	non-SCI	OR(95% CI)	Funding	No funding	OR(95% CI)
0	1	0	0.49(0.02-13.47)	1	0	4.08(0.16-105.52)
1	5	0	2.01(0.10-41.00)	2	3	0.85(0.13-5.65)
2	8	1	1.25(0.13-12.25)	3	6	0.59(0.13-2.71)
3	15	3	0.60(0.11-3.36)	6	12	0.50(0.15-1.71)
4	4	0	1.60(0.08-33.46)	2	2	1.33(0.71-10.39)
5	2	0	0.84(0.04-19.66)	2	0	7.16(0.32-158.02)
6	2	1	0.26(0.02-3.46)	1	2	0.63(0.05-7.50)
7	3	1	0.41(0.04-4.69)	3	1	4.41(0.42-46.05)

the same, the frequency respectively is PubMed/Medline (95%, 100%), EMBASE (67.5%, 83.33%), and Web of Science (35%, 33.33%). There were no statistical differences among the various databases (Tab. 2, Appendix 2-3). Sorting by funding source, the rank is the same as the publication type; there was also no statistical difference (Tab. 2, Appendix 2-4).

The combination of databases

There were 5 articles only retrieving a database, which was PubMed. 9 articles retrieved 2 databases, of them, 7 (19.57%) retrieved "PubMed + EMBASE", and the other two were "PubMed + web of science". In 3 databases searched articles, the most combination was "PubMed + EMBASE + BIOSIS" (13.04%), the second one was "PubMed + EMBASE + web of science" (10.87%), "PubMed + EMBASE + Cochrane library" was the third (4.35%). Each database combination for 4, or 5, or 6 databases appeared once, respectively. In a combination of seven databases, "PubMed + EMBASE + Web of Science + CNKI + CBM + VIP + Wanfang Data" found in 2 articles, the other two just appeared once (Appendix 1-2).

The reporting quality in the retrieval section of all articles

The ICC value of PRESS is shown in Appendix 1-3. None of the articles reported and satisfied all 30-items of PRESS. There is no article about these three items, respectively is 'Is the width of any proximity operators correct' (item 10, 0%), 'Are sub-heading attached to subject heading' (item 15, 0%), 'Are sub-heading used instead of relevant subject heading and vice versa' (item 16, 0%), and it showed that the weakest reported areas (The proportion of reports is less than 50%) within the included literature were the reporting of 'If NOT is used is this likely to result in any unintended exclusions' (item 8, 9%), 'Does the search miss any synonyms' (item 19, 39%), 'Does the search miss truncation or truncate at the wrong point' (item 20, 24%), 'Is starring (restricted to focus) used and if so, is there adequate justification for this' (item 28, 4%). Of course, not all entries have such low reporting rates. Other than the above items, there were 7 items reporting more than 90%, having a comparatively integrated reporting and fit the right method, 8 items between 80 and 90 %, the remaining 8 items are between 50 and 80 % (Tab. 3).

Comparison of PRESS entries based on category of journals and funding source

Since items 10, 15, and 16 are not shown in 46 articles, they are impossible to estimate. By comparing the remaining 27 items, it was found that item 3 and item 28 had statistical significance, while the remaining 25 items had no significant statistical difference. On the whole, the reporting items of SCI papers are not more comprehensive than those of non-SCI papers (Appendix 2-5). Similarly, there was no significant difference between articles with and without funding by comparing items of PRESS reports (Appendix 2-6).

The results of quality assessment of reporting (n = 46)

The ICC value of PRISMA is shown in Appendix 1-4. In general, the report of PRISMA entry is poor (The details for each entry are shown in Tab. 4). There are six items with a reported frequency of 70% or more, they are title (item 1, 94%), rational (item 3, 100%), summary measures (item 13, 78%), results of individual studies (item 20, 70%), summary of evidence (item 24, 100%), conclusions (item 26, 74%), and the compliance ratio is less than 30% in a structured summary (item 2, 26%), objective (item 4, 7%), protocol and registration (item 5, 7%), eligibility criteria (item 6, 28%), information sources (item 7, 11%), search (item 8, 24%), study selection (item 9, 15%), data collection process (item 10, 22%), study characteristics (item 18, 30%), funding (item 27, 24%).

Comparison of PRISMA entries based on Category of journals

In order to explore the difference between report quality in the SCI and non-SCI articles, we compared entries of PRISMA that both are satisfying in the two types of articles. Item 3 and item 2 are fully reported in both SCI and non-SCI articles, therefore, no comparison can be made. However, in addition to these two entries, no significant differences were found after comparing the others (Appendix 2-7).

TABLE 2

The various databases analysis of included literature

database	Publication type			Funding source		
	SCI	Non-SCI	OR(95% CI)	Funding	No funding	OR(95% CI)
BIOSIS	9	1	1.45(0.15-14.07)	5	5	1.40(0.34-5.71)
Cochrane library	6	2	0.35(0.05-2.37)	3	5	0.74(0.15-3.55)
EMBASE	27	5	0.42(0.04-3.93)	14	18	1.04(0.29-3.69)
PubMed/Medline	38	6	1.18(0.05-27.59)	19	25	0.76(0.04-12.95)
Web of Science	14	2	1.08(0.17-6.63)	9	7	2.22(0.65-7.64)
CBM	3	1	0.41(0.04-4.69)	3	1	4.41(0.42-46.05)
CNKI	6	1	0.88(0.09-8.94)	5	2	4.00(0.69-23.30)
VIP	4	1	0.56(0.05-6.02)	3	2	2.12(0.32-14.07)
Wanfang Data	4	1	0.56(0.05-6.03)	3	2	2.12(0.32-14.07)

TABLE 3

The reporting quality of retrieval

PRESS items		Yes		Partial		No		Unreported	
		n (%)	95% CI	n (%)	95% CI	n (%)	95% CI	n (%)	95% CI
Translation of the research question	1. Has the research question been translated correctly into search concepts	41 (0.89)	0.76-0.96	2 (0.04)	0.01-0.15	0 (0)	0-0.08	3 (0.07)	0.01-0.18
	2. Are the search concepts clear?	39 (0.85)	0.71-0.94	1 (0.02)	0-0.12	3 (0.07)	1-0.18	3 (0.07)	0.01-0.18
	3. Are there 'too many' search concepts?	42 (0.91)	0.79-0.98	1 (0.02)	0-0.12	0 (0)	0-0.08	3 (0.07)	0.01-0.18
	4. Are any of the search concepts too narrow or too broad?	40 (0.87)	0.74-0.95	0 (0)	0-0.08	3 (0.07)	1-0.18	3 (0.07)	0.01-0.18
	5. Does the search appear to retrieve too many or too few records?	43 (0.94)	0.82-0.99	0 (0)	0-0.08	0 (0)	0-0.08	3 (0.07)	0.01-0.18
Boolean and proximity operators	6. Are there any mistakes in the use of Boolean or proximity operators?	35 (0.76)	0.61-0.87	0 (0)	0-0.08	0 (0)	0-0.08	11 (0.24)	0.13-0.39
	7. Are there any mistakes in the use of nesting with brackets?	24 (0.52)	0.37-0.67	0 (0)	0-0.08	11 (0.24)	0.13-0.39	11 (0.24)	0.13-0.39
	8. If NOT is used is this likely to result in any unintended exclusions?	4 (0.09)	0.02-0.21	0 (0)	0-0.08	0 (0)	0-0.08	42 (0.91)	0.79-0.98
	9. Could precision be improved by using proximity operators instead of AND.	36 (0.78)	0.64-0.89	0 (0)	0-0.08	0 (0)	0-0.08	10 (0.22)	0.11-0.36
	10. Is the width of any proximity operators correct?	0 (0)	0-0.08	0 (0)	0-0.08	0 (0)	0-0.08	46 (1.00)	0.92-1.00
Subject headings	11. Are the subject heading relevant?	42 (0.91)	0.79-0.98	1 (0.02)	0-0.12	0 (0)	0-0.08	3 (0.07)	0.01-0.18
	12. Are subject headings missing?	28 (0.61)	0.45-0.75	15 (0.33)	0.20-0.48	0 (0)	0-0.08	3 (0.07)	0.01-0.18
	13. Are any subject heading too broad or too narrow?	40 (0.87)	0.74-0.95	3 (0.07)	0.01-0.18	0 (0)	0-0.08	3 (0.07)	0.01-0.18
	14. Are subject headings exploded where necessary and vice versa?	43 (0.94)	0.82-0.99	0 (0)	0-0.08	0 (0)	0-0.08	3 (0.07)	0.01-0.18
	15. Are sub-heading attached to subject heading?	0 (0)	0-0.08	0 (0)	0-0.08	0 (0)	0-0.08	46 (1.00)	0.92-1.00
	16. Are sub-heading used instead of relevant subject heading and vice versa?	0 (0)	0-0.08	0 (0)	0-0.08	0 (0)	0-0.08	46 (1.00)	0.92-1.00
	17. Are both subject headings and natural language terms used for each concept?	28 (0.61)	0.45-0.75	15 (0.33)	0.20-0.48	0 (0)	0-0.08	3 (0.07)	0.01-0.18
Natural language (also free-text or text-word)	18. Does the search miss any spelling variants in free-text?	42 (0.91)	0.79-0.98	2 (0.04)	0.01-0.15	0 (0)	0-0.08	2 (0.04)	0.01-0.15
	19. Does the search miss any synonyms?	18 (0.39)	0.25-0.55	25 (0.54)	0.39-0.69	0 (0)	0-0.08	3 (0.07)	0.01-0.18
	20. Does the search miss truncation or truncate at the wrong point?	11 (0.24)	0.13-0.39	1 (0.02)	0-0.12	32 (0.70)	0.54-0.82	2 (0.04)	0.01-0.15
	21. If an acronym or abbreviation is used, is the full term also included?	37 (0.80)	0.66-0.91	6 (0.13)	0.05-0.26	0 (0)	0-0.08	3 (0.07)	0.01-0.18
	22. Are apparently irrelevant or excessively broad natural language terms used?	43 (0.94)	0.82-0.99	0 (0)	0-0.08	0 (0)	0-0.08	3 (0.07)	0.01-0.18
Spelling, syntax and line numbers	23. Are there any spelling errors?	43 (0.94)	0.82-0.99	0 (0)	0-0.08	0 (0)	0-0.08	3 (0.07)	0.01-0.18
	24. Are there any errors in system syntax or wrong line numbers?	37 (0.80)	0.66-0.91	0 (0)	0-0.08	0 (0)	0-0.08	9 (0.20)	0.09-0.34
Limits and filters	25. Do any of the limits used seem unwarranted?	29 (0.63)	0.48-0.77	0 (0)	0-0.08	0 (0)	0-0.08	17 (0.37)	0.23-0.53
	26. Are any filters used appropriate for the topic?	27 (0.59)	0.43-0.73	0 (0)	0-0.08	0 (0)	0-0.08	19 (0.41)	0.27-0.57
	27. Are any potentially helpful limits or filters missing?	30 (0.65)	0.50-0.79	0 (0)	0-0.08	16 (0.35)	0.21-0.50	0 (0)	0-0.08
	28. Is starring (restrict to focus) used and if so, is there adequate justification for this?	2 (0.04)	0.01-0.15	0 (0)	0-0.08	0 (0)	0-0.08	44 (0.96)	0.85-1.00
Search strategy adaptations	29. Does the searcher indicate that the search strategy has been adapted for additional databases and/or interfaces?	37 (0.80)	0.66-0.91	0 (0)	0-0.08	2 (0.04)	0.01-0.15	7 (0.15)	0.06-0.29
	30. Are the adaptations available for review and correct?	39 (0.85)	0.71-0.94	0 (0)	0-0.08	0 (0)	0-0.08	7 (0.15)	0.06-0.29

