

METHODOLOGY

Open Access



Publication bias in otorhinolaryngology meta-analyses in 2021

Fatemeh Mohammadian¹, Shahin Bastaninejad² and Shirin Irani^{2*}

Abstract

Introduction One concern in meta-analyses is the presence of publication bias (PB) which leads to the dissemination of inflated results. In this study, we assessed how much the meta-analyses in the field of otorhinolaryngology in 2021 evaluated the presence of PB.

Methods Six of the most influential journals in the field were selected. A search was conducted, and data were extracted from the included studies. In cases where PB was not assessed by the authors, we evaluated the risk of its presence by designing funnel plots and performing statistical tests.

Results Seventy-five systematic reviews were included. Fifty-one percent of them used at least one method for assessing the risk of PB, with the visual inspection of a funnel plot being the most frequent method used. Twenty-nine percent of the studies reported a high risk of PB presence. We replicated the results of 11 meta-analyses that did not assess the risk of PB and found that 63.6% were at high risk. We also found that a considerable proportion of the systematic reviews that found a high risk of PB did not take it into consideration when making conclusions and discussing their results.

Discussion Our results indicate that systematic reviews published in some of the most influential journals in the field do not implement enough measures in their search strategies to reduce the risk of PB, nor do they assess the risk of its presence or take the risk of its presence into consideration when inferring their results.

Background

Rationale

The Catalogue of Bias, a collaboration dedicated to describing a wide range of biases and outlining their potential impact on research studies, defines publication bias (PB) as “when the likelihood of a study being published is affected by the findings of the study” [1]. Several factors could lead to PB including selective publication of positive findings, selective publication of statistically

significant findings, selective publication of “interesting” findings, and publication according to the quality of the trial or its funding [2]. A study in 2009 demonstrated that randomized controlled trials with positive findings were 3.9 times more likely to be submitted and published than trials with negative or null findings [3]. The OPEN project (Overcome failure to Publish nEgative fINdings), a project funded by the European Union to investigate the extent and impact of dissemination bias, has encouraged systematic reviewers (SRs) to follow “the best practices” in conducting systematic reviews (SR) (especially practices concerning the assessment of the impact of dissemination bias) and publishing the protocol and results of their SRs publicly. Regarding “the best practices” in conducting SRs, the members of the OPEN project have proposed the Cochrane Handbook for Systematic Reviews of Interventions [4] and the standards for SRs stated by

*Correspondence:

Shirin Irani

sh_irani@razi.tums.ac.ir

¹ Department of Medicine, Tehran University of Medical Sciences, Keshavarz Blvd, Tehran, Iran

² Otorhinolaryngology Research Center, Amiralam Hospital, Tehran University of Medical Sciences, Keshavarz Blvd, Tehran, Iran



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

the Institute of Medicine of the National Academy of Sciences [5]. The main recommendation in these guidelines for avoiding PB is to search for evidence that includes bibliographic databases and other sources such as grey literature, citation indexes, trial data, and other unpublished reports. On the other hand, selection models and graph-based methods have been suggested to be the main methods for assessing the presence of PB [6].

Unfortunately, despite the significance of PB, previous studies show that a considerable proportion of systematic reviewers in different healthcare fields, such as oncology [7], anesthesiology [8], dermatology [9], cardiology [10], and gastroenterology [11], did not try to evaluate its possible presence in their SRs and MAs. Also, a substantial proportion of SRs were found not to search resources other than published materials, hence increasing the risk of PB in their results [12].

Unfortunately, previous reports show that methodological research biases, including biases in conducting randomized trials or gender bias in research and publishing, in the field of otorhinolaryngology (ENT) are still pretty common, and not many improvements have been made in the last few decades despite the significant growth of research methodology awareness in the scientific community [13, 14]. In this study, however, we aimed to assess another form of bias in the field, namely, to what extent the recent SRs with MAs in the field of ENT have taken measures to reduce the risk of PB in their results or have evaluated the probability of its presence in their research. To our knowledge, this is the first study that has evaluated this subject in this field. Our findings may help to understand how much the issue of PB is addressed in the field.

