

ORIGINAL ARTICLE

# Avoiding searching for outcomes called for additional search strategies: a study of Cochrane review searches

Tove Faber Frandsen<sup>a,\*</sup>, Michael Friberg Bruun Nielsen<sup>b</sup>, Mette Brandt Eriksen<sup>c,d</sup>

<sup>a</sup>University of Southern Denmark, Department of Design and Communication, Universitetsparken 1, 5000 Kolding, Denmark

<sup>b</sup>UCL University College, Niels Bohrs Allé 1, 5230 Odense M, Denmark

<sup>c</sup>The University Library of Southern Denmark, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

<sup>d</sup>Centre for Evidence-Based Medicine Odense (CEBMO) and Cochrane Denmark, Department of Clinical Research, University of Southern Denmark, Odense, Denmark

Accepted 24 May 2022; Published online 3 June 2022

## Abstract

**Objectives:** A search strategy for a systematic review that uses the Population, Intervention, Comparison, and Outcome framework should include the population, the intervention(s), and the type(s) of study design. According to existing guidelines, outcome should generally be excluded from the search strategy unless the search is multi-stranded. However, a recent study found that approximately 10% (51) of recent Cochrane reviews on interventions included outcomes in their literature search strategies. This study aims to analyze the alternatives to including outcomes in a search strategy by analyzing these recent Cochrane reviews.

**Study Design and Setting:** This study analyzes the 51 Cochrane reviews that included outcomes in their literature search strategies and analyzes the results of alternative search strategies that follow current recommendations.

**Results:** Despite a small study sample of 51 reviews the results show that many of the reviews excluded some of the recommended elements due to very broadly defined elements (e.g., all interventions or all people). Furthermore, excluding outcomes from the search strategy is followed by an enormous increase in the number of retrieved records making it unmanageable to screen, if using a single-stranded search strategy.

**Conclusion:** Recommendations for search strategies in difficult cases are called for. © 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

**Keywords:** Search strategies; Systematic reviews; Systematic searches; PICO; Alternative search strategies; Cochrane reviews; Multi-stranded search strategies

## 1. Introduction

Systematic reviews and meta-analyses are considered indispensable in the chain of scientific information and

key tools for evidence-based practice. The annual number of published systematic reviews is increasing rapidly, and they become more widely recognized and influential [1,2]. Unfortunately, many reviews use poorly justified methodologies and less-than-rigorous applications of these methods [3,4].

The literature search is the foundation for data collection in a systematic review, and the quality of it has a major impact on the overall quality of the final product [5]. Many studies confirm that there is room for improvement [6–9], and frequently, actions to strengthen the search strategies [10] are called for. One study found that 83% of systematic review search strategies in Cochrane reviews were flawed to an extent that would have affected search recall [11]. In a later study, however, 27% of the search strategies were error-free [9]. Recently, a study found that 78% of systematic review search strategies were flawed to an extent that

Author contributions: Tove Faber Frandsen contributed to conceptualization, methodology, data curation, investigation, project administration, and roles/writing—original draft. Michael Friberg Bruun Nielsen contributed to methodology, data curation, investigation, writing—review and editing. Mette Brandt Eriksen contributed to conceptualization, methodology, data curation, investigation, project administration, and writing—review and editing. All authors approved the final version to be published.

Author declaration: The authors wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

\* Corresponding author. Tel.: +4565501351.

E-mail address: [t.faber@sdu.dk](mailto:t.faber@sdu.dk) (T.F. Frandsen).

**What is new?****Key findings**

- Search strategies that include outcome are often updated reviews.
- Many search strategies include outcome if the recommended elements are very broadly defined (e.g., all interventions or all people) in the sample of 51 reviews investigated in this study.
- The number needed to screen after deduplication ranges from 16 to 24,703.

**What this adds to what is known?**

- Excluding outcome from the search strategy may be followed by an enormous increase in the number of retrieved records, which may increase the workload of screening records significantly.

**What is the implication, what should change now?**

- It may be difficult to avoid using outcome as part of some search strategy, which may call for alternative or additional search strategies.

would affect recall. The errors were related to lacking knowledge of the principles of information retrieval and/or the specific characteristics of searching in PubMed [12].