TABLE 4

The results of quality assessment of reporting

PRISMA		Yes		Partial		No	
Items		**n (%)**	**95% CI**	**n (%)**	**95% CI**	**n (%)**	**95% CI**
Title	1. Title	43 (0.94)	0.82-0.99	1 (0.02)	0.00-0.12	2 (0.04)	0.01-0.15
Abstract	2. Structured summary	12 (0.26)	0.14-0.41	33 (0.72)	0.57-0.84	1 (0.02)	0.00-0.12
Introduction	3. Rational	46 (1.00)	0.92-1.00	0 (0.00)	0.00-0.08	0 (0.00)	0.00-0.08
	4. Objective	3 (0.07)	0.01-0.18	37 (0.80)	0.66-0.91	6 (0.13)	0.05-0.26
Methods	5. Protocol and registration	3 (0.07)	0.01-0.18	5 (0.11)	0.04-0.24	38 (0.68)	0.54-0.80
	6. Eligibility criteria	13 (0.28)	0.16-0.44	30 (0.65)	0.50-0.79	3 (0.07)	0.01-0.18
	7. Information sources	5 (0.11)	0.04-0.24	39 (0.85)	0.71-0.94	2 (0.04)	0.01-0.15
	8. Search	11 (0.24)	0.13-0.39	26 (0.57)	0.41-0.71	9 (0.20)	0.09-0.34
	9. Study selection	7 (0.15)	0.06-0.29	19 (0.41)	0.27-0.57	20 (0.44)	0.29-0.59
	10. Data collection process	10 (0.22)	0.11-0.36	19 (0.41)	0.27-0.57	17 (0.37)	0.23-0.53
	11. Data items	26 (0.57)	0.41-0.71	10 (0.22)	0.11-0.36	10 (0.22)	0.11-0.36
	12. Risk of bias in individual studies	25 (0.54)	0.39-0.69	3 (0.07)	0.01-0.18	18 (0.39)	0.25-0.55
	13. Summary measures	36 (0.78)	0.64-0.89	0 (0.00)	0.00-0.08	10 (0.22)	0.11-0.36
	14. Synthesis of results	27 (0.59)	0.43-0.73	1 (0.02)	0.00-0.12	18 (0.39)	0.25-0.55
	15. Risk of bias across studies	20 (0.44)	0.29-0.59	0 (0.00)	0.00-0.08	26 (0.57)	0.41-0.71
	16. Additional analyses	21 (0.46)	0.31-0.61	0 (0.00)	0.00-0.08	25 (0.54)	0.39-0.69
Results	17. Study selection	27 (0.59)	0.43-0.73	9 (0.20)	0.09-0.34	10 (0.22)	0.11-0.36
	18. Study characteristics	14 (0.30)	0.18-0.46	25 (0.54)	0.39-0.69	7 (0.15)	0.06-0.29
	19. Risk of bias with studies	23 (0.50)	0.35-0.65	5 (0.11)	0.04-0.24	18 (0.39)	0.25-0.55
	20. Results of individual studies	32 (0.70)	0.54-0.82	2 (0.04)	0.01-0.15	12 (0.26)	0.14-0.41
	21. Synthesis of results	24 (0.52)	0.37-0.67	3 (0.07)	0.01-0.18	19 (0.41)	0.27-0.57
	22. Risk of bias across studies	18 (0.39)	0.25-0.55	0 (0.00)	0.00-0.08	28 (0.61)	0.45-0.75
	23. Additional analyses	23 (0.50)	0.35-0.65	0 (0.00)	0.00-0.08	23 (0.50)	0.35-0.65
Discussion	24. Summary of evidence	46 (1.00)	0.92-1.00	0 (0.00)	0.00-0.08	0 (0.00)	0.00-0.08
	25. Limitations	23 (0.50)	0.35-0.65	3 (0.07)	0.01-0.18	20 (0.44)	0.29-0.59
	26. Conclusions	34 (0.74)	0.59-0.86	11 (0.24)	0.13-0.39	1 (0.02)	0.00-0.12
Funding	27. Funding	11 (0.24)	0.13-0.39	15 (0.33)	0.20-0.48	20 (0.44)	0.29-0.59

Comparison of PRISMA entries based on funding

We conducted a comparison between articles based on funding and no funding to investigate if the quality of the report is related to the funding source. Item 3 and item 24 are reported in all articles, comparisons are unnecessary. Since item 21 showed a difference, the remaining 24 entries did not show significant differences, so for most entries, we think the funding may have no significant impact on the reporting (Appendix 2-8).

The results of methodological quality assessment based on AMSTAR 2 checklists

The ICC value of AMSTAR 2 is shown in Appendix 1-5. All components of the PICO have been fully reported in research questions and inclusion criteria of most (82.61%, 38/46) SRs/MAs (item 1). Also, 82.61% (38/46) publications describe the adequate detail information of included studies' characteristics (item 8). Only 17.39%(8/46) of the literature had a report about protocol (item 2), more than 95% (44/46) of the included articles did not provide a detailed description of the study design (item 3), only 13.04% (6/46) articles show the list and rationale for the exclusion (item 7). Almost all articles did not report the source of funding for the included study in the SRs/MAs, except for one (item 10). The majority of articles (86.96%, 40/46) have incomplete reports on search strategies, lack of search methods (search terms), or search only one database, or no supplementary search (item 4). The results presented in Appendix 2-9. In order to discuss whether the quality assessment is related to the type of journal and the source of the funding, the AMSTAR entry was analyzed in the two subgroups (Appendix 2-10 and Appendix 2-11). According to the journal type, the following four items were found to have statistical significance: item4 (0.007), item8 (0.04), item11 (0.04), and item16 (0.04). The

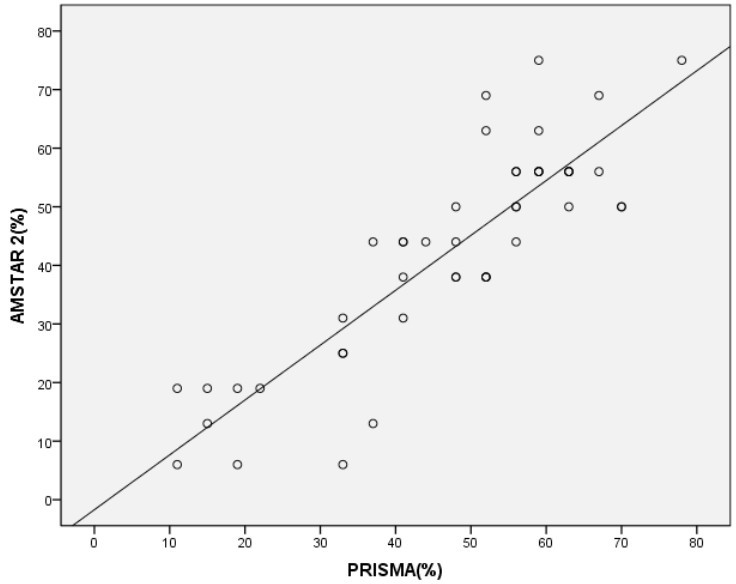


FIGURE 2. Correlation between PRISMA and AMSTAR 2.

