Journal of Clinical Epidemiology 149 (2022) 154−164

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Searching two or more databases decreased the risk of missing relevant studies: a metaresearch study

Hannah Ewald[a,b,*], Irma Klerings[c], Gernot Wagner[c], Thomas L. Heise[d,e], Jan M. Stratil[f], Stefan K. Lhachimi[e,g], Lars G. Hemkens[b,h,i], Gerald Gartlehner[c,j], Susan Armijo-Olivo[k,l], Barbara Nussbaumer-Streit[c]

[a]*University Medical Library, University of Basel, Basel, Switzerland*
[b]*Basel Institute for Clinical Epidemiology and Biostatistics, Department of Clinical Research, University Hospital Basel, University of Basel, Basel, Switzerland*
[c]*Cochrane Austria, Danube University Krems, Krems a.d. Donau, Austria*
[d]*Leibniz Institute for Prevention Research and Epidemiology - BIPS, Bremen, Germany*
[e]*Institute for Public Health and Nursing Research - IPP, Health Sciences Bremen, University of Bremen, Bremen, Germany*
[f]*Institute for Medical Informatics, Biometry and Epidemiology, Pettenkofer School of Public Health, LMU Munich, Munich, Germany*
[g]*Department of Health, Nursing, Management, University of Applied Sciences Neubrandenburg, 17033, Neubrandenburg, Germany*
[h]*Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA*
[i]*Meta-Research Innovation Center Berlin (METRIC-B), Berlin Institute of Health, Berlin, Germany*
[j]*RTI International, Research Triangle Park, NC, USA*
[k]*University of Applied Sciences Osnabrück, Faculty of Economics and Social Sciences − Osnabrück, Germany*
[l]*Faculties of Rehabilitation Medicine and Medicine and Dentistry, University of Alberta, Edmonton, Canada*

Accepted 25 May 2022; Published online 30 May 2022

Abstract

Background and Objectives: Assessing changes in coverage, recall, review, conclusions and references not found when searching fewer databases.

Methods: In randomly selected 60 Cochrane reviews, we checked included study publications' coverage (indexation) and recall (findability) using different search approaches with MEDLINE, Embase, and CENTRAL and related them to authors' conclusions and certainty. We assessed characteristics of unfound references.

Results: Overall 1989/2080 included references, were indexed in ≥1 database (coverage = 96%). In reviews where using one of our search approaches would not change conclusions and certainty (n = 44-54), median coverage and recall were highest (range 87.9%-100.0% and 78.2%-93.3%, respectively). Here, searching ≥2 databases reached >95% coverage and ≥87.9% recall. In reviews with unchanged conclusions but less certainty (n = 2-8): 63.3%-79.3% coverage and 45.0%-75.0% recall. In reviews with opposite conclusions (n = 1-3): 63.3%-96.6% and 52.1%-78.7%. In reviews where a conclusion was no longer possible (n = 3-7): 60.6%-86.0% and 20.0%-53.8%.

Conflict of interest: No conflicts to declare.

Exclusive licences: 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."

Funding source: This work was not directly supported. Employees of Cochrane Austria were supported by internal funds of Cochrane Austria.

Disclaimer: The funders had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript or its submission for publication.

Information on author access to data: Available on request from the corresponding author.

Transparency declaration: The Corresponding Author 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 planned have been explained.

Ethical approval: Not required, this article does not contain any personal medical information about any identifiable living individuals.

Declaration of interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

* Corresponding author. University Medical Library, University of Basel, Spiegelgasse 5, CH-4051 Basel, Switzerland. Tel.: +41 (0)61 2075564.

E-mail address: Hannah.Ewald@unibas.ch (H. Ewald).

https://doi.org/10.1016/j.jclinepi.2022.05.022

0895-4356/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

The 265 references that were indexed but unfound were more often abstractless (30% vs. 11%) and older (28% vs. 17% published before 1991) than found references.