One, very important aspect is the conceptualization or translation of the research question into search concepts using a conceptualization tool such as Population, Intervention, Comparison, and Outcome (PICO) [13]. According to existing guidelines, a search strategy should include the population, the intervention(s), and the type(s) of study design (PISD) [14,15]. The Cochrane Handbook suggests including study type (in addition to appropriate elements from the PICO model) in the search strategy, and several studies have tested and validated various study type search filters [16–18]. On the other hand, the Cochrane Handbook recommends that outcome should generally be excluded from the search strategy [19]. The Population, Intervention, Comparison model performs better than the full PICO model [20] in terms of sensitivity of the search, as the outcome of interest may not always be mentioned in abstracts or subject headings [21,22]. The lower retrieval potential of outcomes has many causes, among these are selective outcome reporting; hence, some outcomes may only be mentioned in the full text of a publication [23–25].

A recent study found that approximately 10% of recent Cochrane reviews on interventions included outcomes in their literature search strategies [6]. This gives reason for concern, especially because the limitations are rarely mentioned. Furthermore, more than 90% of these reviews that include outcomes in their search strategy assessed outcomes that they did

not include in their search strategy. Tsujimoto et al. emphasize that their study of Cochrane reviews which include outcome in the search strategy does not intend to criticize the researchers behind these reviews. In some cases, outcome is included in the search strategy after careful consideration of the type of research question [6].

The aim of this study is to analyze the alternatives to including outcome in a search strategy. More specifically, we analyze the search strategies used in a sample of recent Cochrane reviews that includes outcome in their search strategy to find out if alternatives to this possible recall-reducing approach are available. We aim to determine (i) how often did the authors choose not to search for the PISD, (ii) the number of retrieved items for reviews including outcome in their search strategy, and finally (iii) the number of retrieved items for reviews when excluding outcome from the search strategy.

**2. Methods**

The analyses in this study are based on searches performed for recent Cochrane reviews. The reviews are selected because they all include outcome in their search strategy. From the study by Tsujimoto et al. [6], the 51 Cochrane reviews found to include outcomes in their search strategies were identified in PubMed. From each of the 51 reviews, the following information was extracted into an Excel sheet (data were extracted by T.F.F., and the accuracy of the extraction was verified by M.B.E. and M.F.B.N.):

- The PICOSD elements (population, intervention, comparison, outcome, and type(s) of study design), as stated in the methods section of the original Cochrane review
- The PICOSD elements that were included in the literature search (this information was extracted from the study by Tsujimoto et al [6])
- The number of records identified in the database searches after deduplication (since the actual number of records from the specific MEDLINE/PubMed searches was rarely stated in the reviews)
- Whether the review was an update of a previous review

From the extracted information, it was determined which of the PICOSD elements the authors of the reviews included in their strategy.

To determine how many records would be necessary to screen, if search strategies had followed the Cochrane recommendation of avoiding searching for outcome(s), MEDLINE searches (or PubMed searches, if MEDLINE was searched using PubMed) were replicated, i.e., both the original search stated in the reviews and a search excluding search terms for outcome(s). Furthermore, if the reviews only included randomized controlled trials (RCTs) as types of studies, a study type search filter for

RCTs was added to the search strategy. This also includes the reviews that did not apply a study type search filter in their own search strategy. The following filters were used: the Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); Ovid format or Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); and PubMed format [15,17,26].

Replication of the searches was done as accurately as possible, from the reported search strategies of the 51 reviews in MEDLINE (Ovid) or PubMed (depending on which database was used to perform the original search). However, some of the search histories do not include detailed information on time limits as well as study type search filters. This information is in some cases available in the methods section of the review or in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram; however, in some reviews, it is not clear which time limits or study type search filters were used. Complete modified search strategies for all the 51 reviews are available as supplementary material.

### 3. Results

We start with a short characterization of the reviews. Table 1 shows the number of updated and new reviews. In many cases, an updated review implies updating the search strategy of the previous review. The average number of screened records is about one-third in the updated reviews vs. the new reviews (Table 1). This is to be expected, as the updated reviews can typically limit the publication years and thus the number of records to be screened. However, the huge variation should be noticed. The number of records screened after deduplication ranges from 16 to 24,703.

Next, we address our first research question and determine how often the authors choose not to search for the PISD. Table 2 shows how many of the 51 reviews that included each of the PISD elements in the search strategies. About one-fourth of the reviews did not include one or more of the PISD elements (24%–32%) in the search strategies. The reason for elements being left out is generally that they are broadly defined elements (e.g., all interventions or all people). An example of a broadly defined population is the review of exercise vs. no exercise for the occurrence, severity, and duration of acute respiratory infections that include people of all ages [27]. An example of a broadly defined intervention is the review of organizational interventions for preventing and minimizing aggression directed towards health-care workers that includes all interventions that prevent aggression. In these cases, it would be very difficult, if not impossible, to define these elements with search terms, and the authors were forced to find alternative approaches to limit the retrieved records to get manageable results.