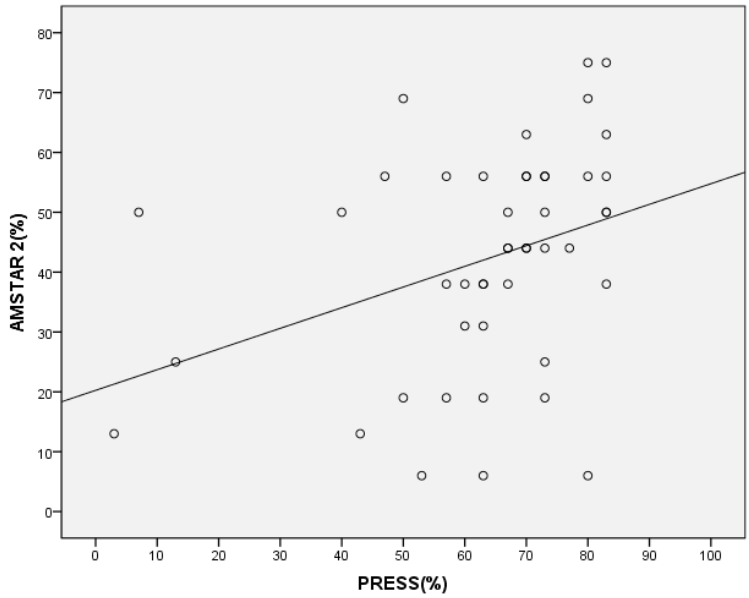


FIGURE 3. Correlation between PRESS and AMSTAR 2.

TABLE 5

Subgroup analyses of methodological quality assessment (n/%)

		Category of journals			Funding source		
items	**Yes (n = 46)**	**SCI (n = 40)**	**non-SCI (n = 6)**	***P*-value**	**funding (n = 20)**	**no-funding (n = 26)**	***P*-value**
item1	38 (0.83)	33 (0.825)	5 (0.83)	0.96	18 (0.90)	20 (0.77)	0.26
item2	8 (0.17)	6 (0.15)	2 (0.33)	0.28	2 (0.10)	6 (0.23)	0.26
item3	2 (0.04)	1 (0.025)	1 (0.17)	0.17	1 (0.05)	1 (0.04)	0.85
item4	5 (0.11)	**2 (0.05)**	**3 (0.50)**	**0.007**	2 (0.10)	3 (0.12)	0.87
item5	23 (0.50)	19 (0.475)	4 (0.67)	0.39	9 (0.45)	14 (0.54)	0.55
item6	17 (0.37)	14 (0.35)	3 (0.50)	0.48	9 (0.45)	8 (0.31)	0.32
item7	6 (0.13)	6 (0.15)	0 (0.00)	0.56	1 (0.05)	5 (0.19)	0.19
item8	38 (0.83)	**35 (0.875)**	**3 (0.50)**	**0.04**	16 (0.80)	22 (0.85)	0.68
item9	30 (0.65)	28 (0.7)	2 (0.33)	0.10	14 (0.70)	16 (0.62)	0.55
item10	1 (0.02)	1 (0.025)	0 (0.00)	0.68	0 (0.00)	1 (0.04)	0.60
item11	33 (0.72)	**31 (0.775)**	**2 (0.33)**	**0.04**	17 (0.85)	16 (0.62)	0.09
item12	21 (0.46)	20 (0.50)	1 (0.17)	0.16	9 (0.45)	12 (0.46)	0.94
item13	15 (0.33)	13 (0.325)	2 (0.33)	0.97	7 (0.35)	8 (0.31)	0.76
item14	27 (0.59)	24 (0.60)	3 (0.50)	0.64	14(0.70)	13 (0.50)	0.18
item15	20 (0.43)	18 (0.45)	2 (0.33)	0.59	10(0.50)	10 (0.38)	0.44
item16	28 (0.61)	**27 (0.675)**	**1 (0.17)**	**0.04**	14(0.70)	14 (0.54)	0.27

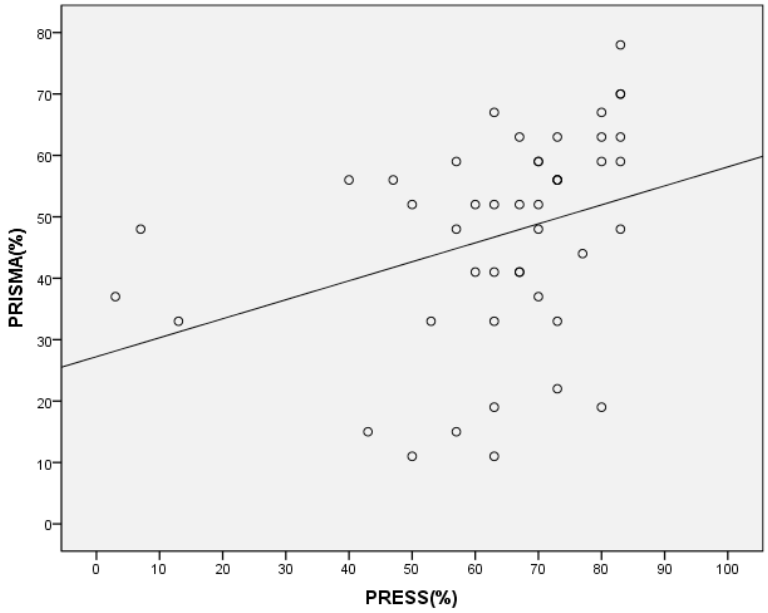


FIGURE 4. Correlation between PRESS and PRISMA.

results showed that possibly the methodological quality of SCI is better than that of non-SCI. The analysis of funding sources did not show significant differences (Tab. 5).

The correlation between PRISMA, AMSTAR 2 and PRESS
The results of the correlation analysis for AMSTAR 2 and PRISMA are shown in Fig. 2. The correlation coefficient was 0.866, and the *p*-value was less than 0.01 (the correlation was significant at the level of 0.01 (both sides)), indicating a strong correlation. The linear equation was: $y = 0.94x - 1.7$.

The linear analysis of AMSTAR 2 and PRESS (Fig. 3) showed that the correlation coefficient was 0.349, and the *p*-value was 0.017 (the correlation was significant at the level of 0.05 (both sides)), indicating that the correlation was possible but not significant. The linear equation was: $y = 0.35x + 20.26$.

The correlation coefficient obtained by linear analysis (Fig. 4) for PRISMA and PRESS was 0.338, and the *p*-value was 0.022 (the correlation was significant at the level of 0.05 (both sides)). It can be seen that there was a possible correlation between them, but the correlation was not strong. The linear equation was: $y = 0.31x + 27.23$.

Discussion

Analyses for resources of retrieval
This study analyzed the retrieval report of SRs/MAs of animal models. In terms of the number of databases, selection of different databases, there is no obvious difference between SCI and non-SCI, as well as between funding and no funding.

The database number of 14 articles was retrieving less than 3, 1 article without searching. Of all databases, PubMed was appearing in 44 articles with the most frequency. The combination of PubMed and EMBASE was used in 31 articles; more than 3 databases were retrieved in 31 articles. Half of all articles did not show the literature screening process. It can be seen that some SRs/MAs are deficient in database retrieval and reporting, which may lead to incomplete results, and affect the transparency and reliability of the results, and even reverse the outcome. Therefore, attention should be paid to a comprehensive, systematic retrieval report in the future. This paper puts forward the following specific suggestions for the retrieval method and reports for reference.

When writing SRs/MAs, it is necessary to retrieve as many databases as possible to ensure comprehensive information. According to the Cochrane Handbook for Systematic Reviews of Interventions (Higgins and Green, 2011) PubMed, EMBASE and the Cochrane library are the three databases that are often retrieved by MAs, so, at least the three databases should be selected as much as possible. It is also recommended to retrieve the professional database related to the topic of the article to ensure the credibility and advancement of the research results.

Analyses for selection process
In the literature selection process, the screening process should be reported. For example, how many people participate in this work and how to screen it? The Cochrane systematic review guide recommends that at least two people filter the retrieved literature (Higgins and Green, 2011).

Analyses for retrieving quality based on PRESS
It is important to focus on the correct translation of the research topic for the retrieval strategy when searching and ensure that the necessary concepts can be retrieved. Each concept should be retrieved simultaneously with the subject and free words, and the retrieval method (subject heading or keywords) should be explained in the report. The retrieval of synonyms is also essential. For example, entry terms in PubMed, and all the words below synonyms for the topic retrieval interface of EMBASE.com are synonyms of related subject headings, which can be selected and retrieved. If an acronym or abbreviation is used, the full term also should be included, Note the use of the spelling variant. When necessary, the appropriate qualifiers or filters can be used to make the retrieval results more accurate.