Approaches to deal with publication bias

There are two approaches for dealing with PB: selection models and graph-based methods [6]. Selection models use the weighted distribution theory to model the publication process and, thus, develop estimation procedures that account for the selection process [15]. The Hedges model [16] is the first and most known selection model used for assessing PB. Unfortunately, these models rely on largely untestable assumptions [17] and, thus, are rarely used in practice for assessing PB. Instead, they are used in sensitivity analyses [15]. On the other hand, graph-based methods are widely used. These methods are based on a funnel plot which usually presents effect sizes plotted against their standard errors (SE) or precision [18]. In the presence of PB, the plot is expected to be asymmetrical. However, a variety of other factors may also lead to asymmetrical funnel plots, such as inflated effects in smaller studies, true heterogeneity, artefactual effects, and chance [19]. Another issue with the visual

inspection of funnel plots for detecting PB is the subjectivity of the approach, which in turn leads to errors in interpretation [20]. Thus, statistical tests have also been proposed to detect funnel plot asymmetries, such as Begg's rank test [21], Egger's regression [19], Harbord's regression [22], Peters' regression [23], and Deeks' regression [24]. Trim-and-fill method, a nonparametric rank-based correction method, was proposed to recover symmetry by "trimming" observed studies and subsequently imputing missing studies [25].

It has been proposed that reviewers use various tests and methods to detect PB in their research because different tests make different assumptions on the association between the effect sizes and their precision measures [26]. The handbook of Cochrane [27] has also made some other recommendations:

- Tests should be used only when there are at least 10 studies included in the meta-analysis.
- Tests should not be used if studies have similar SE of effect estimates.
- Results of tests should be interpreted in the light of visual inspection of the funnel plot.
- When there is evidence of plot asymmetry from a test, PB should be considered the only one of several possible explanations.

Objectives

To what extent the MAs in the field of ENT in 2021 have searched for evidence other than bibliographic databases and assessed the risk of PB?

Methods

SR selection

Due to limited resources (mostly reviewers), only six journals were included in this research. These six journals were selected from the ten journals with the highest impact factor in the field of ENT (discovered through Google Scholar) following a consensus between a group of four attending ENT surgeons at the Tehran University of Medical Sciences. The attending ENT surgeons that selected these journals were blinded to the aim of the study and were told to select six of those ten journals that they believe to (1) have been the most influential in the field in the past decade and (2) are known to them and their peers as publishing some of SRs with the highest impact on their clinical practice. Although this process had the potential to introduce some selection bias, we figured that it might be the best tactic for assessing the presence of PB in some of the most influential SRs in the field, given our limited review resources. The selected journals included the following: *Audiology*

and *Neurotology, Ear & Hearing, International Forum of Allergy & Rhinology, Otolaryngology-Head and Neck Surgery, Rhinology, and The Laryngoscope*. PubMed was searched for the papers published in these journals in the year 2021. The “Systematic Review” and “Meta-Analysis” filters of PubMed were activated to narrow the search results. Also, those journals were hand-searched to make sure that no related studies were missed in our search. The results of the search were imported into the Mendeley Reference Manager application, a software designed for the management of citations. Then, using the “Look up metadata by DOI” feature of the application, all the metadata of the records were updated.

Two reviewers retrieved the full texts of the records and independently assessed them for eligibility. Discrepancies were resolved through discussion. Studies that had an MA component were included.

Data items

Two reviewers independently extracted the data from eligible studies. Discrepancies were resolved through discussion. The following data were gathered:

- Study characteristics
- Reporting guideline
- Bibliographic databases and citation indexes searched
- Other sources of data (clinical trials registries, grey literature, citation checking, etc.)
- Whether studies in a language other than English were considered eligible for inclusion
- The number of studies included in the SR
- If assessed, the method used to assess PB
- Number of papers that cited the SR (as a measure of the influence the SR had in the field)
- Review type

The review type was determined according to the 10 categories proposed by a typology study [28]: effectiveness, experiential, costs/economic evaluation, prevalence and/or incidence, diagnostic test accuracy (DTA), etiology and/or risk, expert opinion/policy, psychometric, prognostic, and methodology. In cases where one SR fell into more than one category, we considered the main objective of the SR to determine the review type.

Publication bias assessment

We recorded the methods used for assessing PB in each SR. It must be noted that PB is an issue at the outcome level rather than the study level. In this study, we only assessed the primary outcome in each SR for PB. For SRs that did not evaluate the risk of PB, we assessed it by replicating the results of their MAs (using the reported effect measures for each included study), designing funnel

plots, using the trim-and-fill method, and performing Begg’s rank test and Egger’s regression (or Deeks’ regression in case of DTA SRs). Only SRs that included at least 10 studies in their MA were assessed in this way because when there are fewer than 10 studies, the power of the statistical tests for assessing the risk of PB is relatively low [4].

We used R version 4.2 [29] “meta” package [30] to perform analyses and assess PB in SRs where it was not assessed. The significance level for PB tests was set at $P < 0.10$, as it is the commonly used threshold for such tests due to their low statistical power.

Results

Study flow

The search on PubMed returned 188 records. After updating the metadata, 57 records were removed as they were not published in 2021. After screening the full-text papers, 56 reports were also excluded. Especially, an SR that was conducted on individual patient data (IPD) was excluded because, although these studies are also at risk of PB [31], the unit of analysis in our research is a study and not an individual patient. Finally, 75 SRs were included in this study (check the [Appendix](#) for the characteristics of the included studies). Figure 1 shows the flow of study selection.