Finally, we address our two latter research questions and explore the number of retrieved items for reviews including outcome in their search strategy and the number of retrieved items for reviews when excluding outcome from the search strategy. Table 3 shows the search strategy elements in the searches when including outcome in the search strategy, as was done in the reviews. Furthermore, the number of records retrieved when following Cochrane recommendations and excluding outcome as a search element (thus including the PISD elements).

Summing up the results shows that excluding outcome from the search strategy in most cases is followed by an enormous increase in the number of retrieved records.

### 4. Discussion

The results of this study show that some of the reviews excluded certain elements because a very wide definition of an element was used (e.g., all interventions). In addition, results show that excluding outcomes from the search strategy (as recommended by Cochrane) in this sample of reviews leads to an enormous increase in the number of retrieved records, making manual screening nearly impossible. Finally, the results show that the number of records retrieved in MEDLINE/PubMed and the number needed to screen vary considerably across reviews.

We find a high number of updated reviews in our data sample. Searches in systematic reviews have improved considerably over time, but the search strategy in an updated review is typically not developed for that specific review but for the original review. Hence, maybe less attention may have been paid to the quality and adjustment of the search strategy. In some cases, it is difficult to understand why outcome would be included in a search strategy resulting in only 16 records to screen after removal of duplicates. But for an updated review, the search strategy is maybe not revised. On the other hand, it seems understandable to try to narrow down the number needed to screen when this number is as high as 24,703 even when outcome is part of the search strategy. We can see that across the different variations of element combinations, the average number of retrieved records would not be manageable in a screening process. The average number of records retrieved in PubMed/MEDLINE across the 51 reviews when excluding outcome from the search strategies is 136,978, which would not be advisable to screen. Choosing to include outcome in the search strategy could, of course, be due to a lack of knowledge of the fact that outcome often reduces sensitivity of the search [21] or the possibility of performing multistranded search strategies, as mentioned in Cochrane Handbook [15]. More focus on alternative search strategies as part of—or as an addendum to—the traditional single-stranded search strategy, when identifying evidence for complex topics, could perhaps be relevant.

**Table 1.** Updated and new reviews

Updates and new reviews	Updated reviews	New reviews
Number of reviews <i>n</i> (%)	30 (59%)	21 (41%)
Average number of screened records after deduplication according to review	1,936.3	6,276.3
Range	(63–9,100)	(16–24,703)

Before discussing the implications of the results further, we need to consider the limitations of this study. First, replicating the search strategies defined and described in the 51 reviews proved to be difficult, and it involved making some decisions regarding the use of time limits and study type search filters. A greater level of detail in the search histories would have been preferable, but it is worth noting that we are exploring the differences in the number of retrieved records between search strategies that either include or exclude outcome. Consequently, it is the relative numbers that we are particularly interested in. Second, we are working with a small sample of 51 reviews, and we should be careful making general conclusions based solely on the search strategies used in these reviews. However, the results of this study can be used to address the limitations of the current recommendations, and hopefully, further studies may cast a light on the extent of the problem.

Existing guidelines recommend that search strategies should consider including the population, the intervention(s), and the type(s) of study design to maximize sensitivity [15]. In some cases, outcome can be excluded, and the results would still be manageable. However, as we can see from the results of this study, it can be difficult to avoid searching for outcome as the number needed to screen increases enormously in some cases when excluding outcome. Two-thirds of the reviews including outcome in their search strategy deal with prevention [6]. This could imply that prevention reviews struggle more with developing single-stranded search strategies that follow current recommendations to exclude outcome. In such cases, a range of compensatory approaches, e.g., multistranded search strategies and citation searches, might be relevant [15]. In this study we find that most reviews (49 out of 51) supplemented their database searches with alternative approaches and sources. Citation searching was the most used alternative approach. Citation searching uses the citation network surrounding a source study to identify additional studies [28]. Consulting reference lists in included studies and existing, relevant systematic reviews also known as backward citation

searching were used in almost all reviews (49 out of 51), whereas forward citation searching was only used in 9 of the 51 reviews. Hand searching or screening of proceedings, journals, and websites as well as consulting experts were used in 21 and 28 reviews, respectively. We did not find examples of multistranded or multifaceted approaches where the reviewers conduct a series of searches, with different combinations of concepts, to capture a complex research question [15].