Conclusion: Searching ≥2 databases improves coverage and recall and decreases the risk of missing eligible studies. If researchers suspect that relevant articles are difficult to find, supplementary search methods should be used. © 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Systematic review; Rapid review; Recall; Database coverage; Search strategy; Literature search

1. Introduction

Systematic reviews are crucial to inform healthcare decision-making. Their purpose is to identify, appraise, and synthesize all evidence on a specific research question using explicit, systematic methods [1]. A comprehensive systematic literature search is the basis for every systematic review. It aims to reduce bias in study selection by using a variety of information sources. These searches usually comprise of searching multiple bibliographic databases and supplementary search methods (hand searches, searches in study registers, web searching, etc.). The Methodological Expectations of Cochrane Intervention Reviews (MECIR) require searching at least MEDLINE, CENTRAL, and Embase (depending on access availability), study registers as well as reference lists of included studies and relevant systematic reviews [2].

Recent methods studies evaluating research studies of medical or psychological topics [3−5] found that the majority of relevant references included in Cochrane reviews were available (indexed) in MEDLINE (i.e., database coverage). Removing studies that were not indexed in MEDLINE from the meta-analyses had only little impact on the effect estimates [3,4,6]. However, these results may be overestimating the impact of searching only one database as they have not assessed the recall of the indexed studies. Recall is, in this context, a more meaningful measure: It indicates the amount of relevant indexed references that were actually found by the search strategy in a particular database. Comprehensive searches for systematic reviews aim for high recall by applying broad searches to find all relevant references. As this usually retrieves many irrelevant references, practical concerns require balancing the sensitivity of the searches with the available time and budget [1].

Combining multiple information sources often improves recall compared to single-database searches [7−10]. Bramer et al. [8] assessed systematic search results in various medical topics, from therapeutic effectiveness and diagnostic accuracy to ethics and public health. They found high database coverage in MEDLINE and Embase, but reported that in some cases even the cumulative recall of MEDLINE and Embase was below 50%.

This project is part of a larger methodological project that examined how conclusions of 60 published Cochrane systematic reviews [11,12] and treatment effect estimates [13] changed when the number of information sources is limited (details below). Here, we extend the published results by focusing on database coverage and recall and possible reasons as to why some studies cannot be retrieved anymore.

Our aim was to assess how coverage and recall relate to conclusions in reviews and why some references could not be found anymore when searching in a limited number of databases.

2. Methods

2.1. Methods of the main study

Here we briefly summarize the most important methodological aspects of the underlying project previously reported in detail [11−13]. We randomly selected 60 Cochrane reviews that included a replicable search strategy, had clear conclusions, and reported a meta-analysis [12]. We determined the clinical topic of each review and categorized all reviews according to the type of intervention (pharmacological or nonpharmacological). From each review, we used all references reported in the "References to studies included in this review" section as a gold standard-set to assess 14 different search approaches using a limited number of sources (i.e., the 14 possible combinations of MEDLINE, Embase, CENTRAL with and without screening reference lists). We identified which studies were not found when we limited the number of databases and recalculated the estimates from meta-analyses for all outcomes of the main summary of findings table for each review accordingly [11,13]. Then, we asked the original authors whether the resulting summary of findings tables would alter their conclusions. Possible answers were (1) conclusion does not change ("same conclusion/same certainty"), (2) same conclusion but with less certainty, (3) conclusion with opposite direction ("opposite conclusion"), and (4) conclusion not possible anymore [11].

3. Methods of this sub-study

For this sub-study, we used the same gold standard-set as mentioned above. We first determined what the individual and cumulative database coverage and recall would have been, had the searches been done in a limited number of databases or database combinations, namely: (1) MEDLINE-only (M),

What is new?

Key findings

- Searching at least two databases improves coverage and recall and decreases the chances of making an inappropriate conclusion in a review.

What this adds to what is known

- A systematic review not always needs to find all eligible studies to arrive at the same conclusion with the same certainty as compared to including all eligible studies. However, in some cases even combined searches of MEDLINE, Embase, and CENTRAL to identify research may not be sufficient to draw any conclusion.
- Many studies, almost a third, which were indexed but not found by searching MEDLINE, Embase and CENTRAL had no abstract.

What is the implication, what should change now?

- In most cases, searching in a limited number of databases will lead to the same conclusions in comprehensive systematic reviews of RCTs.
- If researchers suspect that some relevant articles may be difficult to find or have no abstract, supplementary strategies rather than additional databases should be used.
- Researchers should carefully pick suitable electronic databases especially with respect to topic and relevant subject headings or consult an information specialist.

(2) Embase-only (E), (3) CENTRAL-only (C); (4) MEDLINE + Embase (M + E), (5) MEDLINE + CENTRAL (M + C), (6) Embase + CENTRAL (E + C), and (7) MEDLINE + Embase + CENTRAL (M + E + C).