The correct retrieval strategy can reproduce the retrieval process. The Boolean operation, proximity operators, and "NOT", must be used correctly in the retrieval strategy, parentheses can be added when necessary. Generally, use cutting word symbols, to ensure the full of the same root word retrieval. It must report all database retrieval strategies or indicate that the search strategy has been adapted for additional databases, and the adaptations were available and correct.

After subgroup analysis of item conformance, it was found that the retrieval report of SCI articles was not better than non-SCI articles, and there was no significant difference based on funding analysis. This shows that the quality of the literature included in high-quality journals is indeed better, and the funding does not appear to have any significant impact on the quality of the articles, which is precisely reflecting that there is no interesting relationship with the article.

Reporting quality based on PRISMA
Our research has found that SRs/MAs of animal models have prominent flaws in reports, such as structured summary, objective, protocol and registration, eligibility criteria, information sources, search, study selection, data collection process, study characteristics, funding. For all included articles, their reporting ratios are even lower than 30%. Poor compliance with PRISMA entries may lead to incomplete and opaque reporting of information related to system reviews, thus affecting the quality of the article reports and reducing the credibility of the results. A survey of the impact of the funding source and journal type on the report shows that there is no obvious correlation between the SRs/MAs of animal model reports and these two factors. From this perspective, it seems that the credibility of the article results will not be reduced at least, because there is no need to consider the risk of bias arising from the two factors of funding source and journal types.

Methodological quality based on AMSTAR 2
Using AMSTAR2 as an evaluation tool, after assessment of all included publications, it was found that there is a lack of explanation for the following items: protocol, a detailed description for study design, the list and rationale for the exclusion, source of funding. In this study, there was almost no complete report item overall when evaluated with AMSTAR 2, and even a fairly large number of items had poor

conformity. This may be due to two reasons: First, there is a problem with the quality of the articles themselves, there may be incomplete reports or omissions; second, the use of AMSTAR2, some entries may not be suitable for evaluating some of the articles included, when in use no corresponding changes have been made to improve the applicability of the entries, and it is therefore misleading to assume that the report appears to be flawed and results in deviations from actual results.

Correlation of the three checklists

Previous studies have shown that full reports are strongly correlated with quality of research, or that PRISMA and AMSTAR 2 results are positively correlated (Tunis *et al.*, 2013; van der Pol *et al.*, 2015; Riado Minguez *et al.*, 2017). Through analyses, there is a certain correlation between the three checklists; especially the correlation between PRISMA and AMSTAR 2 is very strong. Because both AMSTAR 2 and PRISMA are evaluations of the full text, when an article has a good report quality, it shows that its content is complete, which indicates that the implementation process is described in a comprehensive manner, which reflects the rigor of methodology. On the other hand, when the methodological quality of an article is very high, many details can be reported, so the relationship between the two is closely taken for granted. The above positive relationship between PRISMA and AMSTAR 2 is only theoretical, but in fact, an SR can have a good report quality, but possible low methodological quality, because the assessment of the method quality of the SRs/MAs was based on what the authors had reported, but the actual implementation process may vary (Ge *et al.*, 2018). On the other hand, when the report quality is poor, it is also difficult to determine the methodological quality (Pieper *et al.*, 2015). PRESS is just a checklist for retrieval, and its contents can be considered as only part of the PRISMA and AMSTAR 2. Therefore, when its report quality is good, it may contribute to the improvement of the other two qualities but may be relatively small.

Advantages and limitations

This paper focuses on three checklists for the first time, and comprehensively carries out the assessments of report quality for the full text and retrieval section, methodological quality for full text, and the correlation of the three checklists was analyzed in a linear analysis; Although only EMBASE and Medline databases are searched, we still think the results of retrieval can represent most literature, because we believe that the articles contained in them are comprehensive. But due to the small number of SRs/MAs included, in later subgroup analyses, especially funding analysis, the number of no funding is only six, which may have a certain impact on the comparison results; in the evaluation of the checklist, although carefully checked, it was inevitable that there were some errors.

Conclusions

In conclusion, the report of SRs/MAs of animal models is still inadequate. To improve the quality of the retrieval report, it is suggested that it is better to use PRESS to evaluate the retrieval strategy and present retrieval. The best way to search is a combination of free words and subject words. In order to improve the overall quality of the article, the report and methodological description should be as clear and comprehensive as possible to make it more transparent and credible, and it is best to strictly adhere to the PRISMA and AMSTAR 2 checklists. Due to the correlation, if the items of each checklist are followed in the implementation process, the quality of articles will be improved to the maximum. Many articles have confirmed a correlation between PRISMA and AMSTAR theoretically, but the correlation between PRESS and PRISMA and AMSTAR remains to be further verified in the future.

Acknowledgments

We are very grateful to the editors and reviewers for their contribution to our manuscript.

Funding and Conflict of Interest

None.

Contribution

S-ZS, and ML were responsible for literature screening, W-JM., and YG for data extraction, LG, and X-PS. for data analysis, J-RW, J-HZ, and J-HT for overall planning, and everyone participated in the writing of the paper.

References

Aly SS, Zhao J, Li B, Jiang J (2014). Reliability of environmental sampling culture results using the negative binomial intraclass correlation coefficient. *Springer Plus* **3**: 40.

Baker KA, Weeks SM (2014). An overview of systematic review. *Journal of PeriAnesthesia Nursing* **29**: 454-458.

Beedie SL, Diamond AJ, Fraga LR, Figg WD, Vargesson N (2017). Vertebrate embryos as tools for anti-angiogenic drug screening and function. *Reproductive Toxicology* **70**: 49-59.

Cardona PJ, Williams A (2017). Experimental animal modelling for TB vaccine development. *International Journal of Infectious Diseases* **56**: 268-273.

Ge L, Tian JH, Li XX, Song F, Li L, Zhang J, Li G, Pei GQ, Qiu X, Yang KH (2016). Epidemiology characteristics, methodological assessment and reporting of statistical analysis of network meta-analyses in the field of cancer. *Scientific Reports* **6**: 37208.

Ge L, Tian JH, Li YN, Pan JX, Li G, Wei D, Xing X, Pan B, Chen YL, Song FJ, Yang KH (2018). Association between prospective registration and overall reporting and methodological quality of systematic reviews: a meta-epidemiological study. *Journal of Clinical Epidemiology* **93**: 45-55.

Harris JD, Quatman CE, Manring MM, Siston RA, Flanigan DC (2014). How to write a systematic review. *The American Journal of Sports Medicine* **42**: 2761-2768.

Higgins J, Green S (Eds.) (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0, The Cochrane Collaboration, London. Available at: handbook-5-1. cochrane.org/.

Li XX, Zheng Y, Chen YL, Yang KH, Zhang ZJ (2015). The reporting characteristics and methodological quality of Cochrane reviews about health policy research. *Health Policy* **119**: 503-510.

McGonigle P, Ruggeri B (2014). Animal models of human disease: challenges in enabling translation. *Biochemical Pharmacology* **87**: 162-171.

McGowan J, Sampson M, Lefebvre C (2010). An evidence based checklist for the peer review of electronic search strategies (PRESS EBC). *Evidence Based Library and Information Practice* **5**: 149-154.

Moher D, Liberati A, Tetzlaff J, Altman DG (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine* **151**: 264-269.

Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG (2007). Epidemiology and reporting characteristics of systematic reviews. *PLoS Medicine* **4**: e78.

Mueller KF, Briel M, Strech D, Meerpohl JJ, Lang B, Motschall E, Gloy V, Lamontagne F, Bassler D (2014). Dissemination bias in systematic reviews of animal research: a systematic review. *PLoS One* **9**: e116016.

Nunes DO, Marques VB, Almenara CCP, Marcarini WD, Ribeiro Júnior RF, Padilha AS (2018). Linoleic acid reduces vascular reactivity and improves the vascular dysfunction of the small mesentery in hypertension. *Journal of Nutritional Biochemistry* **62**: 18-27.

Pathuri G, Hedrick AF, Awasthi V, Cowley Jr BD, Gali H (2016). Evaluation of [^{18}F]PFH PET renography to predict future disease progression in a rat model of autosomal dominant polycystic kidney disease. *Nuclear Medicine and Biology* **43**: 1-5.

Pieper D, Buechter RB, Li L, Prediger B, Eikermann M (2015). Systematic review found AMSTAR, but not R(evised)-AMSTAR, to have good measurement properties. *Journal of Clinical Epidemiology* **68**: 574-583.

Riado Minguez D, Kowalski M, Vallve Odena M, Longin Pontzen D, Jelicic Kadic A, Jeric M, Dosenovic S, Jakus D, Vrdoljak M, Poklepovic Pericic T, Sapunar D, Puljak L (2017). Methodological and reporting quality of systematic reviews published in the highest ranking journals in the field of pain. *Anesthesia & Analgesia* **125**: 1348-1354.

Sampson M, McGowan J, Tetzlaff J, Cogo E, Moher D (2008). No consensus exists on search reporting methods for systematic reviews. *Journal of Clinical Epidemiology* **61**: 748-754.

Sena CM, Cipriano MA, Botelho MF, Seiça RM (2018). Lipoic acid prevents high-fat diet-induced hepatic steatosis in Goto Kakizaki rats by reducing oxidative stress through Nrf2 activation. *International Journal of Molecular Sciences* **19**: 2706.

Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E, Henry DA (2017). AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* **358**: j4008.

Swartz MK (2011). The PRISMA statement: a guideline for systematic reviews and meta-analyses. *Journal of Pediatric Health Care* **25**: 1-2.

Tania M, Khan MA, Xia K (2014). Recent advances in animal model experimentation in autism research. *Acta Neuropsychiatrica* **26**: 264-271.

Tunis AS, McInnes MD, Hanna R, Esmail K (2013). Association of study quality with completeness of reporting: have completeness of reporting and quality of systematic reviews and meta-analyses in major radiology journals changed since publication of the PRISMA statement? *Radiology* **269**: 413-426.

van der Pol CB, McInnes MD, Petrcich W, Tunis AS, Hanna R (2015). Is quality and completeness of reporting of systematic reviews and meta-analyses published in high impact radiology journals associated with citation rates? *PLoS One* **10**: e0119892.

Appendix 1-1 General characteristics

Category	Items	n	Proportion (95%CI)
Year (n = 46)	2003	1	0.02 (0.00-0.12)
	2005	2	0.04 (0.00-0.15)
	2007	1	0.02 (0.00-0.12)
	2008	1	0.02 (0.00-0.12)
	2009	1	0.02 (0.00-0.12)
	2011	4	0.09 (0.02-0.21)
	2013	2	0.04 (0.01-0.15)
	2014	8	0.17 (0.07-0.31)
	2015	9	0.20 (0.09-0.34)
	2016	7	0.15 (0.06-0.29)
	2017	10	0.22 (0.11-0.36)
Number of authors	≤ 2	3	0.07 (0.01-0.18)
	≥ 3	43	0.93 (0.82-0.99)
Language of database searched (n = 45)*	Only English	38	0.84 (0.71-0.94)
	Chinese + English	7	0.16 (0.07-0.30)
Presented search strategy (n = 31)	Appendix	7	0.23 (0.10-0.41)
	In the full text	24	0.77 (0.59-0.90)
Supplemental literature search (n = 28)	References	25	0.89 (0.72-0.98)
	Contact the author or relevant organization	4	0.14 (0.04-0.33)
	Conference proceedings	7	0.25 (0.11-0.45)
	Google Scholar/internet search engines	4	0.14 (0.04-0.33)
Funding (n = 46)	Yes	20	0.43 (0.29-0.59)
	No	26	0.57 (0.41-0.71)
SCI (n = 46)	Yes	40	0.87 (0.74-0.95)
	No	6	0.13 (0.05-0.26)
Impact factor(IF) (n = 40)#	≤ 2.0	2	0.05 (0.01-0.17)
	2.0 < IF ≤ 5.0	25	0.63 (0.46-0.77)
	5.0 < IF ≤ 10.0	9	0.23 (0.11-0.39)
	> 10.0	4	0.10 (0.03-0.24)
Categories of disease (n = 46)	Diseases of the nervous system	20	0.43 (0.29-0.59)
	Diseases of the circulatory system	4	0.09 (0.02-0.21)
	Diseases of the respiratory system	3	0.07 (0.01-0.18)
	Endocrine, nutritional and metabolic diseases	4	0.09 (0.02-0.21)
	others	15	0.33 (0.20-0.48)

*:1 article no database retrieved. #: Only categorize the impact factors of 40 SCI articles

Appendix 1-2 Combination of database

1 database							n	(%)
PubMed							5	10.87
2 databases								
PubMed	EMBASE						7	15.22
PubMed	web of science						2	4.35
3 databases								
PubMed	EMBASE	Cochrane library					2	4.35
PubMed	EMBASE	web of science					5	10.87
PubMed	EMBASE	BIOSIS					6	13.04
Scopus	PsycINFO	PubMed/Medline					1	2.17
PubMed	ProQuest	National Theses Database System					1	2.17
PsycINFO	EMBASE	Ovid					1	2.17
EBSCO host	PubMed/Medline	Cochrane library					1	2.17
Scopus	PubMed	BIOSIS					1	2.17
4 databases								
PubMed	web of science	CNKI	Wanfang Data				1	2.17
Scopus	CINAHL	PubMed/Medline	Web of Science				1	2.17
PubMed	EMBASE	Web of Science	BIOSIS				1	2.17
PubMed	EMBASE	Cochrane library	CINAHL				1	2.17
5 databases								
PubMed	EMBASE	Cochrane library	web of science	Library of Congress			1	2.17
PubMed	EMBASE	Cochrane library	Korean medical databases	CNKI			1	2.17
6 databases								
PubMed	EMBASE	Web of Science	CNKI	Wanfang Data	VIP		1	2.17
PubMed	EMBASE	BIOSIS	CNKI	CBM	VIP		1	2.17
PubMed	EMBASE	Cochrane library	Web of Science	AMED	CINAHL		1	2.17
7 databases								
PubMed	EMBASE	Web of Science	BIOSIS	CAB	NTIS	SIGLE	1	2.17
PubMed	EMBASE	Web of Science	CNKI	CBM	VIP	Wanfang Data	2	4.35
PubMed	EMBASE	Cochrane library	CNKI	CBM	VIP	Wanfang Data	1	2.17
Total							45	97.83

* CNKI: China National Knowledge Infrastructure CBM: China Biology Medicine

Wanfang Data: Wanfang Data Knowledge Service Platform

VIP: Database for Chinese Technical Periodicals

AMED: Allied and Alternative Medieine

NTIS: National Technical Information Service

SIGLE: System for Information on Grey Literature

CINAHL: Cumulative Index to Nursing and Allied Health Literature

Appendix 1-3 The ICC value of PRESS

item	ICC	lower	upper	item	ICC	lower	upper
item1	0.945	0.901	0.970	item16	1.000	1.000	1.000
item2	0.945	0.901	0.970	item17	0.930	0.874	0.961
item3	0.961	0.929	0.978	item18	1.000	1.000	1.000
item4	0.956	0.920	0.976	item19	0.953	0.914	0.974
item5	0.978	0.961	0.988	item20	0.993	0.987	0.996
item6	1.000	1.000	1.000	item21	1.000	1.000	1.000
item7	0.984	0.971	0.991	item22	0.958	0.924	0.977
item8	1.000	1.000	1.000	item23	1.000	1.000	1.000
item9	1.000	1.000	1.000	item24	1.000	1.000	1.000
item10	1.000	1.000	1.000	item25	1.000	1.000	1.000
item11	1.000	1.000	1.000	item26	1.000	1.000	1.000
item12	0.972	0.949	0.984	item27	0.994	0.989	0.997
item13	0.961	0.929	0.978	item28	1.000	1.000	1.000
item14	0.939	0.890	0.966	item29	0.922	0.859	0.957
item15	1.000	1.000	1.000	item30	0.954	0.917	0.975

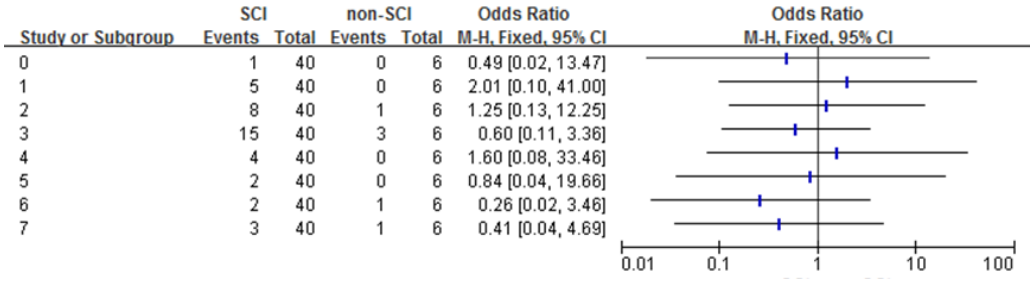
Appendix 1-4 The ICC value of PRISMA

item	ICC	lower	upper	item	ICC	lower	upper
item1	0.940	0.891	0.967	item15	0.989	0.980	0.994
item2	0.932	0.878	0.963	item16	0.972	0.949	0.984
item3	1.000	1.000	1.000	item17	0.956	0.921	0.976
item4	0.943	0.897	0.969	item18	0.975	0.955	0.986
item5	0.919	0.854	0.955	item19	0.994	0.989	0.997
item6	0.923	0.861	0.957	item20	0.986	0.975	0.992
item7	1.000	1.000	1.000	item21	0.989	0.979	0.994
item8	0.930	0.873	0.961	item22	0.989	0.980	0.994
item9	0.977	0.959	0.987	item23	0.989	0.981	0.994
item10	1.000	1.000	1.000	item24	1.000	1.000	1.000
item11	0.992	0.985	0.996	item25	0.973	0.950	0.985
item12	0.988	0.978	0.993	item26	0.975	0.955	0.986
item13	0.983	0.970	0.991	item27	0.984	0.970	0.991
item14	0.963	0.933	0.980				