The development and availability of automation tools to support the preparation of systematic reviews is increasing. Some tools can support the screening process of systematic reviews (see overview of tools in the study by Khalil et al [29]), thereby making it manageable to screen a larger number of records compared to only manual screening. This development is also highlighted in the PRISMA 2020 flow diagram, which states that use of automation tools for screening of records can be indicated [30]. However automation tools do have limitations as mentioned by Khalil et al., e.g., integration with other tools and lack of availability [29]. Only 1 of the 51 reviews used an automation tool to screen the search result for non-RCTs [31]. A study investigated the potential barriers towards uptake of automation tools and found that among other things, a steep learning curve and cost of licensing were among the barriers [32]. Despite potential barriers to their use, automation tools do have the potential to speed up the process of completing systematic reviews, e.g., a case study by Clark et al. completed a systematic review in 2 weeks, using automation tools [33]. Automation tools may therefore be part of the solution to the problem of dealing with large search results, thus allowing searchers to avoid including outcome concepts in the search strategy to decrease the search result.

## 5. Conclusion

This study examines 51 reviews that include outcome in their search strategy, despite current guidelines

**Table 2.** PISD elements included in search strategy in the 51 included reviews

Elements	P (patients)	I (intervention)	SD (type of study design)
Included in search strategy			
No. of reviews <i>n</i> (%)	35 (68%)	37 (73%)	39 (76%)

**Table 3.** Elements in search strategies following recommendation to use PISD elements

Elements	PISD(O)	PSD(O)	ISD(O)	PI(O)	I(O)
Number of reviews <i>n</i> (%)	17 (33%)	10 (20%)	19 (37%)	4 (8%)	1 (2%)
Average number of retrieved records in MEDLINE/PubMed including outcome	1,332.5	2,313.3	2,093.9	773.3	233.0
(Range)	(11–5,023)	(114–14,583)	(89–12,020)	(295–1,518)	-
Average number of retrieved records in MEDLINE/PubMed excluding outcome	57,833.6	192,971.4	63,449.9	309,439.3	59,486
(Range)	(12–726,373)	(1,198–1,631,548)	(186–231,265)	(4,064–1,198,555)	-

*Abbreviations:* PISD, population, intervention, type of study design; PSD, population, type of study design; ISD, intervention, type of study design; PI, population, intervention; I, intervention.

recommending the exclusion of this search element. The results, which are based on a relatively small sample, show that it can be very difficult to avoid including outcome in the search strategy, if using a single-stranded search strategy. When excluding outcome, broadly defined populations and interventions result in massive increases in the number of retrieved records, making it impossible to screen the records resulting from a single-stranded search strategy. Despite possible aid from automation tools when further processing the retrieved records, the increase in the number of retrieved records when excluding outcome in the search strategy was often massive. The Cochrane Handbook suggests the use of multistranded search approaches for complex research questions [15] and we suggest putting more emphasis on these recommendations to encourage alternative search approaches than including outcome concepts in a single-stranded search strategy. Hence, it would be very valuable to develop methods for developing efficient search strategies for reviews with broadly defined populations and interventions. A prevention review is an example of a type of review that may require additional work with the development of a search strategy that does not include outcome.

## References

- [1] Patsopoulos NA, Analatos AA, Ioannidis JP. Relative citation impact of various study designs in the health sciences. *JAMA* 2005;293:2362–6.
- [2] Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med* 2010; 7(9):e1000326.
- [3] Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q* 2016; 94(3):485–514.
- [4] Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. *Nature* 2018;555:175–82.
- [5] Leclercq V, Hilgismann M, Parisi G, Beaudart C, Tirelli E, Bruyère O. Best-worst scaling identified adequate statistical methods and literature search as the most important items of AMSTAR2 (A measurement tool to assess systematic reviews). *J Clin Epidemiol* 2020;128:74–82.
- [6] Tsujimoto Y, Tsutsumi Y, Kataoka Y, Banno M, Furukawa TA. Around ten percent of most recent Cochrane reviews included outcomes in their literature search strategy and were associated with potentially exaggerated results: a research-on-research study. *J Clin Epidemiol* 2021;141:74–81.
- [7] Borges Migliavaca C, Stein C, Colpani V, Barker TH, Munn Z, Falavigna M. How are systematic reviews of prevalence conducted? A methodological study. *BMC Med Res Methodol* 2020;20:1–9.
- [8] Harari MB, Parola HR, Hartwell CJ, Riegelman A. Literature searches in systematic reviews and meta-analyses: a review, evaluation, and recommendations. *J Vocat Behav* 2020;118:103377.
- [9] Franco JVA, Garrote VL, Escobar Liquitay CM, Vietto V. Identification of problems in search strategies in Cochrane Reviews. *Res Synth Methods* 2018;9(3):408–16.
- [10] Neilson CJ. Adoption of peer review of literature search strategies in knowledge synthesis from 2009 to 2018: an overview. *Health Info Libr J* 2021;38:160–71.
- [11] Sampson M, McGowan J. Errors in search strategies were identified by type and frequency. *J Clin Epidemiol* 2006;59:1057–63.
- [12] 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;107:210–21.
- [13] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol* 2016;75:40–6.
- [14] European Network for Health Technology Assessment (EUnetHTA). Process of information retrieval for systematic reviews and health technology assessments on clinical effectiveness. EUnetHTA methodological guidance version 1.2; 2017. Available at <https://www.eunetha.eu/process-of-information-retrieval-for-systematic-reviews-and-health-technology-assessments-on-clinical-effectiveness>. Accessed May 16, 2022.
- [15] Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, et al. Chapter 4: searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0. Cochrane; 2021. Available at [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook). Accessed May 16, 2022.
- [16] McKibbin KA, Wilczynski NL, Haynes RB. Retrieving randomized controlled trials from medline: a comparison of 38 published search filters. *Health Info Libr J* 2009;26:187–202.
- [17] Glanville J, Kotas E, Featherstone R, Dooley G. Which are the most sensitive search filters to identify randomized controlled trials in MEDLINE? *J Med Libr Assoc* 2020;108(4):556–63.
- [18] Rosumeck S, Wagner M, Wallraf S, Euler U. A validation study revealed differences in design and performance of search filters for qualitative research in PsycINFO and CINAHL. *J Clin Epidemiol* 2020;128:101–8.
- [19] Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. *Cochrane handbook for systematic reviews of interventions* version 6.0 2019. Available at [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook). Accessed May 16, 2022.
- [20] Agoritsas T, Merglen A, Courvoisier DS, Combescure C, Garin N, Perrier A, et al. Sensitivity and predictive value of 15 PubMed search