We then put database coverage and recall of these limited searches in relation to changes in review conclusions. Finally, we assessed indexation status, publication date, abstract availability, language, intervention type and topic of references that could not be found.

As the focus of this study is coverage and recall, we did not consider search combinations with reference list checking.

From the Cochrane reviews and our previous study [11], we extracted year of publication, publication type, study design, topic, and intervention types (pharmacological or nonpharmacological) for each reference included in the gold standard-set. From studies not found, we specifically checked indexation status, abstract availability, publication date, language, and topic.

3.1. Data analysis

We used the Endnote X8 (Clarivate Analytics) and Excel 2016 (Microsoft) for data management and analysis. For each review, we saved the included studies to EndNote and exported the bibliographic data as tables to Excel. For each reference, we noted the availability in the target databases using the database-specific identifiers. We used these identifiers to determine if available studies were found by the original search strategies.

We calculated database coverage and recall as ratios (Appendix 1): The numerator for database coverage is the number of relevant references indexed in a database, and the numerator for recall is the individual or cumulative number of relevant references found in a database or a database combination, respectively, at the time of the original search. As denominator for both these metrics, we used the gold standard-set.

To evaluate overall coverage and recall, we performed descriptive statistical analyses for each of the seven search approaches. We calculated the mean and the median and interquartile range (IQR) across all reviews. We contrasted coverage and recall of the search approaches to the conclusions of the respective Cochrane reviews (‘‘same conclusion/same certainty’’, ‘‘same conclusion/less certainty’’, ‘‘opposite conclusion’’, ‘‘no conclusion possible’’). We also analyzed these results by intervention types, i.e., pharmaceutical and nonpharmaceutical.

We examined references that were not found in at least one of the three databases with respect to their indexation status, publication date, abstract availability, language, intervention type, and topic. Between indexed references that were found with one of the limited database search approaches and those that were not found, we compared abstract availability and publication date (before 1991/1991 or later; rationale for date cut-off: Introduction of ‘‘Randomized Controlled Trial’’ and ‘‘Clinical Trial’’ as Publication Types in Medline in 1991).

4. Results

4.1. Study characteristics

Each of the 60 Cochrane reviews used between 2 and 18 (median 6) information sources in addition to M + E + C (including other databases, grey literature, citation tracking [11]). Fifty-six reviews included solely randomized controlled trials (RCTs) or quasi-RCTs; four reviews additionally included nonrandomized controlled trials or observational studies (designs as defined by Cochrane review authors). The review characteristics (number of reviews, included studies, and proportion of pharmacological interventions) are reported in relation to changes in reviews’ conclusions in Appendix 2.

The 60 Cochrane reviews included 2,080 references of studies (the gold standard-set) published between 1960 and 2016: 2014 (96.8%) were references to RCTs or

Table 1. Characteristics of references of studies included by the 60 Cochrane reviews

Characteristics	Subcategory	References 2,080 (100%)
Intervention	Pharmacological	1,052 (50.6%)
	Nonpharmacological	1,028 (49.4%)
Study design	RCT/quasi-RCT	2,014 (96.8%)
	Other design[a]	66 (3.2%)
Publication date	1960 to 1990	375 (18.0%)
	1991 to 2016	1,671 (80.3%)
	Undated document	34 (1.6%)
Publication type	Journal article	1,807 (86.9%)
	Conference abstract	180 (8.7%)
	Study register entry	62 (3.0%)
	Pharma[b]	4 (0.2%)
	Book	3 (0.1%)
	Thesis	19 (0.9%)
	Other[c]	5 (0.2%)

Abbreviation: RCT, randomized controlled trial.
[a] non-randomized controlled clinical trial, controlled cohort study, before and after study, interrupted time series.
[b] Documents provided by pharmaceutical companies.
[c] Other unpublished documents.

quasi-randomized trials. Most references were journal articles (1,807/2,080, 87%), although 273 (13.1%) referred to other publication types (e.g., conference abstracts, study registries; Table 1).

4.2. Coverage

Of the 2,080 references in the gold standard-set, 1,989 (95.6%) were indexed in at least one of the three databases (median coverage = 98.5%; Table 2).