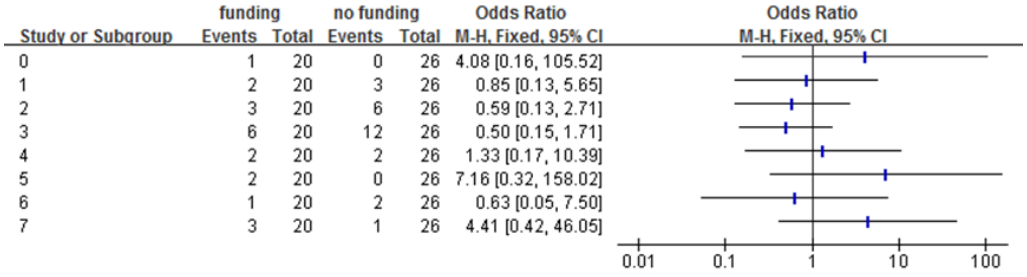
Appendix 1-5 The ICC value of AMSTAR 2

item	ICC	lower	upper	item	ICC	lower	upper
item1	0.959	0.927	0.978	item9	0.950	0.910	0.972
item2	1.000	1.000	1.000	item10	1.000	1.000	1.000
item3	1.000	1.000	1.000	item11	0.948	0.906	0.971
item4	0.958	0.924	0.977	item12	0.956	0.921	0.976
item5	0.912	0.842	0.951	item13	0.929	0.871	0.961
item6	0.977	0.958	0.987	item14	0.954	0.917	0.975
item7	1.000	1.000	1.000	item15	0.978	0.960	0.988
item8	1.000	1.000	1.000	item16	1.000	1.000	1.000

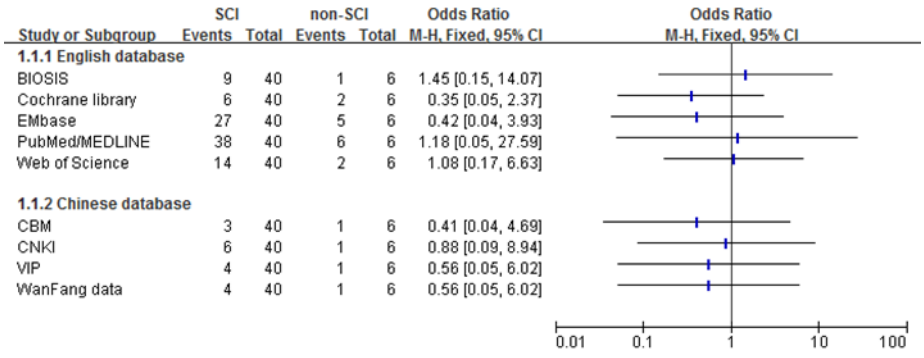
Appendix 2-1 SCI versus non-SCI papers for the databases number



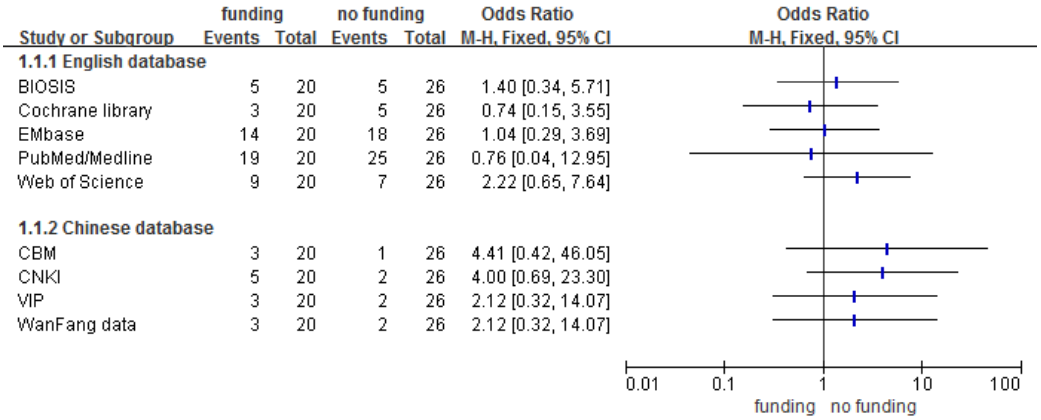
Appendix 2-2 Funding and no funding papers for the databases number



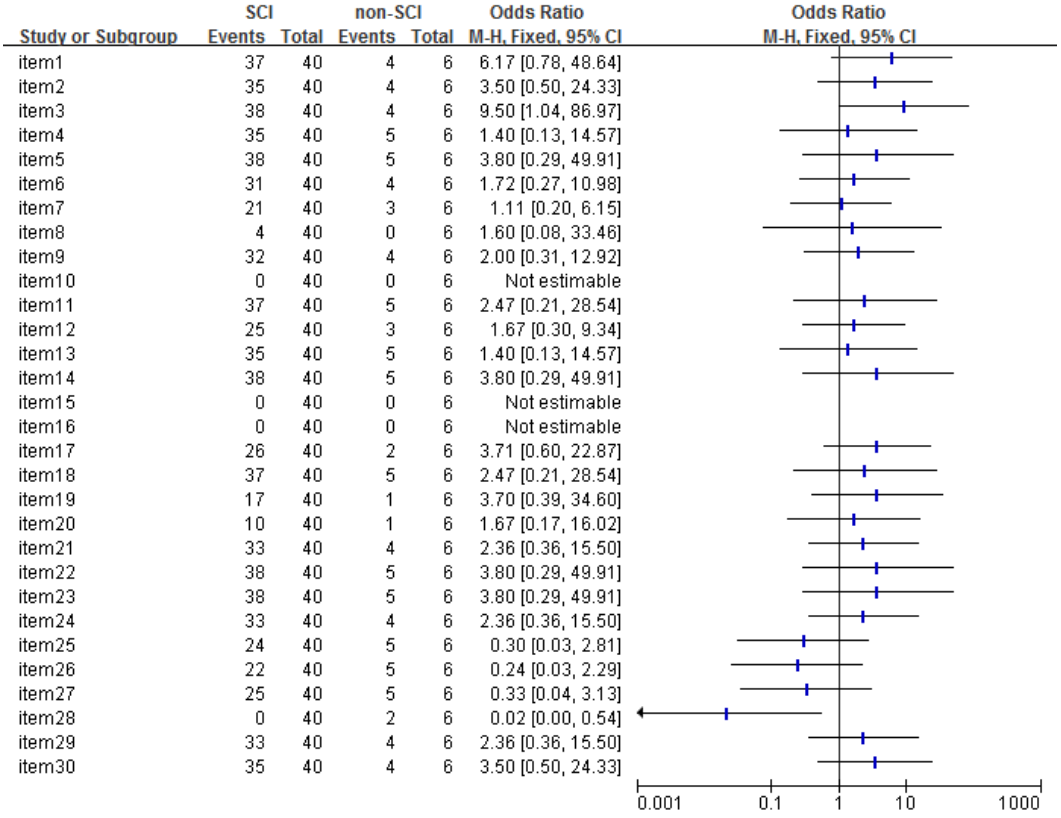
Appendix 2-3 SCI versus non-SCI papers for English or Chinese databases



Appendix 2-4 Funding versus no funding papers for English or Chinese databases

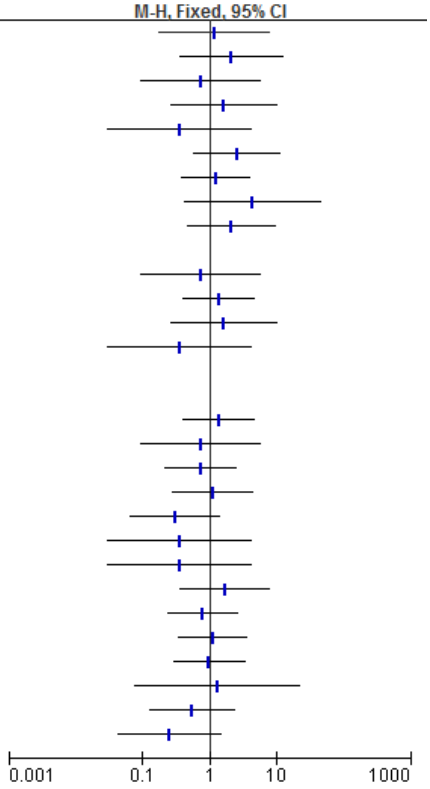


Appendix 2-5 SCI versus non-SCI papers for PRESS entries



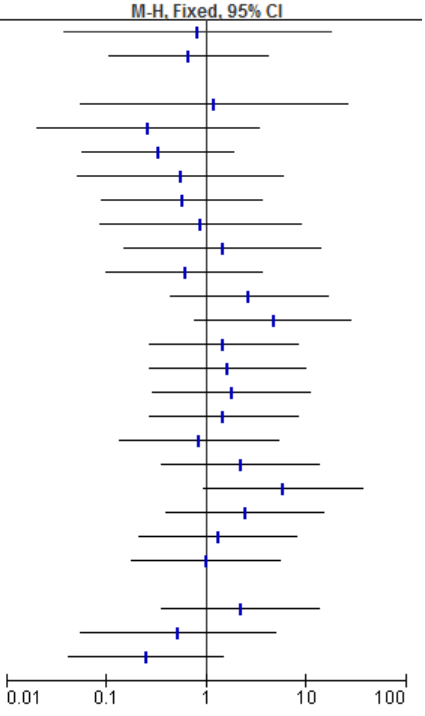
Appendix 2-6 Funding versus no funding papers for PRESS entries

