This article has been accepted for publication in BMJEBM following peer review.
The definitive copyedited, typeset version is available online at
10.1136/bmjebm-2024-113527

Improving peer review of systematic reviews and related review types by involving librarians and information specialists as methodological peer reviewers: a randomized controlled trial

Melissa L Rethlefsen (0000-0001-5322-9368), Sara Schroter (0000-0002-8791-8564), Lex M Bouter (0000-0002-2659-5482), Jamie J Kirkham (0000-0003-2579-9325), David Moher (0000-0003-2434-4206), David Blanco (0000-0003-2961-9328), Ana Patricia Ayala (0000-0002-3613-2270), Tara J Brigham (0000-0002-2175-4819), Holly K Grossetta Nardini (0000-0001-5563-6848), Shona Kirtley (0000-0002-7801-5777), Kate Nyhan (0000-0001-5397-2303), Whitney Townsend (0000-0002-4749-9369), Maurice P Zeegers (0000-0002-2387-083X)

Health Sciences Library & Informatics Center, University of New Mexico, MSC 09 5100, 1 University of New Mexico, Albuquerque, NM 87131-0001, Melissa L Rethlefsen, executive director
BMJ, BMA House, Tavistock Square, London, WC1H 9JR, UK, Sara Schroter, senior researcher
Amsterdam Universities Medical Centers, Department of Epidemiology and Data Science, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands, Lex M Bouter, professor emeritus
Centre for Biostatistics, The University of Manchester, Manchester Academic Health Science Centre, Manchester, United Kingdom, Jamie J Kirkham, professor
Centre of Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, 501 Smyth Road, PO BOX 201B, Ottawa, Ontario, K1H 8L6, Canada, David Moher, director
Department of Physiotherapy, Universitat Internacional de Catalunya, Barcelona, Spain, David Blanco, associate professor
Gerstein Science Information Centre, University of Toronto, 9 King's College Cr., Toronto, Ontario, M5S 1M5, Canada, Ana Patricia Ayala, research services librarian
Mayo Clinic Libraries, Mayo Clinic, 4500 San Pablo Road S., Jacksonville, FL 32223, USA, Tara J Brigham, supervising librarian
Harvey Cushing/John Hay Whitney Medical Library, Yale University and Yale New Haven Hospital, PO Box 208240, New Haven, CT 06520-8240, Holly K Grossetta Nardini, associate director
UK EQUATOR Centre, Centre for Statistics in Medicine (CSM), Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), University of Oxford, Botnar Research Centre, Windmill Road, Oxford OX3 7LD, Shona Kirtley, knowledge and information manager
Harvey Cushing/John Hay Whitney Medical Library, Yale University, P.O. Box 208014, New Haven, CT 06520-8014, Kate Nyhan, research and education librarian
Taubman Health Sciences Library, University of Michigan, 1135 Catherine St, Ann Arbor, MI 48109, Whitney Townsend, informationist
Department of Epidemiology, Care and Public Health Research Institute, Maastricht University, PO Box 616, 6200 MD, Maastricht, the Netherlands, Maurice P Zeegers, professor

Correspondence to: M L Rethlefsen, mlrethlefsen@gmail.com

Abstract

Objective: To evaluate the impact of adding librarians and information specialists (LIS) as methodological peer reviewers to the formal journal peer review process on the quality of search reporting and risk of bias in systematic review searches in the medical literature.
Design: Pragmatic two-group parallel randomized controlled trial
Setting: Three biomedical journals
Participants: Systematic reviews and related evidence synthesis manuscripts submitted to *The BMJ*, *BMJ Open*, and *BMJ Medicine* and sent out for peer review from January 3, 2023 to September 1, 2023. Randomization (allocation ratio, 1:1) was stratified by journal and used permuted blocks (block size = 4). Of 2,670 manuscripts sent to peer review during study enrollment, 400 met inclusion criteria and were randomized (62 *The BMJ*, 334 *BMJ Open*, 4 *BMJ Medicine*). 76 manuscripts were revised and resubmitted in the intervention group and 90 in the control group by January 2, 2024.
Interventions: All manuscripts followed usual journal practice for peer review, but those in the intervention group had an additional (LIS) peer reviewer invited.
Main outcome measures: The primary outcomes are the differences in first revision manuscripts between intervention and control groups in the quality of reporting and risk of bias. Quality of reporting was measured using four pre-specified PRISMA-S items. Risk of bias was measured using ROBIS Domain 2. Assessments were done in duplicate and assessors were blinded to group allocation. Secondary outcomes included differences between groups for each individual PRISMA-S and ROBIS Domain 2 item. Difference in the proportion of manuscripts rejected as the first decision post-peer review between the intervention and control groups was an additional outcome.
Results: Differences in the proportion of adequately reported searches (4.4% difference, 95% CI: -2.0% to 10.7%) and risk of bias in searches (0.5% difference, 95% CI: -13.7% to 14.6%) showed no statistically significant differences between groups. By four months post-study, 98 intervention and 70 control group manuscripts had been rejected after peer review (13.8% difference, 95% CI: 3.9% to 23.8%).
Conclusions: Inviting LIS peer reviewers did not impact adequate reporting or risk of bias of searches in first revision manuscripts of biomedical systematic reviews and related review types, though LIS peer reviewers may have contributed to a higher rate of rejection after peer review.
Trial registration: Open Science Framework: https://doi.org/10.17605/OSF.IO/W4CK2