When limiting the number of databases, single-database coverage across all 60 Cochrane reviews was highest in CENTRAL (median: 91.3%, IQR: 75.0%-100.0%). The median coverage in CENTRAL and EMBASE was higher for reviews on pharmacological than nonpharmacological interventions although this was the opposite for MEDLINE (Table 2).

The combination of databases increased the median coverage. The highest coverage could be attained with M + E + C (Median 98.5%; IQR 91.9%-100.0%) (Table 2).

For nearly half of the reviews under investigation (n = 29; 16 pharmacological, 13 nonpharmacological), there was at least one database or combination with a coverage of 100% (Appendix 3).

4.3. Recall

Of the 2,080 references in the gold standard-set, 1,724 (83%) could be found searching in M + E + C (Table 2).

Across all 60 Cochrane reviews, recall was markedly lower than coverage. The highest median single-database recall was in MEDLINE (median: 75.5%, IQR: 56.8%-87.0%, Table 2) which seemed balanced between reviews of pharmacological and nonpharmacological interventions. CENTRAL and Embase median recall was higher for reviews on pharmacological interventions (Table 2).

There was a gradual increase in recall with more databases included across evaluated search approaches. The highest median recall was 90.0% in M + E + C (IQR: 77.7%-100.0%; pharmacological interventions: n = 30, median 94.1%, IQR 82.4%-100.0%; nonpharmacological interventions: n = 30, median 87.0%, IQR 64.0%-97.7%) (Table 2).

In 18 Cochrane reviews (11 pharmacological, 7 nonpharmacological), at least one search strategy or combination had a recall of 100% (Appendix 3).

4.4. Coverage in relation to reviews' conclusions

Table 3 shows how database coverage relates to the conclusions of the Cochrane review when searching in a limited number of databases.

In reviews where the conclusion and certainty did not change (depending on search approach n = 44 to 54 of 60), CENTRAL and each combination of at least two databases reached over 95% median coverage (although a mean coverage of at least 95% was only reached with C + E and M + E + C).

In reviews where the conclusion did not change but the authors were less certain (n = 2 to 8 of 60), median coverage was much lower ranging from 63.3% to 78.3% from single database searches to M + E + C (Table 3).

When limiting the information sources to M + E + C, the conclusion did not change (i.e., with the same or less certainty) in 56 reviews including 1,969 references. Half of these reviews (28/56; 50%) included at least one reference that was not indexed and hence not found (80/1,969, 4%; coverage = 78.2%-78.5%; Table 3).

In reviews where searching in a limited number of databases led to an opposite conclusion (n = 1 to 3 of 60), median coverage ranged between 63.3% (M) and 96.6% (C + E, M + E, M + E + C) (Table 3).

In reviews where a conclusion could no longer be drawn (n = 3 to 7 of 60 reviews), we found the lowest median coverage which ranged from 60.6% (C) to 86.0% (M + E).

When limiting the information sources to M + E + C, a conclusion was no longer possible or changed to the opposite of the original review in three [14–16] and one [17] reviews, respectively. Median coverage of these reviews ranged between 64.4% and 96.6%. Three of these four reviews included at least one reference that was not indexed (overall 11 of 111, 10%), but coverage differed greatly between individual reviews (54%-100%; Appendix 5).

4.5. Recall in relation to reviews' conclusions

Analysis of recall by conclusion (Table 4) showed that search approaches that led to the same conclusion with the same certainty had a markedly higher median recall than those where any change occurred (‘‘same direction

Table 2. Database coverage and recall of all references included in the 60 Cochrane reviews and stratified according to pharmacological and non-pharmacological interventions