Basic characteristics of contributing SRs

The basic characteristics of the included SRs are presented in Table 1.

The SR types were as follows: 34 effectiveness, 24 etiology and/or risk, 10 prevalence and/or incidence, 5 prognostic, and 2 DTA. Most of the SRs followed PRISMA [32] or MOOSE [33] reporting guidelines. Only 9 SRs did not report the use of a standard reporting guideline.

Included SRs were cited by other papers on a median of 4 citations per SR (interquartile range (IQR): 1–8). The highest number of citations to an SR was 107 [34], while 9 SRs were not cited by any other paper at all.

Literature search

Bibliographic databases and citation indexes

On average, 3.8 bibliographic databases and citation indexes were searched in the included SRs. The data sources searched included APA PsychINFO, CAB Abstracts, CINAHL, CNKI, Cochrane Library, CQVIP, Embase, EmCare, Google Scholar, LILACS, MEDLINE, OTseeker, ProQuest, PubMed, SciELO, ScienceDirect, Scopus, Wanfang Data, and Web of Science. PubMed was the most searched database with 56 SRs (74.7%) searching it, followed by Embase which was searched by 51 SRs (68.0%). Figure 2 shows how many SRs went through searching each of these databases.

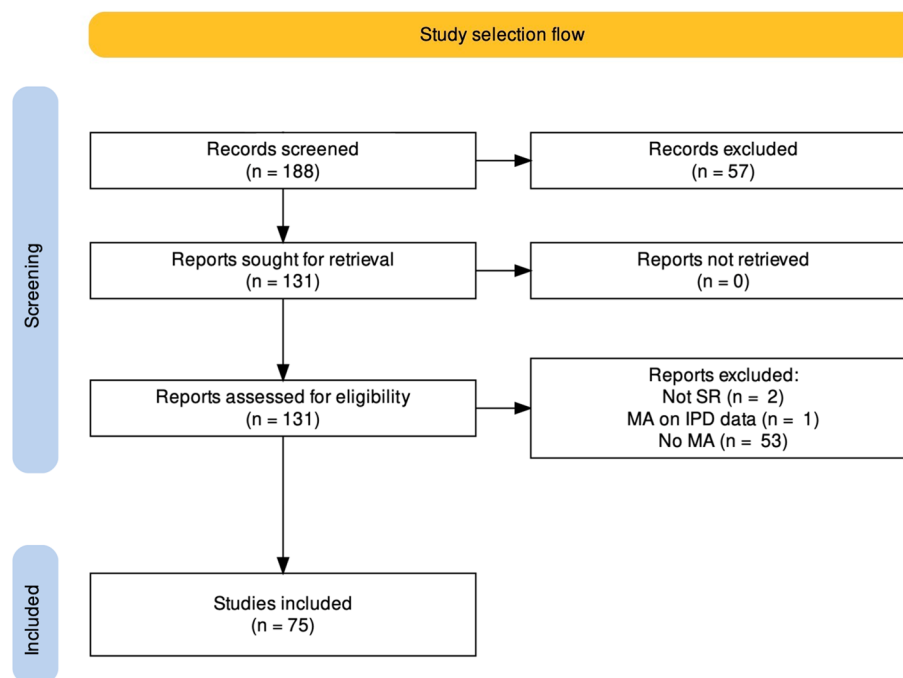


Fig. 1 The flow of study selection

Table 1 Basic characteristics of contributing studies. *SR* systematic reviews

Journal	# SRs	# Citations			
<i>Audiology and Neurotology</i>	1	0			
<i>Ear & Hearing</i>	2	27			
<i>International Forum of Allergy & Rhinology</i>	6	15			
<i>Otolaryngology-Head and Neck Surgery</i>	22	208			
<i>Rhinology journal</i>	7	16			
<i>The Laryngoscope</i>	37	394			
SR type	Effectiveness	Etiology and/or risk	Prevalence and/or incidence	Prognostic	Diagnostic test accuracy
	34	24	10	5	2
Reporting guideline used	Yes		No		
	66		9		

Other sources and languages

Backward citation checking was the most prevalent technique used for finding other potentially eligible records (50 SRs, 66.7%). On the other hand, forward citation checking was rarely used (4 SRs, 5.3%). Also, 3 SRs (4.0%) hand-searched the relevant journals or contacted experts in the field. Finally, 1 SR used the “related articles” feature of PubMed to search for other studies.

Search for trial registers and trial results registers was also rarely used (10 SRs, 13.3%). The registries searched were ClinicalTrials.gov (9 SRs) and ICTRP (International Clinical Trials Registry Platform) (1 SR).