Summary boxes

What is already known on this topic

- Literature searches are poorly reported and low quality in systematic reviews and related review types
- Methodological peer reviewers, including statisticians, as well as librarians and information specialists, bring specialized expertise to peer review
- Librarians and information specialists are underutilized as peer reviewers of systematic reviews and related review types

What this study adds

- Involving librarians and information specialists in peer review did not impact systematic review search reporting quality or risk of bias in this study
- Librarians and information specialists may influence editorial decision making when considering manuscript rejection

How this study might affect research, practice or policy

- Journal editors could consider inviting librarians and information specialists as methodological peer reviewers

Introduction

Systematic reviews (SRs) and related review types rely on highly sensitive, robust literature searches to identify available evidence to answer research questions using prespecified methods and eligibility criteria. Many systematic review guidelines recommend working with a librarian or information specialist (LIS) who is an expert on creating complex search strategies in order to reduce risk of bias [1-5]. Despite available guidance on creating high quality searches and recommendations to work with expert searchers [1-3 6 7], systematic reviews are frequently published that have used poor quality search methods [8-10]. Since the search is a critical component of systematic review methodology, these systematic reviews may be biased due to missing relevant studies [11 12].

A contributing factor may be poor quality reporting [13-18]; search strategies cannot be properly assessed for risk of bias if they are not reported well enough to assess. Reporting guidelines describe what and how to report components of systematic review search strategies for maximum transparency and reproducibility [19-21], but most systematic reviews fail to report all required search-related items [14-18]. Many published search strategies also contain errors, from simple typographical ones to missing key terminology [18 22 23]. Peer reviewers may not have the specialized methodological expertise to recognize these errors or omissions or to evaluate search strategies [24].

Inviting methodological experts, generally statisticians, is a key component of many journals' peer review processes [25-28]. Methodological reviewers can provide distinctly different comments and can contribute to improved reporting [25 29 30]. LIS, who are experts in search methods, are an underutilized resource for methodological peer review [31]. There is some preliminary evidence that LIS provide more comments on methods and recommend rejection more than other peer reviewers when peer reviewing systematic reviews [32].

Though LIS have expressed interest in being methodological peer reviewers [31], there have been no studies to test whether inviting LIS to participate as methodological peer reviewers improves the quality of reporting or reduces the risk of bias in systematic reviews. Therefore, we sought to determine the impact of inviting an LIS as an additional reviewer on the quality of reporting and risk of bias of search strategies in systematic reviews and related review types.

Methods

Overview

We used a pragmatic randomized controlled trial parallel group design. This study was declared not human subjects research (University of New Mexico HRRP 21-060) and therefore not subject to ethical review. We published the full protocol with detailed methods prior to randomization [33]. In keeping with open science best practices, we created an Open Science Framework project folder where the

registration, protocol, and other trial related information are available [34]. CONSORT for pragmatic trials was used to report the trial [35].

Eligibility criteria

Eligible manuscripts included systematic reviews and related review types newly submitted to *The BMJ*, *BMJ Open*, or *BMJ Medicine* and sent out for peer review between January 3, 2023 and September 1, 2023 (*The BMJ*, *BMJ Open*) or between May 23, 2023 and September 1, 2023 (*BMJ Medicine*). Though these journals are not necessarily representative of all biomedical research, they are general medical journals that publish studies in most medical research areas. Related reviews included scoping reviews, rapid reviews, and other methods-based review types. Manuscripts were excluded if they were a protocol or a resubmission. Protocols were excluded as most journals do not publish systematic review protocols and thus our results would be less generalizable. Additional exclusion criteria were added for *BMJ Open* during the recruitment period to accommodate *BMJ Open*'s editorial processes (see details in protocol amendments in the supplementary materials).

Study implementation

Each journal's manuscript management system was monitored (MLR) for new eligible manuscript submissions. Using computer generated [36] sequence generation and allocation concealment, we randomized each eligible manuscript to either the control or intervention group. Randomization was stratified by journal and in permuted blocks of four (1:1 allocation). The computer generated an individual four-character code for each manuscript when enrolled which was used to identify the manuscript during the evaluation process to mask assessors to the group allocation.

Manuscript authors and peer reviewers were not notified that their specific manuscripts/peer review reports were included in this study, but all submitting authors and invited peer reviewers are routinely notified that their manuscripts and reviews may be used for research purposes. Journal editors were not masked to manuscript inclusion in the study nor to the identity of peer reviewers submitting reports. The lead author (MLR) monitored all manuscripts throughout the study and used the journals' manuscript management systems to assign LIS peer reviewers in the intervention group.

Sample size

We anticipated that 5% of control group manuscripts would meet all criteria for adequate reporting [33 37] and calculated that we would need 150 manuscripts completing the first round of revisions (75 per group) to detect a 15% difference in adequately reported searches and risk of bias between groups (80% power, alpha 0.05). We needed to randomize more than 150 submissions to account for manuscripts that would not be revised. Based on the rejection rates observed throughout the first seven months of the study, we anticipated that when we had randomized 400 total submissions, we would have at least 75 first revision manuscripts per group. We thus randomized manuscripts through September 1, 2023

when we reached a total of 400 randomized manuscripts. We monitored these through January 2, 2024 to ensure a minimum of 75 first revision manuscripts per group.