		Database coverage in %				Recall in %		
Database combination	**Coverage N References**	**Mean**	**Median (IQR) per review**		**Recall N References**	**Mean**	**Median (IQR) per review**	
Gold standard-set	2,080				2,080			
MEDLINE	1,650	79.3	87.0	(72.2-93.0)	1,354	65.1	75.5	(56.8-87.0)
Embase	1,602	77.0	83.0	(62.6-92.4)	1,276	61.3	71.0	(46.5-84.3)
CENTRAL	1,839	88.4	91.3	(75.0-100.0)	1,502	72.2	75.0	(65.1-92.6)
M + E	1,809	87.0	94.1	(75.0-100.0)	1,531	73.6	85.4	(66.6-95.1)
M + C	1,942	93.4	95.7	(88.5-100.0)	1,628	78.3	85.5	(74.0-94.2)
C + E	1,959	94.2	96.6	(88.7-100.0)	1,656	79.6	89.0	(72.6-98.4)
M + E + C	1,989	95.6	98.5	(91.9-100.0)	1,724	82.9	90.0	(77.7-100.0)
Pharmacological interventions (n = 30)								
Gold standard-set	1,052				1,052			
MEDLINE	799	76.0	85.7	(68.8-89.6)	745	70.8	75.7	(62.8-86.6)
Embase	842	80.0	86.6	(66.3-94.4)	767	72.9	74.6	(62.5-83.7)
CENTRAL	967	91.9	94.0	(86.2-100.0)	870	82.7	82.3	(72.2-95.8)
M + E	892	84.8	90.5	(76.8-97.3)	850	80.8	84.5	(71.3-95.4)
M + C	994	94.5	96.7	(89.4-100.0)	907	86.2	88.4	(75.7-95.6)
C + E	1,016	96.6	96.9	(93.7-100.0)	960	91.3	93.3	(82.4-100.0)
M + E + C	1,018	96.8	98.8	(93.7-100.0)	966	91.8	94.1	(82.4-100.0)
Nonpharmacological interventions (n = 30)								
Gold standard-set	1,028				1,028			
MEDLINE	851	82.8	89.6	(73.5-95.6)	609	59.2	75.5	(52.2-88.4)
Embase	760	73.9	82.0	(60.7-91.2)	508	49.4	61.3	(30.5-83.3)
CENTRAL	872	84.8	81.2	(70.1-100.0)	632	61.5	70.7	(55.7-89.6)
M + E	917	89.2	95.9	(78.8-100.0)	680	66.1	85.1	(60.1-94.4)
M + C	948	92.2	94.7	(86.4-100.0)	721	70.1	80.5	(64.0-92.9)
C + E	943	91.7	96.3	(81.8-100.0)	695	67.6	80.6	(59.4-93.9)
M + E + C	971	94.5	98.5	(91.8-100.0)	759	73.7	87.0	(64.0-97.7)

Abbreviations: C, CENTRAL; E, Embase; IQR, Interquartile range; M, MEDLINE; Gold standard-set, references included in the 60 Cochrane reviews using the original search methods; Mean, The total number of included references indexed or found in a database divided by the total number of included references found by the comprehensive Cochrane review search; Median, The median value of database coverage or recall per review; N, number of.

but less certainty'', ''opposite conclusion'' or ''no conclusion possible''). The category ''no conclusion possible'' had the lowest median recall across all limited database search approaches (median ranged from 20.0% to 53.8%). Depending on which approach was used (single database search, two databases, or combination of all three), there were always three to seven Cochrane reviews where a conclusion was no longer possible (median recall ranged from 20.0% to 53.8%) and one to three reviews where the opposite conclusion was drawn (median recall ranged from 52.1% to 78.7%).

When limiting the information sources to M + E + C, the conclusion did not change (i.e., with the same or less certainty) in 56 reviews including 1,969 references. A sixth of these references had not been found (328/1,969, 17%). Median recall per review ranged between 62.4% and 100% (Table 4). Of the 1,969 references, 1,706 were journal articles and 263 grey literature articles of which 238 (14%) and 90 (34%) were not found, respectively. The 90/263 grey literature references that could not be found when limiting the information sources to M + E + C and the conclusion remained unchanged were 15/60 study register entries (25%), 52/177 conference abstracts (29%), 11/14 theses (79%), 3/3 books (100%), 4/4 pharma documents (100%), and 5/5 other document types (100%). Of the 111 study references (101 journal articles and 10 grey literature articles) included in the four reviews where a conclusion was no longer possible or changed to the opposite of the original review, 28 were not found (Appendix 4).

4.6. Characteristics of references not found

Overall, 356 of the 2,080 (17%) references were not found when using any of the seven search approaches in a limited number of databases.