The search for unpublished studies and grey literature was also less than satisfactory. One study searched for the abstracts from the related conferences, one study searched for theses (using ProQuest), one study searched the National Rehabilitation Information Center, and one study searched the Chulalongkorn Medical Library.

Most of the included SRs also used language restrictions in their search strategy. Only 18 SRs (24.0%) searched for studies in languages other than English, while 6 SRs did not report if they utilized any language restrictions in their search.

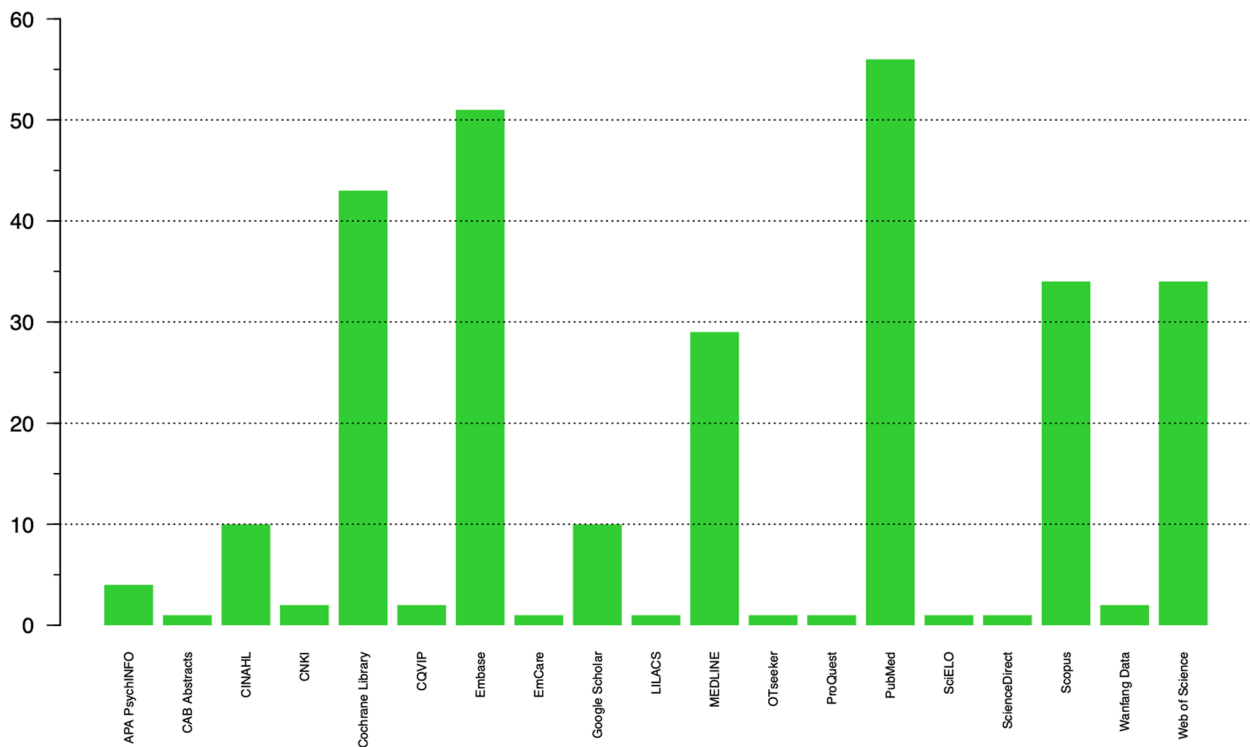


Fig. 2 Number of included systematic reviews (SRs) that searched each bibliographic database and citation index

Included studies

On average, 24.8 studies were included in the SRs (median 16, IQR 10.5–30.5). The maximum number of studies included in an SR was 98, while the minimum was 3. For MAs, an average number of 21.9 studies were used (median 13, IQR 7.5–26.5). The maximum and the minimum number of studies included in an MA were similar to the aforementioned numbers.

Publication bias

Assessment for publication bias

Almost half of the included SRs used at least one method to assess the potential risk of PB (38/75, 50.7%), while the other SRs failed to acknowledge PB at all. Figure 3 shows the ratio of SRs that assessed PB per journal.

Visual inspection of the funnel plot was the most used method for assessing PB across the included SRs (34/38, 89.5%). Among tests used to assess funnel plot asymmetry, Egger’s regression was the most widely used (25/38, 65.8%), followed by Begg’s rank test (4/38, 10.6%). Two SRs used Harbord’s regression and one used Peters’ regression. Finally, 4 SRs used the trim-and-fill method to assess the effect of missing studies.