Intervention

Manuscripts in the control group received peer review as usual. Intervention group manuscripts received peer review as usual, plus an LIS was invited as an additional peer reviewer. LIS peer reviewers were invited using the same standard email as all other peer reviewers and received no additional or special instructions. We invited the LIS from the Librarian Peer Reviewer Database [38], a database of LIS's who had voluntarily registered to conduct peer review on systematic reviews and other types of evidence synthesis, as part of a wider initiative to increase librarian peer review of systematic reviews. To invite individual LIS peer reviewers for each intervention manuscript, we worked through the Librarian Peer Reviewer Database list sequentially, contacting each LIS until an LIS accepted; the required number of peer reviewer report were received, regardless of whether the LIS peer reviewer had submitted their review; or ten LIS peer reviewers had already been invited. Invitations were sent through October 17, 2023.

Outcomes

The primary outcomes were differences in the proportions of adequately reported searches and searches with low risk of bias between groups in first revision manuscripts. Adequate search reporting was measured using a combination of four key PRISMA-S reporting items (items 1, 8, 9, and 13) [21], all of which must be adequately reported (Table 1). PRISMA-S is an extension to the PRISMA 2020 reporting guideline focused on transparently reporting literature searches [19 21]. These four PRISMA-S items were selected because they would be applicable to any systematic review and related review types where at least one database search was conducted to identify studies. All outcomes related to PRISMA-S and ROBIS were measured on revised manuscripts submitted by January 2, 2024.

Table 1. The four PRISMA-S items essential for adequately reporting searches in bibliographic databases, the most common information source used in systematic reviews [21]

PRISMA-S Item Number	**Section**	**PRISMA-S Item**	**Short Explanation**
1	Information Sources and Methods	Name each individual database searched, stating the platform for each.	Is the database and the associated platform for each listed in the manuscript or supplemental materials? Is it clear exactly which database(s) are searched? If multiple databases are searched on one

			platform, are all of the databases searched listed?
8	Search Strategies	Include the search strategies for each database and information source, copied and pasted exactly as run.	Are the full, copied and pasted, unadulterated search strategies from each bibliographic database searched available in the manuscript or supplemental materials? If a generic search strategy is included that is for all databases, is it clear that this search strategy could be reproduced in each database?
9	Search Strategies	Specify that no limits were used, or describe any limits or restrictions applied to a search (e.g., date or time period, language, study design) and provide justification for each use.	Does the manuscript text indicate that searches were limited using database limits? Does the manuscript explicitly state that no limits were applied to the search? Does the full search strategy for each database match the text and/or indicate that limits were applied?
13	Search Strategies	For each search strategy, provide the date when the last search occurred.	Does the manuscript or supplemental material explicitly state the date that the last search was executed in each bibliographic database?

Risk of bias was assessed using four ROBIS Domain 2 signaling questions (2.1-2.4; Table 2) [12]. The overall risk of bias was calculated based on the responses to the individual signaling questions, each of which must show low risk of bias to receive an overall low risk of bias rating.

Secondary outcomes included the difference in the proportion of papers rejected after the first round of peer review between groups and the difference in the proportion of papers requiring revisions not resubmitted after the first round of peer review. These two outcomes were measured at follow up on all randomized manuscripts. In addition, we determined the difference in the proportion of first revision manuscripts meeting each individual PRISMA-S item and assessments of low risk of bias in each individual ROBIS Domain 2 signaling question between groups.

Table 2. ROBIS Domain 2 Signaling Questions [12]

ROBIS Number	Signaling Question	Rating

2.1	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Yes / Probably Yes / Probably No / No / No Information
2.2	Were methods additional to database searching used to identify relevant reports?	Yes / Probably Yes / Probably No / No / No Information
2.3	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Yes / Probably Yes / Probably No / No / No Information
2.4	Were restrictions based on date, publication format, or language appropriate?	Yes / Probably Yes / Probably No / No / No Information

Evaluation

Seven assessors (APA, DB, TJB, SK, HKGN, KN, and WT) used an online evaluation form to assess the included first revision manuscripts. The assessors were provided with the manuscripts and supplementary materials. Manuscripts were not modified to mask names or institutions from assessors, but assessors were unaware of the group allocation. Assessors met as a group prior to the onset of the evaluation to iteratively test and develop the form and the instructions for assessors using a set of four published systematic reviews. The group began the assessment of included manuscripts by jointly assessing three of the included manuscripts to calibrate responses. Two assessors each independently evaluated the remaining manuscripts. The initial level of concordance between assessors was analyzed using Cohen's kappa. Conflicts were resolved through discussion amongst the reviewer pairs and, when needed, a third reviewer. Major conflicts of interpretation were discussed by all reviewers. The online form and the instructions for assessors are available on our Open Science Framework project site [34].

Data analysis

We analyzed differences between first revision manuscripts in the control and intervention groups by examining the proportions of adequately reported searches and low risk of bias searches in each group. We followed up on all randomized manuscripts and calculated the differences in the proportion of papers rejected after the first round of peer review (excluding withdrawn papers and papers where editors had yet not made a first decision). Similarly, we compared the papers where revisions were requested as the first decision. We examined the between groups proportion of those where the manuscript was not resubmitted.