Study or Subgroup	funding Events	Total	no funding Events	Total	Odds Ratio M-H, Fixed, 95% CI
item1	18	20	23	26	1.17 [0.18, 7.79]
item2	18	20	21	26	2.14 [0.37, 12.41]
item3	18	20	24	26	0.75 [0.10, 5.84]
item4	18	20	22	26	1.64 [0.27, 9.98]
item5	18	20	25	26	0.36 [0.03, 4.28]
item6	17	20	18	26	2.52 [0.57, 11.10]
item7	11	20	13	26	1.22 [0.38, 3.93]
item8	3	20	1	26	4.41 [0.42, 46.05]
item9	17	20	19	26	2.09 [0.46, 9.38]
item10	0	20	0	26	Not estimable
item11	18	20	24	26	0.75 [0.10, 5.84]
item12	13	20	15	26	1.36 [0.41, 4.54]
item13	18	20	22	26	1.64 [0.27, 9.98]
item14	18	20	25	26	0.36 [0.03, 4.28]
item15	0	20	0	26	Not estimable
item16	0	20	0	26	Not estimable
item17	13	20	15	26	1.36 [0.41, 4.54]
item18	18	20	24	26	0.75 [0.10, 5.84]
item19	7	20	11	26	0.73 [0.22, 2.45]
item20	5	20	6	26	1.11 [0.28, 4.34]
item21	14	20	23	26	0.30 [0.07, 1.42]
item22	18	20	25	26	0.36 [0.03, 4.28]
item23	18	20	25	26	0.36 [0.03, 4.28]
item24	17	20	20	26	1.70 [0.37, 7.85]
item25	12	20	17	26	0.79 [0.24, 2.65]
item26	12	20	15	26	1.10 [0.34, 3.60]
item27	13	20	17	26	0.98 [0.29, 3.34]
item28	1	20	1	26	1.32 [0.08, 22.41]
item29	15	20	22	26	0.55 [0.13, 2.37]
item30	15	20	24	26	0.25 [0.04, 1.46]

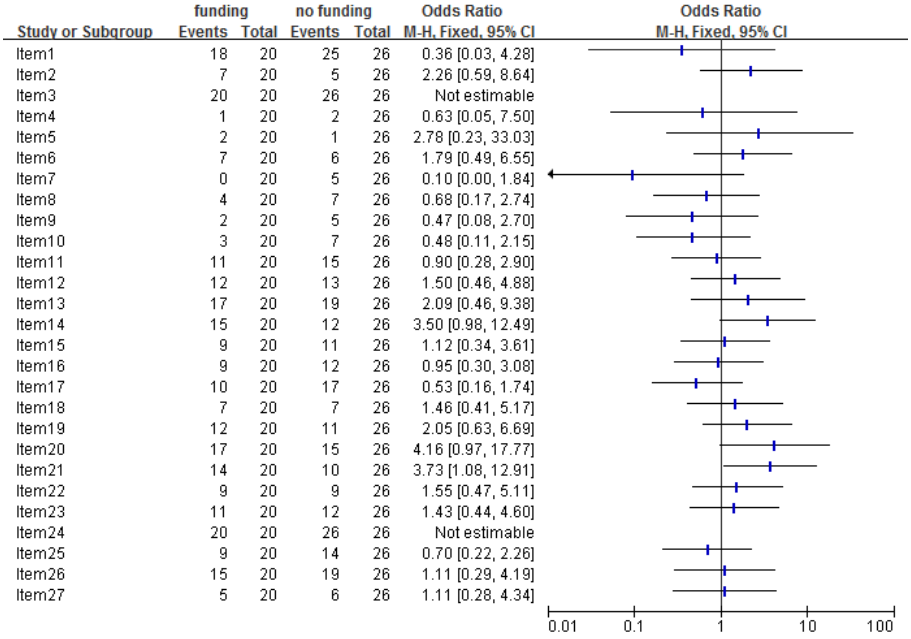


Appendix 2-7 SCI versus non-SCI papers for PRISMA entries

Study or Subgroup	SCI Events	Total	non-SCI Events	Total	Odds Ratio M-H, Fixed, 95% CI
Item1	37	40	6	6	0.82 [0.04, 17.90]
Item2	10	40	2	6	0.67 [0.11, 4.21]
Item3	40	40	6	6	Not estimable
Item4	3	40	0	6	1.21 [0.06, 26.35]
Item5	2	40	1	6	0.26 [0.02, 3.46]
Item6	10	40	3	6	0.33 [0.06, 1.92]
Item7	4	40	1	6	0.56 [0.05, 6.02]
Item8	9	40	2	6	0.58 [0.09, 3.70]
Item9	6	40	1	6	0.88 [0.09, 8.94]
Item10	9	40	1	6	1.45 [0.15, 14.07]
Item11	22	40	4	6	0.61 [0.10, 3.73]
Item12	23	40	2	6	2.71 [0.44, 16.52]
Item13	33	40	3	6	4.71 [0.78, 28.41]
Item14	24	40	3	6	1.50 [0.27, 8.38]
Item15	18	40	2	6	1.64 [0.27, 9.98]
Item16	19	40	2	6	1.81 [0.30, 11.03]
Item17	24	40	3	6	1.50 [0.27, 8.38]
Item18	12	40	2	6	0.86 [0.14, 5.33]
Item19	21	40	2	6	2.21 [0.36, 13.47]
Item20	30	40	2	6	6.00 [0.95, 37.86]
Item21	22	40	2	6	2.44 [0.40, 14.91]
Item22	16	40	2	6	1.33 [0.22, 8.16]
Item23	20	40	3	6	1.00 [0.18, 5.56]
Item24	40	40	6	6	Not estimable
Item25	21	40	2	6	2.21 [0.36, 13.47]
Item26	29	40	5	6	0.53 [0.06, 5.03]
Item27	8	40	3	6	0.25 [0.04, 1.48]