Of 38 SRs that assessed PB, 26 (68.4%) used at least two of the mentioned methods, while 6 (15.8%) used at least three methods. Also, 4 SRs did not design a funnel plot

but relied solely on statistical tests for funnel plot asymmetry to assess the potential risk of PB. This finding was interesting because it is always advised that the results of tests for funnel plot asymmetry must be interpreted in the light of visual inspection of the plot [4]. Also, both included DTA SRs used Egger’s regression test, instead of Deeks’ regression, to assess funnel plot asymmetry.

Presence of publication bias

Of the 38 SRs that assessed PB, 11 (28.9%) reported a considerable risk for its presence. We tried to assess the risk of PB in the remaining 37 SRs. In the process, we found that 22 SRs had included less than 10 studies in their MAs for any single outcome and, thus, were ineligible for PB risk assessment through statistical tests. Furthermore, replicating the MA for 4 SRs was impossible mostly due to an incomplete report of the data or statistical methods used. Eventually, we managed to assess the risk of PB in 11 additional SRs. Considerably, 7 of these SRs (63.6%) were found to be at a considerable risk of PB. Overall, 49 SRs were assessed for the risk of PB, out of which 18 (36.7%) were found to be at considerable risk.

We also examined the SRs that assessed the presence of PB in their review and found a high risk for its presence, in more depth to check if they took measures to assess its impact on their MA results, or if their

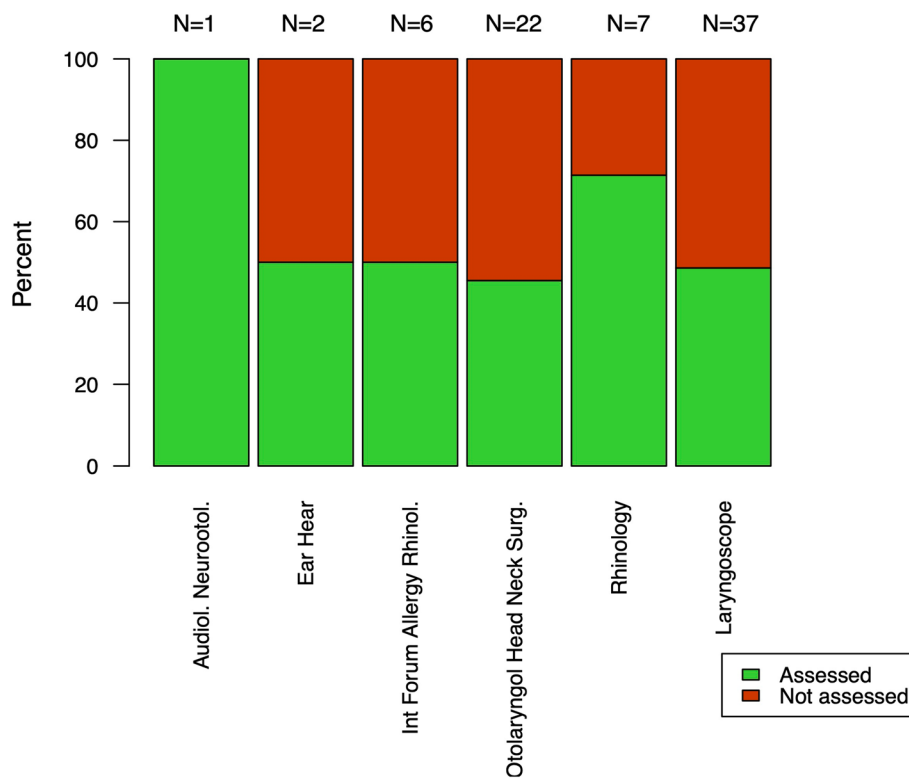


Fig. 3 Ratio of included systematic reviews (SRs) that assessed publication bias per journal

conclusions were made taking PB into consideration. Of the 11 SRs that reported a high risk of PB in their included studies, 3 (27.2%) used further tests to estimate intervention effects “corrected” for the effects of PB, such as the trim-and-fill method or conducting sensitivity analyses. Also, only 7 of those 11 studies (63.3%) that found a high risk of PB took this risk into consideration when making conclusions, by either downgrading the certainty of the evidence or discussing the potential impact that PB might have on their results.

Possible factors contributing to the risk of publication bias

To test for the possible factors that may have affected the risk of PB, a series of Pearson’s chi-square tests and *t*-tests were conducted. However, none of these tests were found to indicate a possible statistically significant correlation. The results of these tests are provided in Table 2.