Exploratory multivariate logistic regression was used to examine associations between the global and individual PRISMA-S and ROBIS Domain 2 ratings and potential predictors. The potential predictors included number of authors, journal, reported LIS involvement in the systematic review, LIS authorship of the systematic review, citation of the main PRISMA Statement or one of its extensions [19 39], the

citation of an alternate reporting guideline (e.g., MOOSE [40]), citation of a conduct guideline (e.g., the Cochrane Handbook [2]), or the citation of the PRESS Guideline [41].

Primary analyses were conducted using an intention-to-treat model. We conducted a per protocol sensitivity analysis for all primary and secondary outcomes comparing manuscripts with and without peer review reports completed by LIS. Data analysis was done in jamovi 1.6.23 [42]. Additional details and complete protocol amendments are available in the supplementary materials.

Patient and public involvement

To ensure readability and applicability for patients and the public, we invited a member of the public with experience as a patient peer reviewer for the BMJ Publishing Group (RH) to provide feedback on the initial study design as part of protocol development and the final manuscript.

Results

During the study recruitment period, 2,670 submissions were sent for peer review. Of those, 2,270 were excluded for not meeting eligibility criteria (see Figure 1). Four hundred eligible manuscripts were randomized; three were excluded from all analyses after randomization because they were later found to not meet the eligibility criteria. As of January 2, 2024, 166 first revision manuscripts were available for assessment, 76 from the intervention group and 90 from the control group (Table 3). A decision or revision was still pending on 27 intervention manuscripts and 41 control manuscripts. No *BMJ Medicine* manuscripts were included in the PRISMA-S or ROBIS evaluation outcomes as no manuscripts were resubmitted prior to the cut-off date for outcome assessment (January 2, 2024); 12 and 154 manuscripts were included from *BMJ* and *BMJ Open*, respectively.

Figure 1. Study Flow Diagram

[INSERT FIG. 1 HERE]

Table 3. Characteristics of manuscripts included in the PRISMA-S and ROBIS assessment (n = 166)

	Control Group N (%)	**Intervention Group N (%)**
Number of manuscripts	90 (100)	76 (100)
Journal		
BMJ Open	83 (92.2)	71 (93.4)
BMJ	7 (7.8)	5 (6.6)
LIS author	9 (10.0)	8 (10.5)
LIS involvement		

	Control Group N (%)	**Intervention Group N (%)**
Number of manuscripts	90 (100)	76 (100)
Acknowledged by name	7 (7.8)	11 (14.5)
Acknowledged, not named	12 (13.3)	10 (13.2)
Number of Authors		
Mean	6.98 (SD 3.58)	7.24 (SD 4.41)
Range	2-21	3-24
Guidance and Methods Cited		
PRISMA Guidelines	87 (96.7)	74 (97.4)
Other Reporting Guidelines	11 (12.2)	9 (11.8)
Conduct Guidelines	41 (45.6)	37 (48.7)
PRESS [41]	1 (1.1)	2 (2.6)

Primary outcomes

After consensus, two (2.2%) revised control group manuscripts adequately reported all four PRISMA-S items (Table 4), compared to five (6.6%) intervention group manuscripts (4.4% difference, 95% CI: -2.0%, 10.7%). No significant difference was detected in overall ROBIS Domain 2 outcomes (0.5% difference, 95% CI: -13.7%, 14.6%). Low risk of bias was present in 28 (31.1%) of control and 24 (31.6%) intervention manuscripts, compared to high or unclear risk of bias (control, n=62 (68.9%); intervention, n=52 (68.4%)).

Secondary outcomes

Three individual PRISMA-S items (Table 4) similarly showed no significant differences between intervention and control groups. There was a small difference favoring the intervention group in adequately reporting PRISMA-S item 13 (dates of search). All ROBIS items were similar across groups (Table 4). Initial assessor agreement in the PRISMA-S and ROBIS evaluation process ranged depending on the topic (median K=0.44; range 0.295-1.0) with agreement on limits and restrictions-related items the poorest due to differences in interpretation (see [34] for details).

To assess differences in rejection rates and decisions not to revise, all randomized manuscripts (n=397) were tracked four months beyond the primary outcome assessments. As of May 2, 2024, five manuscripts did not have a first decision and three had requested revisions pending. There were 98 manuscripts rejected as the first decision in the intervention group and 70 in the control group, a 13.8% difference (95% CI: 3.9%, 23.8%). Authors declined to revise seven manuscripts in the control group and

eight in the intervention group, which was a non-significant difference (2.4% difference, 95% CI: -4.6%, 9.4%).

Table 4. Reporting Quality and Risk of Bias in First Revision Manuscript Searches by Individual Item