- strategies to answer clinical questions rated against full systematic reviews. *J Med Internet Res* 2012;14(3):e85.
- [21] Frandsen TF, Bruun Nielsen MF, Lindhardt CL, Eriksen MB. Using the full PICO model as a search tool for systematic reviews resulted in lower recall for some PICO elements. *J Clin Epidemiol* 2020;127:69–75.
- [22] Frandsen TF, Lindhardt CL, Eriksen MB. Performance of conceptual framework elements for the retrieval of qualitative health literature: a case study. *J Med Libr Assoc* 2021;109:388–94.
- [23] Jin L, Hua F, Cao Q. Reporting quality of randomized controlled trial abstracts published in leading laser medicine journals: an assessment using the CONSORT for abstracts guidelines. *Lasers Med Sci* 2016;31(8):1583–90.
- [24] Duyx B, Swaen GM, Urlings MJ, Bouter LM, Zeegers MP. The strong focus on positive results in abstracts may cause bias in systematic reviews: a case study on abstract reporting bias. *Syst Rev* 2019;8(1):174.
- [25] Duyx B, Swaen GM, Urlings MJ, Bouter LM, Zeegers MP. The strong focus on positive results in abstracts may cause bias in systematic reviews: a case study on abstract reporting bias. *Syst Rev* 2019;8(1):1–8.
- [26] Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M, et al. Technical supplement to chapter 4: searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston MS, Li T, Page MJ, et al, editors. *Cochrane Handbook for Systematic Reviews of Interventions Version 6*. Cochrane; 2019. Available at [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook). Accessed May 16, 2022.
- [27] Grande AJ, Keogh J, Silva V, Scott AM. Exercise versus no exercise for the occurrence, severity, and duration of acute respiratory infections. *Cochrane Database Syst Rev* 2020;4(4):CD010596.
- [28] Briscoe S, Bethel A, Rogers M. Conduct and reporting of citation searching in Cochrane systematic reviews: a cross-sectional study. *Res Synth Methods* 2020;11(2):169–80.
- [29] Khalil H, Ameen D, Zarnegar A. Tools to support the automation of systematic reviews: a scoping review. *J Clin Epidemiol* 2021;144:22–42.
- [30] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
- [31] Jefferson T, Del Mar CB, Dooley L, Ferroni E, Al-Ansary LA, Bawazeer GA, et al. Physical interventions to interrupt or reduce the spread of respiratory viruses. *Cochrane Database Syst Rev* 2020;11(11):Cd006207.
- [32] Van Altena A, Spijker R, Olabarriaga S. Usage of automation tools in systematic reviews. *Res Synth Methods* 2019;10(1):72–82.
- [33] Clark J, Glasziou P, Del Mar C, Bannach-Brown A, Stehlik P, Scott AM. A full systematic review was completed in 2 weeks using automation tools: a case study. *J Clin Epidemiol* 2020;121:81–90.