As these results indicate, the chances of PB were slightly altered following the inclusion of languages other than English and sources other than bibliographic databases in the search strategy, but none of these correlations were statistically significant. The number of bibliographic databases and citation indexes searched

Table 2 The possible sources of publication bias. OR odds ratio

Factor		Risk for publication bias		OR	p Value
		High risk	Low risk		
Language restriction	Only English	12	20	1.35	0.899
	Other languages too	4	9		
Sources other than bibliographic databases (citation searching not included)	Searched	3	6	0.83	0.922
	Not searched	15	25		
Sources other than bibliographic databases (citation searching included)	Searched	13	22	1.06	0.976
	Not searched	5	9		
Number of databases searched	Mean (SD)	3.72 (0.89)	3.84 (1.13)	-	0.692

was also not different in a statistically significant manner between the studies with low risk of PB vs. those at high risk of PB.

Discussion

Interpretation of results

In this study, we aimed to evaluate to what extent the SRs and MAs published in some of the most influential journals of ENT use methods to reduce the risk of PB and different techniques to assess the risk of its presence. Our findings revealed that this issue is not addressed optimally in a considerable proportion of the SRs.

First, the search strategies used in these SRs were not comprehensive enough to mitigate the risk of PB. Most SRs restricted their search to papers published in English, thus suffering from a great risk of language bias. Although previous studies have shown that the impact of language bias is negligible on the results of an SR in most circumstances [35–37], exceptions have also been observed [38–40]. As a result, Cochrane recommends that language restrictions should not be used unless in the setting of rapid reviews, and even in that setting, its use should be justified by the reviewers [4]. Also, most of the SRs did not search for other sources of data or grey literature. This issue is of great importance as it has been found that such data can seriously affect the results of an SR [41, 42]. Specifically, including a grey literature search should be seriously considered when conducting an SR because an association between “statistically significant” results and publication has been documented in previous studies [4].

Another finding of interest was that almost half of the SRs did not assess the risk of PB. This finding becomes bolder knowing that our analyses revealed that the risk of PB was considerably higher in the SRs that did not assess the risk of its presence (63.6% vs 28.9%). The reason behind this phenomenon is unknown, but some of the potential reasons could be as follows: (a) reviewers trying not to downgrade the confidence in their results; (b) lack of methodological expertise and knowledge for assessing the risk of PB which also resulted in designing poor search strategies; and (c) solely due to chance. Nevertheless, the journal editors and reviewers should ask the authors to assess the risk of PB in their SRs whenever feasible.

More importantly, we saw that in a lot of the cases where reviewers found a high risk for PB in their SRs, they did not try to estimate the intervention effect corrected for the impact of PB, take the risk of PB in making conclusions, or expand their search to reduce the risk of PB presence. This issue should be specifically noted by journal editors, asking the authors to include other sources of data as well when the risk of PB was assessed

to be high, in an attempt to avoid publishing inflated results as much as possible.

Another finding of interest was the inappropriate use of methods to assess the risk of PB. Although this problem was not frequent across the SRs, some used inappropriate tests to assess funnel plot asymmetry, such as using Egger’s regression test instead of Deeks’ regression in the setting of DTA SRs or using statistical tests alone with no visual inspection of the funnel plot beforehand. Both journal editors and reviewers must note that the results of statistical tests for funnel plot asymmetry should be interpreted in light of the visual inspection of the plot, as all these tests are known to have low statistical power [27]. Other factors should also be considered for using such tests, such as the fact that they are not recommended for cases when there are less than 10 studies included in the MA or that they should not be used when studies are of similar size [27]. Using contour-enhanced funnel plots is also highly desirable as they help with differentiating the reasons for funnel plot asymmetry [43].

Finally, we assessed some possible factors that might have contributed to the risk of PB presence. Surprisingly though, none of those factors (language restriction, a search of sources other than bibliographic databases, and the number of databases searched) had a statistically significant correlation with the presence of PB. This could be due to some possible reasons: First, it might be due to the small sample size of SRs included in the test. Another reason could be that some risk of PB was inevitable even in the absence of language restriction of the search, seeking other sources of data, and searching a large number of databases. Nevertheless, the results of these tests do not exclude the fact that implementing these measures will most probably reduce the risk of PB.

Implications

Our findings indicate the lack of methodological sufficiency for conducting high-quality SRs in the most influential journals of the field, which in turn might have led to the possible dissemination of inflated results. We strongly encourage future reviewers and editors of journals to take the issue of PB seriously and demand authors to take measures to reduce its risk and use appropriate methods to assess its possible presence. As PB is an issue at the outcome level, we also encourage future reviewers who want to conduct a study similar to ours in their fields to also assess if the SRs that evaluated the risk of PB for their primary outcome did the same for the secondary outcomes in their study as well. Finally, if feasible, we encourage future researchers who want to conduct a similar study to use more robust selection criteria for including SRs, as our criteria, which was a necessity due to the lack of enough review resources in our team, might have

introduced some degree of selection bias in the results. Overall, the issue of PB is a serious issue that can result in the dissemination of inflated results, and thus, the whole scientific community is encouraged to take this phenomenon into more careful consideration, especially when conducting an SR.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13643-023-02404-0>.