Outcome	Control Group (n=90)	Intervention Group (n=76)		
	n (%)	n (%)	% Difference	% Difference 95% CI
Primary Outcomes				
PRISMA-S Global	2 (2.2)	5 (6.6)	4.4	-2.0, 10.7
ROBIS Domain 2 Global	28 (31.1)	24 (31.6)	0.5	-13.7, 14.6
Secondary Outcomes				
PRISMA-S Item 1 (Database Name)	11 (12.2)	16 (21.1)	8.9	-2.6, 20.2
PRISMA-S Item 8 (Search Strategies)	42 (46.7)	35 (46.1)	-0.6	-15.8, 14.6
PRISMA-S Item 9 (Limits and Restrictions)	46 (51.1)	36 (47.4)	-3.7	-19.0, 11.5
PRISMA-S Item 13 (Dates of Search)	33 (36.7)	40 (52.6)	16	0.96, 31.0
ROBIS Signaling Question 2.1 (Range of Sources)	74 (82.2)	62 (81.6)	-0.6	-12.4, 11.1
ROBIS Signaling Question 2.2 (Additional Methods)	71 (78.9)	57 (75.0)	-3.9	-16.8, 9.0
ROBIS Signaling Question 2.3 (Terms and Structure)	38 (42.2)	33 (43.4)	1.2	-13.9, 16.3
ROBIS Signaling Question 2.4 (Appropriateness of Restrictions)	50 (55.6)	34 (44.7)	-10.8	-26.0, 4.4

Table 4. Secondary outcomes include the differences in proportions of global and individual PRISMA-S items adequately reported and ROBIS Domain 2 items with low risk of bias between control and intervention groups. PRISMA-S items are described in Table 1. ROBIS Domain 2 Signaling Questions 2.1-2.4 are described in Table 2.

Exploratory analysis: factors predicting adequately reported searches and low risk of bias

Other factors were associated with adequate reporting, including the journal and whether there was an LIS author, using both bivariate and multivariate logistic regression (Table S1). In a multivariate model, the odds ratio for adequately reporting the search if submitted to *BMJ* versus *BMJ Open* was 10.5 (95% CI: 1.5, 95.3) and, when there was an LIS author, 20.3 (95% CI: 3.3-126.3). Other factors, including citing

any type of methods or reporting guidance, did not demonstrate an effect, perhaps due to the small number of manuscripts that were adequately reported. Only two variables were identified that were associated with low risk of bias in all four ROBIS Domain 2 signaling questions, LIS author and LIS involvement, specifically the involvement of an LIS who was named (Table S1).

Bivariate logistic regression results for the global PRISMA-S and ROBIS outcomes and results for multivariate regressions for each individual PRISMA-S and ROBIS item are available via Open Science Framework [34]. Multivariate analyses of individual PRISMA-S and ROBIS items identified similar predictive variables (Table S2).

Sensitivity analysis

LIS peer reviewers submitted reports for 62 (81.6%) intervention group manuscripts; no control group manuscripts had LIS peer reviewers. Per protocol tests on proportions of all and individual PRISMA-S and ROBIS (Table S3) items produced similar results to the intention to treat analysis, though the difference in proportion of adequate reporting of PRISMA-S Item 13 was no longer significant (13.3% difference, 95% CI: -2.6%, 29.3%). Sensitivity analysis of rejection proportions between intervention and control group manuscripts remained marginally significant (10.6% difference, 95% CI: 0.1%, 21.2%).

Discussion

We investigated whether LIS peer reviewers affected reporting quality and risk of bias using a pragmatic randomized controlled trial design. The results are statistically non-significant; LIS peer reviewers did not impact reporting quality or risk of bias in this study. However, LIS peer reviewers may help editors reject more manuscripts as a first decision. The primary factors that were associated with adequate reporting were having a librarian co-author and the journal of submission. Librarian co-authorship and acknowledging an LIS by name were associated with less risk of bias.

The only other study looking at LIS as methodological peer reviewers is a small, single journal case study, but it similarly demonstrated that LIS peer reviewers may contribute to rejection rates [32]. Few randomized controlled trials have been conducted on peer review in general, but only two have been conducted on methodological peer review, both in the journal, Medicina Clinica (Barc) [29 43]. Statisticians were the focus, and the larger trial found that the inclusion of a statistician as an additional peer reviewer improved quality [29]. Differences exist between their study and ours, namely that their trials investigated changes to manuscripts between submission and revision instead of assessing differences between groups at the revision stage [29 43]. Similarly, Ibragimova et al's case study looked at changes made to manuscripts to determine LIS peer reviewers' impact [32].

Our additional findings that LIS authors are correlated with adequate reporting and low risk of bias of searches correspond with prior research, as does our finding that involving an LIS was associated with low risk of bias [10 44-49]. Prior research has shown a difference in search quality (equivalent to ROBIS 2.3 and 2.4) between LIS mentioned in the manuscript versus in the acknowledgements without

differentiating between named LIS and generic mentions [45 50]. We chose to look at whether acknowledging LIS by name instead of as a generic mention (e.g., thanking the library) made a difference; we believe that this analysis is unique to our study. Others have shown that higher impact journals are more likely to have better search reporting [51], so finding that manuscripts under review in *The BMJ* had higher quality reporting than in *BMJ Open* is not surprising.

Limitations

Our analysis plan was based on a sample size of 75 manuscripts in each group, which we achieved, but only 62 intervention manuscripts had a peer review report completed by an LIS. It is possible that this impacted the study's results and made it difficult to detect an effect. In addition, we did not examine changes between original submission and revised manuscript or peer reviewer reports and whether suggested changes had been made. Such an analysis would reveal more nuance and could potentially highlight similar findings to Ibragimova et al [32]. Our choice of measures (PRISMA-S and ROBIS Domain 2 (2.1-2.4)) had moderate to low initial interrater reliability. ROBIS in particular is based on judgment [12 52 53], which accounts for its lower interrater reliability, but that makes it difficult to standardize.