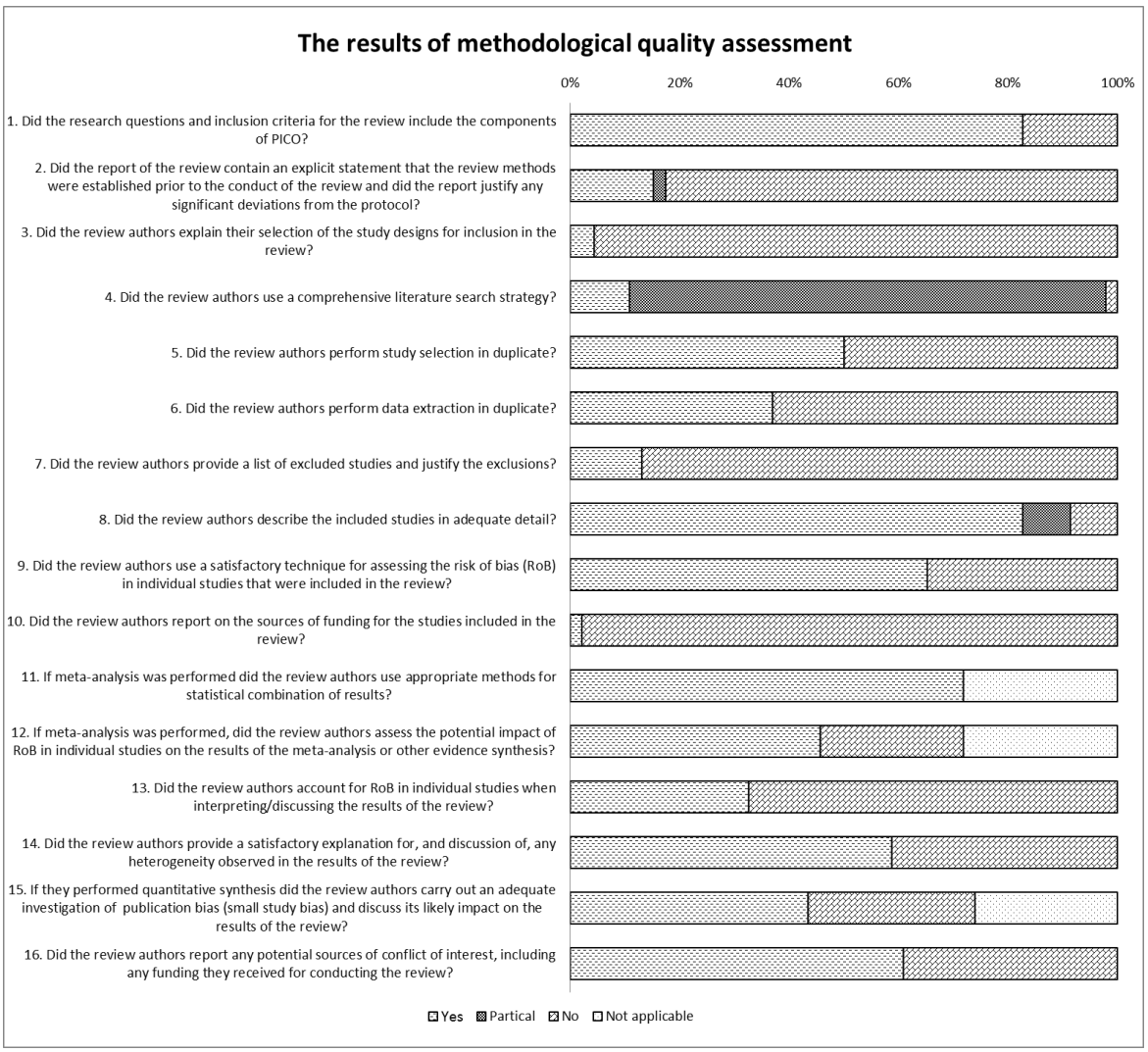


Appendix 2-8 Funding versus no funding papers for PRISMA entries

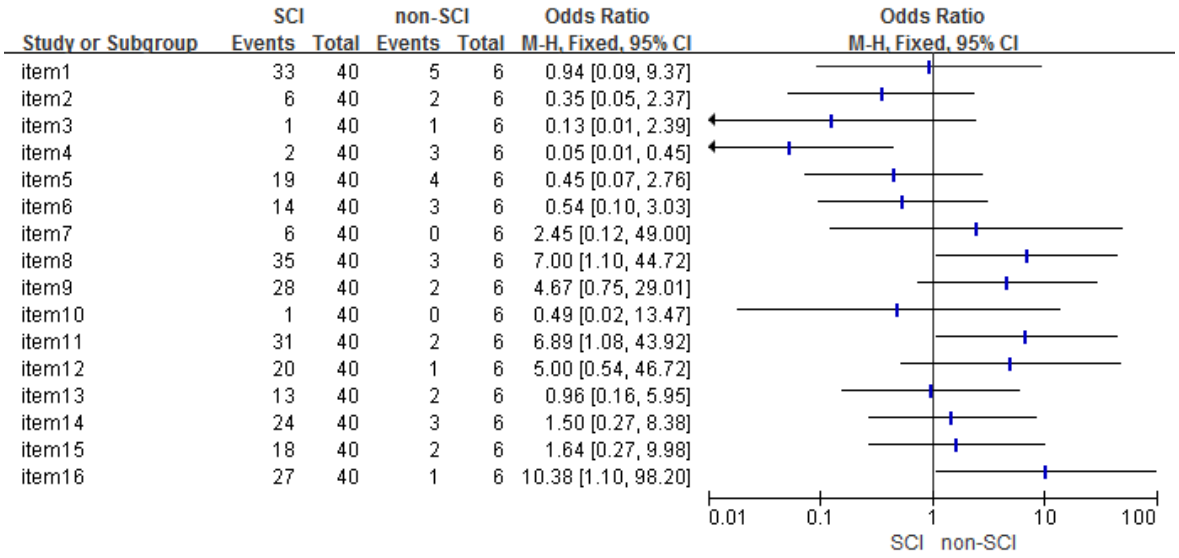


Study or Subgroup	funding Events	Total	no funding Events	Total	Odds Ratio M-H, Fixed, 95% CI
Item1	18	20	25	26	0.36 [0.03, 4.28]
Item2	7	20	5	26	2.26 [0.59, 8.64]
Item3	20	20	26	26	Not estimable
Item4	1	20	2	26	0.63 [0.05, 7.50]
Item5	2	20	1	26	2.78 [0.23, 33.03]
Item6	7	20	6	26	1.79 [0.49, 6.55]
Item7	0	20	5	26	0.10 [0.00, 1.84]
Item8	4	20	7	26	0.68 [0.17, 2.74]
Item9	2	20	5	26	0.47 [0.08, 2.70]
Item10	3	20	7	26	0.48 [0.11, 2.15]
Item11	11	20	15	26	0.90 [0.28, 2.90]
Item12	12	20	13	26	1.50 [0.46, 4.88]
Item13	17	20	19	26	2.09 [0.46, 9.38]
Item14	15	20	12	26	3.50 [0.98, 12.49]
Item15	9	20	11	26	1.12 [0.34, 3.61]
Item16	9	20	12	26	0.95 [0.30, 3.08]
Item17	10	20	17	26	0.53 [0.16, 1.74]
Item18	7	20	7	26	1.46 [0.41, 5.17]
Item19	12	20	11	26	2.05 [0.63, 6.69]
Item20	17	20	15	26	4.16 [0.97, 17.77]
Item21	14	20	10	26	3.73 [1.08, 12.91]
Item22	9	20	9	26	1.55 [0.47, 5.11]
Item23	11	20	12	26	1.43 [0.44, 4.60]
Item24	20	20	26	26	Not estimable
Item25	9	20	14	26	0.70 [0.22, 2.26]
Item26	15	20	19	26	1.11 [0.29, 4.19]
Item27	5	20	6	26	1.11 [0.28, 4.34]

Appendix 2-9 The results of methodological quality assessment

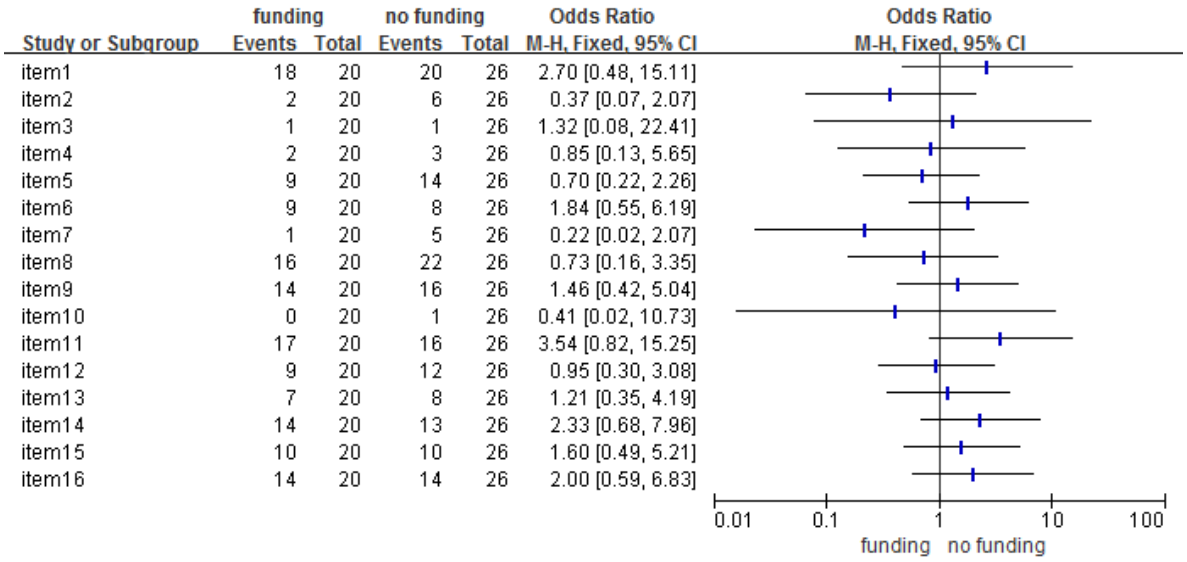


Appendix 2-10 SCI versus non-SCI papers for AMSTAR 2 entries



Study or Subgroup	SCI Events	SCI Total	non-SCI Events	non-SCI Total	Odds Ratio M-H, Fixed, 95% CI
item1	33	40	5	6	0.94 [0.09, 9.37]
item2	6	40	2	6	0.35 [0.05, 2.37]
item3	1	40	1	6	0.13 [0.01, 2.39]
item4	2	40	3	6	0.05 [0.01, 0.45]
item5	19	40	4	6	0.45 [0.07, 2.76]
item6	14	40	3	6	0.54 [0.10, 3.03]
item7	6	40	0	6	2.45 [0.12, 49.00]
item8	35	40	3	6	7.00 [1.10, 44.72]
item9	28	40	2	6	4.67 [0.75, 29.01]
item10	1	40	0	6	0.49 [0.02, 13.47]
item11	31	40	2	6	6.89 [1.08, 43.92]
item12	20	40	1	6	5.00 [0.54, 46.72]
item13	13	40	2	6	0.96 [0.16, 5.95]
item14	24	40	3	6	1.50 [0.27, 8.38]
item15	18	40	2	6	1.64 [0.27, 9.98]
item16	27	40	1	6	10.38 [1.10, 98.20]

Appendix 2-11 Funding versus no funding papers for AMSTAR 2 entries



Study or Subgroup	funding Events	funding Total	no funding Events	no funding Total	Odds Ratio M-H, Fixed, 95% CI
item1	18	20	20	26	2.70 [0.48, 15.11]
item2	2	20	6	26	0.37 [0.07, 2.07]
item3	1	20	1	26	1.32 [0.08, 22.41]
item4	2	20	3	26	0.85 [0.13, 5.65]
item5	9	20	14	26	0.70 [0.22, 2.26]
item6	9	20	8	26	1.84 [0.55, 6.19]
item7	1	20	5	26	0.22 [0.02, 2.07]
item8	16	20	22	26	0.73 [0.16, 3.35]
item9	14	20	16	26	1.46 [0.42, 5.04]
item10	0	20	1	26	0.41 [0.02, 10.73]
item11	17	20	16	26	3.54 [0.82, 15.25]
item12	9	20	12	26	0.95 [0.30, 3.08]
item13	7	20	8	26	1.21 [0.35, 4.19]
item14	14	20	13	26	2.33 [0.68, 7.96]
item15	10	20	10	26	1.60 [0.49, 5.21]
item16	14	20	14	26	2.00 [0.59, 6.83]