Additional file 1: Appendix. The title, journal, and type of the review of the 75 systematic reviews included in this study are as follows.

Authors' contributions

FM: conduct, analysis, and presentation of the research.

SB: design and presentation of the research.

SI: design, conduct, and presentation.

Funding

This study was not funded.

Availability of data and materials

The authors confirm that the data supporting the findings of this study are available as needed.

Declarations

Competing interest

All authors declare no potential conflicts of interest regarding this study and its outcomes.

Received: 23 July 2023 Accepted: 30 November 2023

Published online: 02 January 2024

References

- DeVito NJ, Goldacre B. Catalogue of bias: publication bias. *BMJ Evid Based Med.* 2019;24(53–54):20181206. <https://doi.org/10.1136/bmjebm-2018-111107>.
- Bassler D, Mueller KF, Briel M, et al. Bias in dissemination of clinical research findings: structured OPEN framework of what, who and why, based on literature review and expert consensus. *BMJ Open.* 2016;6(e010024):20160121. <https://doi.org/10.1136/bmjopen-2015-010024>.
- Hopewell S, Loudon K, Clarke MJ, et al. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database Syst Rev.* 2009;2009(MR000006):20090121. <https://doi.org/10.1002/14651858.MR000006.pub3>.
- Higgins JPT, Thomas J, Chandler J, et al. *Cochrane handbook for systematic reviews of interventions.* 2019.
- Jill E, Laura L, Alfred B and Sally M. Finding what works in health care. 2011.
- Marks-Anglin A, Chen Y. A historical review of publication bias. *Res Synth Methods.* 2020;11(725–742):20200917. <https://doi.org/10.1002/jrsm.1452>.
- Herrmann D, Sinnott P, Holmes J, et al. Statistical controversies in clinical research: publication bias evaluations are not routinely conducted in clinical oncology systematic reviews. *Ann Oncol.* 2017;28:931–7. <https://doi.org/10.1093/annonc/mdw691>.
- Hedin RJ, Umberham BA, Detweiler BN, et al. Publication bias and nonreporting found in majority of systematic reviews and meta-analyses in anesthesiology journals. *Anesth Analg.* 2016;123:1018–25. <https://doi.org/10.1213/ANE.0000000000001452>.
- Atakop P, Vassar M. Publication bias in dermatology systematic reviews and meta-analyses. *J Dermatol Sci.* 2016;82(69–74):20160224. <https://doi.org/10.1016/j.jdermsci.2016.02.005>.
- Palma S, Delgado-Rodriguez M. Assessment of publication bias in meta-analyses of cardiovascular diseases. *J Epidemiol Community Health.* 2005;59:864–9. <https://doi.org/10.1136/jech.2005.033027>.
- Heavener T, Vassar M. A review of publication bias in the gastroenterology literature. *Indian J Gastroenterol.* 2018;37(58–62):20180227. <https://doi.org/10.1007/s12664-018-0824-2>.
- Pradhan R, Garnick K, Barkondaj B, et al. Inadequate diversity of information resources searched in US-affiliated systematic reviews and meta-analyses: 2005–2016. *J Clin Epidemiol.* 2018;102(50–62):20180604. <https://doi.org/10.1016/j.jclinepi.2018.05.024>.
- Pasick LJ, Yeakel H, Sataloff RT. Sex bias in laryngology research and publishing. *J Voice.* 2022;36(389–395):20200718. <https://doi.org/10.1016/j.jvoice.2020.06.021>.
- Peters JPM, Stegeman I, Grolman W, Hooft L. The risk of bias in randomized controlled trials in otorhinolaryngology: hardly any improvement since 1950. *BMC Ear Nose Throat Disord.* 2017;17(3):20170418. <https://doi.org/10.1186/s12901-017-0036-x>.
- Lin L, Chu H. Quantifying publication bias in meta-analysis. *Biometrics.* 2018;74(785–794):20171115. <https://doi.org/10.1111/biom.12817>.
- Hedges LV. Estimation of effect size under nonrandom sampling: the effects of censoring studies yielding statistically insignificant mean differences. *J Educ Stat.* 2016;9:61–85. <https://doi.org/10.3102/1076998609001061>.
- Sutton AJ, Duval SJ, Tweedie RL, et al. Empirical assessment of effect of publication bias on meta-analyses. *BMJ.* 2000;320:1574–7. <https://doi.org/10.1136/bmj.320.7249.1574>.
- Richard JL and David BP. *Summing Up: The Science of Reviewing Research* Harvard University Press: Cambridge, MA, 1984, xiii+ 191 pp. *Educ Res* 1986; 15: 16–17, publisher = Sage Publications Sage CA Thousand Oaks, CA.
- Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315:629–34. <https://doi.org/10.1136/bmj.315.7109.629>.
- Lau J, Ioannidis JP, Terrin N, et al. The case of the misleading funnel plot. *BMJ.* 2006;333:597–600. <https://doi.org/10.1136/bmj.333.7568.597>.
- Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics.* 1994;50:1088. <https://doi.org/10.2307/2533446>.
- Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med.* 2006;25:3443–57. <https://doi.org/10.1002/sim.2380>.
- Peters JL, Sutton AJ, Jones DR, et al. Comparison of two methods to detect publication bias in meta-analysis. *JAMA.* 2006;295:676–80. <https://doi.org/10.1001/jama.295.6.676>.
- Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol.* 2005;58:882–93. <https://doi.org/10.1016/j.jclinepi.2005.01.016>.
- Duval S, Tweedie R. A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. *J Am Stat Assoc.* 2000;95:89–98. <https://doi.org/10.1080/01621459.2000.10473905>.
- Lin L, Chu H, Murad MH, et al. Empirical comparison of publication bias tests in meta-analysis. *J Gen Intern Med.* 2018;33(1260–1267):20180416. <https://doi.org/10.1007/s11606-018-4425-7>.
- Page MJ, Higgins JPT and Sterne JAC. Assessing risk of bias due to missing results in a synthesis. *Cochrane Handb Syst Rev Interv.* 2022.
- Munn Z, Stern C, Aromataris E, et al. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. *BMC Med Res Methodol.* 2018;18(5):20180110. <https://doi.org/10.1186/s12874-017-0468-4>.
- Team RC. R: A language and environment for statistical computing. 2013.
- Guido S. meta: an R package for meta-analysis. *R news.* 2007;7:40–5.
- Tierney JF, Clarke M, Stewart LA. Is there bias in the publication of individual patient data meta-analyses? *Int J Technol Assess Health Care.* 2000;16:657–67. <https://doi.org/10.1017/s0266462300101217>.

32. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev*. 2021;10(89):20210329. <https://doi.org/10.1186/s13643-021-01626-4>.
33. 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: 2008–2012. <https://doi.org/10.1001/jama.283.15.2008>.
34. Saniasiaya J, Islam MA, Abdullah B. Prevalence of olfactory dysfunction in coronavirus disease 2019 (COVID-19): a meta-analysis of 27,492 patients. *Laryngoscope*. 2021;131(865–878):20201205. <https://doi.org/10.1002/lary.29286>.
35. Hartling L, Featherstone R, Nuspl M, et al. Grey literature in systematic reviews: a cross-sectional study of the contribution of non-English reports, unpublished studies and dissertations to the results of meta-analyses in child-relevant reviews. *BMC Med Res Methodol*. 2017;17(64):20170419. <https://doi.org/10.1186/s12874-017-0347-z>.
36. Morrison A, Polisena J, Husereau D, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *Int J Technol Assess Health Care*. 2012;28:138–44. <https://doi.org/10.1017/S0266462312000086>.
37. Nussbaumer-Streit B, Klerings I, Dobrescu AI, et al. Excluding non-English publications from evidence-syntheses did not change conclusions: a meta-epidemiological study. *J Clin Epidemiol*. 2020;118(42–54):20191104. <https://doi.org/10.1016/j.jclinepi.2019.10.011>.
38. Egger M, Juni P, Bartlett C, et al. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technol Assess (Winchester, England)*. 2003; 7:1–76.
39. Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technol Assess*. 2003;7:1–90. <https://doi.org/10.3310/hta7410>.
40. Pham B, Klassen TP, Lawson ML, Moher D. Language of publication restrictions in systematic reviews gave different results depending on whether the intervention was conventional or complementary. *J Clin Epidemiol*. 2005;58:769–76. <https://doi.org/10.1016/j.jclinepi.2004.08.021>.
41. Halfpenny NJ, Quigley JM, Thompson JC, Scott DA. Value and usability of unpublished data sources for systematic reviews and network meta-analyses. *Evid Based Med*. 2016;21(208–213):20160929. <https://doi.org/10.1136/ebmed-2016-110494>.
42. Wieseler B, Kerekes MF, Vervoelgyi V, et al. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. *BMJ*. 2012;344(d8141):20120103. <https://doi.org/10.1136/bmj.d8141>.
43. Peters JL, Sutton AJ, Jones DR, et al. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J Clin Epidemiol*. 2008;61(991–996):20080606. <https://doi.org/10.1016/j.jclinepi.2007.11.010>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