Prior studies examining methodological peer review were implemented by journal editors [25 28 29 32 43]. Here, our intervention was conducted by an external investigator, though the peer review invitations came directly from the peer review platform. Editors were told we were conducting a trial, but not what the trial was, and were asked to continue their usual processes. Throughout the process, editors continued their usual practices and made changes that impacted the intervention (e.g., reducing the number of reviewers required after it was increased for the study, making a decision prior to receiving all reviews). In addition, because editors did not oversee the intervention, they may have placed less weight on LIS peer reviewer reports than in other methodological peer review studies. Furthermore, this study may not be generalizable to other journals or other disciplines. All three journals studied use open peer review, which may have impacted study outcomes.

Based on the number of manuscripts that required inviting multiple LIS (as many as eleven) before one accepted a review invitation, LIS peer review could potentially add time to the review process. Adding LIS as paid consultants (similar to how some journals work with statisticians) could be a way to speed up the peer review process and help increase the number of LIS-reviewed publications, but this would mean additional costs for journals.

Future research

To better understand both the feasibility of LIS peer review and the impact (or lack thereof) of LIS peer review, more research is required. Understanding the feasibility will require further investigation of the when, how, and who of accepting and completing peer review reports. For example, 89 (43.2%) of the 206 LIS invited to peer review in this study completed a peer review report. Why these individuals chose to (or declined to) accept an invitation is not understood. What is understood is the small number of available LIS peer reviewers globally [38]; this would affect any journal wishing to regularly invite LIS as

methodological peer reviewers. How journals can implement this type of review without overburdening LIS should be considered. Understanding the impact of LIS peer review could be better understood by assessing peer review reports, editor decisions, and manuscript changes throughout the peer review process, both as a follow up to this study and for new research.

Conclusions

Involving librarians and information specialists in peer review did not impact systematic review search reporting quality or risk of bias in this study. LIS may contribute to editorial decision making, especially when considering manuscript rejection, potentially acting as a safety net to help editors keep lower quality reviews out of their journals. Librarian and information specialist authorship was the only identified factor correlating with better reporting quality and lower risk of bias.

Data supplement

Contains protocol amendments, supplemental methods, and three tables

Acknowledgements

We thank Rebecca Harmston for her invaluable perspective and contributions to this manuscript. We acknowledge Shelley de Kock and Jos Kleijnen for providing us additional information on how KSR Evidence used ROBIS to assess Domain 2. Their guidance was added to our instructions for assessors. We thank all of the LIS peer reviewers in the Librarian Peer Reviewer Database for their contributions to improving search reporting and risk of bias in evidence syntheses. MLR also thanks Jim Moriarty for his help.

Ethical approval

This study was declared not human subjects research (University of New Mexico HRRP 21-060) and therefore not subject to ethical review.

Data availability statement

All anonymized data and materials from this study are available in the Open Science Framework [34].

Transparency statement

MLR affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned have been explained in the manuscript and the supplementary materials.

Contributorship

MLR: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, validation, visualization, writing – original draft, writing – review & editing; SS: Conceptualization, methodology, resources, supervision, writing – review & editing; LMB: Conceptualization, methodology, supervision, writing – review & editing; JJK: Formal analysis, methodology, supervision, writing – review & editing; DM: Methodology, supervision, writing – review & editing; DB: Formal analysis, investigation, methodology, writing – review & editing; APA: Investigation, methodology, writing – review & editing; TJB: Investigation, methodology, writing – review & editing; HKGN: Investigation, methodology, writing – review & editing; SK: Investigation, methodology, writing – review & editing; KN: Investigation, methodology, writing – review & editing; WT: Investigation, methodology, writing – review & editing; MPZ: Conceptualization, methodology, supervision, writing – review & editing

The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted, and is the guarantor.

Copyright/license for publication

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

Competing interests: All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/disclosure-of-interest/ and declare: no financial support for this specific study; MLR is the lead author of PRISMA-S and has received funding from the Medical Library Association and Baltic American Freedom Foundation to teach PRISMA-S; DM is a member of the BMJ North American advisory committee; SS is a full-time employee of BMJ Publishing Group, regularly undertakes research on the publication process but is not involved in decision-making on manuscripts; KN, HKGN, and MLR are managers of the Librarian Peer Reviewer Database; this research was conducted under a confidentiality agreement and is part of an ongoing PhD collaboration between BMJ Publishing Group and the team Meta-Research at Maastricht University (UM) on the responsible conduct of publishing scientific research; BMJ Publishing Group Ltd “BMJ ” is a private limited company incorporated in England and Wales, registration number 3102371; UM is a public legal entity in the Netherlands, registration number 50169181; this is part of MLR's self-funded BMJ/UM PhD program through Maastricht University. Authors express their own opinions and not necessarily of their funders or employers.

Role of the funding source

No specific funding was received for this study.

References

1. Aromataris E, Lockwood C, Porritt K, et al. JBI Manual for Evidence Synthesis: JBI; 2024 [cited 2024 6 May]. https://synthesismanual.jbi.global.
2. Higgins JPT, Thomas J, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023): Cochrane; 2023 [cited 2024 6 May]. www.training.cochrane.org/handbook.
3. Institute of Medicine (U.S.). Committee on Standards for Systematic Reviews of Comparative Effectiveness Research, Eden J. Finding What Works in Health Care : Standards for Systematic Reviews. Washington, D.C.: National Academies Press 2011.
4. Methods Group of the Campbell Collaboration. Methodological expectations of Campbell Collaboration intervention reviews: reporting standards: Campbell Collaboration; 2019 [updated 29 Oct; cited 2023 11 Dec]. https://dx.doi.org/10.4073/cpg.2016.
5. Relevo R, Balshem H. Finding evidence for comparing medical interventions: AHRQ and the Effective Health Care Program. *J Clin Epidemiol* 2011;64(11):1168-77. doi: 10.1016/j.jclinepi.2010.11.022
6. Kugley S, Wade A, Thomas J, et al. Searching for studies: a guide to information retrieval for Campbell systematic reviews. *Campbell Syst Rev* 2017;13(1):1-73. doi: 10.4073/cmg.2016.1
7. Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. 1.3 Undertaking the review: CRD, University of York; 2009 [cited 2024 6 May]. https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf.
8. Page MJ, Shamseer L, Altman DG, et al. Epidemiology and reporting characteristics of systematic reviews of biomedical research: A cross-sectional study. *PLoS Med* 2016;13(5):e1002028-e28. doi: 10.1371/journal.pmed.1002028
9. de Kock S, Stirk L, Ross J, et al. Systematic review search methods evaluated using the Preferred Reporting of Items for Systematic Reviews and Meta-Analyses and the Risk Of Bias In Systematic reviews tool. *Int J Technol Assess Health Care* 2020;37:e18. doi: 10.1017/S0266462320002135
10. Koffel JB. Use of recommended search strategies in systematic reviews and the impact of librarian involvement: a cross-sectional survey of recent authors. *PLoS One* 2015;10(5):e0125931. doi: 10.1371/journal.pone.0125931
11. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;358:j4008. doi: 10.1136/bmj.j4008
12. Whiting P, Savović J, Higgins JP, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016;69:225-34. doi: 10.1016/j.jclinepi.2015.06.005
13. Koffel JB, Rethlefsen ML. Reproducibility of search strategies is poor in systematic reviews published in high-impact pediatrics, cardiology and surgery journals: a cross-sectional study. *PLoS One* 2016;11(9):e0163309. doi: 10.1371/journal.pone.0163309

14. Maggio LA, Tannery NH, Kanter SL. Reproducibility of literature search reporting in medical education reviews. *Acad Med* 2011;86(8):1049-54. doi: 10.1097/ACM.0b013e31822221e7
15. Mullins MM, DeLuca JB, Crepaz N, et al. Reporting quality of search methods in systematic reviews of HIV behavioral interventions (2000-2010): are the searches clearly explained, systematic and reproducible? *Res Synth Methods* 2014;5(2):116-30. doi: 10.1002/jrsm.1098
16. Toews LC. Compliance of systematic reviews in veterinary journals with Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) literature search reporting guidelines. *J Med Libr Assoc* 2017;105(3):233-39. doi: 10.5195/jmla.2017.246
17. Yoshii A, Plaut DA, McGraw KA, et al. Analysis of the reporting of search strategies in Cochrane systematic reviews. *J Med Libr Assoc* 2009;97(1):21-9. doi: 10.3163/1536-5050.97.1.004
18. Rethlefsen ML, Brigham TJ, Price C, et al. Systematic review search strategies are poorly reported and not reproducible: a cross-sectional metaresearch study. *J Clin Epidemiol* 2023;166:111229. doi: 10.1016/j.jclinepi.2023.111229
19. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71
20. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021;372:n160. doi: 10.1136/bmj.n160
21. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Syst Rev* 2021;10(1):39. doi: 10.1186/s13643-020-01542-z
22. Salvador-Oliván JA, Marco-Cuenca G, Arquero-Avilés R. Errors in search strategies used in systematic reviews and their effects on information retrieval. *J Med Libr Assoc* 2019;107(2):210-21. doi: 10.5195/jmla.2019.567
23. Sampson M, McGowan J. Errors in search strategies were identified by type and frequency. *J Clin Epidemiol* 2006;59(10):1057-63. doi: 10.1016/j.jclinepi.2006.01.007
24. Townsend W, MacEachern M, Song J. A critical analysis of peer reviewer comments on systematic review search strategies. *Deep Blue Documents* 2024 doi: 10.7302/23162
25. Day FC, Schriger DL, Todd C, et al. The use of dedicated methodology and statistical reviewers for peer review: a content analysis of comments to authors made by methodology and regular reviewers. *Ann Emerg Med* 2002;40(3):329-33. doi: 10.1067/mem.2002.127326
26. Hardwicke TE, Goodman SN. How often do leading biomedical journals use statistical experts to evaluate statistical methods? The results of a survey. *PLoS One* 2020;15(10):e0239598. doi: 10.1371/journal.pone.0239598
27. Katz KA, Crawford GH, Lu DW, et al. Statistical reviewing policies in dermatology journals: results of a questionnaire survey of editors. *J Amer Acad Dermatol* 2004;51(2):234-40. doi: 10.1016/j.jaad.2004.02.015
28. Schriger DL, Cooper RJ, Wears RL, et al. The effect of dedicated methodology and statistical review on published manuscript quality. *Ann Emerg Med* 2002;40(3):334-7. doi: 10.1067/mem.2002.127328
29. Cobo E, Selva-O'Callagham A, Ribera JM, et al. Statistical reviewers improve reporting in biomedical articles: a randomized trial. *PLoS One* 2007;2(3):e332. doi: 10.1371/journal.pone.0000332
30. Cobo E, Cortes J, Ribera JM, et al. Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. *BMJ* 2011;343:d6783. doi: 10.1136/bmj.d6783
31. Grossetta Nardini HK, Batten J, Funaro MC, et al. Librarians as methodological peer reviewers for systematic reviews: results of an online survey. *Res Integr Peer Rev* 2019;4:23. doi: 10.1186/s41073-019-0083-5

32. Ibragimova I, Fulbright H. Librarians and information specialists as methodological peer-reviewers: a case-study of the International Journal of Health Governance. *Res Integr Peer Rev* 2024;9(1):1. doi: 10.1186/s41073-023-00142-4
33. Rethlefsen ML, Schroter S, Bouter LM, et al. Improving peer review of systematic reviews by involving librarians and information specialists: protocol for a randomized controlled trial. *Trials* 2021;22(1):791. doi: 10.1186/s13063-021-05738-z
34. Rethlefsen ML, Schroter S, Bouter LM, et al. Librarian and information specialists as methodological peer reviewers: randomized controlled trial. OSF. 2024. https://dx.doi.org/10.17605/OSF.IO/ZY547.
35. Zwarenstein M, Treweek S, Gagnier JJ, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;337:a2390. doi: 10.1136/bmj.a2390
36. Study Randomizer Wageningen, Netherlands: Phased Locked Software; 2021 https://www.studyrandomizer.com.
37. Golder S, Loke Y, McIntosh HM. Poor reporting and inadequate searches were apparent in systematic reviews of adverse effects. *J Clin Epidemiol* 2008;61(5):440-8. doi: 10.1016/j.jclinepi.2007.06.005
38. Nyhan K, Haugh D, Grossetta Nardini HK, et al. Librarian Peer Reviewer Database 2020 https://sites.google.com/view/mlprdatabase/home.
39. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6(7):e1000100. doi: 10.1371/journal.pmed.1000100
40. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;283(15):2008-12. doi: 10.1001/jama.283.15.2008
41. McGowan J, Sampson M, Salzwedel DM, et al. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. *J Clin Epidemiol* 2016;75:40-6. doi: 10.1016/j.jclinepi.2016.01.021
42. The jamovi project. Jamovi 1.6.23. 2024 [cited 2024 5 May]. https://www.jamovi.org.
43. Arnau C, Cobo E, Ribera JM, et al. [Effect of statistical review on manuscript quality in Medicina Clínica (Barcelona): a randomized study]. *Med Clin (Barc)* 2003;121(18):690-4. doi: 10.1016/s0025-7753(03)74064-0
44. Aamodt M, Huurdeman H, Strømme H. Librarian co-authored systematic reviews are associated with lower risk of bias compared to systematic reviews with acknowledgement of librarians or no participation by librarians. *Evid Based Libr Inf Pract* 2019;14(4):103-27. doi: 10.18438/eblip29601
45. Brennan E, Whitney R, Gordis T, et al. The effect of librarian collaboration on the quality of otolaryngology systematic reviews and meta-analyses. Medical Library Association. New Orleans, LA, 2022.
46. Pawliuk C, Cheng S, Zheng A, et al. Librarian involvement in systematic reviews was associated with higher quality of reported search methods: a cross-sectional survey of authors from Vancouver, BC. *J Clin Epidemiol* 2023:111237. doi: 10.1016/j.jclinepi.2023.111237
47. Rethlefsen ML, Farrell AM, Osterhaus Trzasko LC, et al. Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. *J Clin Epidemiol* 2015;68(6):617-26. doi: 10.1016/j.jclinepi.2014.11.025
48. Schellinger J, Sewell K, Bloss JE, et al. The effect of librarian involvement on the quality of systematic reviews in dental medicine. *PLoS One* 2021;16(9):e0256833. doi: 10.1371/journal.pone.0256833
49. Meert D, Torabi N, Costella J. Impact of librarians on reporting of the literature searching component of pediatric systematic reviews. *J Med Libr Assoc* 2016;104(4):267-77. doi: 10.3163/1536-5050.104.4.004

50. Whitney R, Shih MC, Gordis T, et al. Effect of librarian collaboration on otolaryngology systematic review and meta-analysis quality. *J Med Libr Assoc* 2024;112(3):261-74. doi: 10.5195/jmla.2024.1774
51. Faggion CM, Jr., Huivin R, Aranda L, et al. The search and selection for primary studies in systematic reviews published in dental journals indexed in MEDLINE was not fully reproducible. *J Clin Epidemiol* 2018;98:53-61. doi: 10.1016/j.jclinepi.2018.02.011
52. Pereira AG, Martins CC, Campos JR, et al. Critical appraisal of systematic reviews of intervention studies in periodontology using AMSTAR 2 and ROBIS tools. *J Clin Exp Dent* 2023;15(8):e678-e94. doi: 10.4317/jced.60197
53. Pieper D, Puljak L, González-Lorenzo M, et al. Minor differences were found between AMSTAR 2 and ROBIS in the assessment of systematic reviews including both randomized and nonrandomized studies. *J Clin Epidemiol* 2019;108:26-33. doi: 10.1016/j.jclinepi.2018.12.004