Table 3. Database coverage in relation to search approach and conclusion category

	MEDLINE (M)					Embase (E)					CENTRAL (C)				
		References		Coverage per review in %			References		Coverage per review in %			References		Coverage per review in %	
Conclusions	**N reviews**	**Incl.**	**Index.**	**Mean**	**Median (IQR)**	**N reviews**	**Incl.**	**Index.**	**Mean**	**Median (IQR)**	**N reviews**	**Incl.**	**Index.**	**Mean**	**Median (IQR)**
Same conclusion, same certainty	48	1,753	1,397	79.7	87.9 (76.1-93.7)	44	1,561	1,244	79.7	87.9 (74.6-93.5)	47	1,762	1,611	91.4	96.2 (85.2-100.0)
Same conclusion, less certainty	6	87	58	66.7	63.3 (57.3-78.6)	6	220	159	72.3	66.0 (52.5-76.1)	8	199	133	66.8	67.8 (62.7-78.5)
Opposite conclusion	2	119	91	76.5	63.3 (50.0-76.6)	3	145	110	75.9	79.2 (55.2-85.1)	1	89	81	91.0	91.0 (N/A)
No conclusion possible	4	121	104	86.0	84.0 (65.9-94.7)	7	154	88	57.1	62.5 (48.8-81.3)	4	30	14	46.7	60.6 (37.7-81.3)
	M + E					M + C					C + E				
Same conclusion, same certainty	50	1,794	1,562	87.1	95.3 (86.3-100.0)	53	1,912	1,802	94.2	96.4 (90.0-100.0)	50	1,816	1,739	95.8	98.0 (92.0-100.0)
Same conclusion, less certainty	5	76	53	69.7	71.9 (57.1-73.7)	3	57	42	73.7	71.9 (68.1-81.4)	6	153	121	79.1	79.3 (71.2-88.2)
Opposite conclusion	1	89	86	96.6	96.6 (N/A)	1	89	84	94.4	94.4 (N/A)	1	89	86	96.6	96.6 (N/A)
No conclusion possible	4	121	108	89.3	86.0 (65.9-97.7)	3	22	14	63.6	75.0 (64.4-87.5)	3	22	13	59.1	75.0 (60.6-87.5)
	M + E + C														
Same conclusion, same certainty			54		1,923			1,853		96.4					100.0 (93.5-100.0)
Same conclusion, less certainty			2		46			36		78.3					78.3 (78.2-78.5)
Opposite conclusion			1		89			86		96.6					96.6 (N/A)
No conclusion possible			3		22			14		63.6					75.0 (64.4-87.5)

Abbreviations: C, CENTRAL; E, Embase; Incl., References included by reviews; Index., Reverences indexed in databases; IQR, Interquartile range; M, MEDLINE; Mean, The total number of included references indexed in a database divided by the total number of included references; Median: The median value of database coverage per review; N/A, not applicable (only 1 review).

Table 4. Recall in relation to search approach and conclusion category

	MEDLINE (M)					Embase (E)					CENTRAL (C)				
		N References		Recall per review in %			N References		Recall per review in %			N References		Recall per review in %	
Conclusions	**N reviews**	**Incl.**	**Found**	**Mean**	**Median (IQR)**	**N reviews**	**Incl.**	**Found**	**Mean**	**Median (IQR)**	**N reviews**	**Incl.**	**Found**	**Mean**	**Median (IQR)**
Same conclusion, same certainty	48	1,753	1,208	68.9	81.4 (66.7-88.1)	44	1,561	1,112	71.2	78.2 (62.5-88.1)	47	1,762	1,313	74.5	83.3 (70.2-95.7)
Same conclusion, less certainty	6	87	51	58.6	50.0 (48.0-64.1)	6	220	49	22.3	45.0 (31.4-57.5)	8	199	111	55.8	59.5 (47.6-66.7)
Opposite conclusion	2	119	73	61.3	52.1 (42.7-61.4)	3	145	85	58.6	66.3 (44.1-72.7)	1	89	65	73.0	73.0 (N/A)
No conclusion possible	4	121	22	18.2	21.4 (9.1-41.8)	7	154	29	18.8	20.0 (13.8-35.4)	4	30	13	43.3	29.3 (9.4-53.4)
	M + E					M + C					C + E				
Same conclusion, same certainty	50	1,794	1,387	77.3	88.6 (73.8-97.1)	53	1,912	1,508	78.9	87.9 (75.0-95.1)	50	1,816	1,484	81.7	92.5 (79.2-100.0)
Same conclusion, less certainty	5	76	46	60.5	50.0 (50.0-57.9)	3	57	41	71.9	75.0 (66.1-78.4)	6	153	90	58.8	60.4 (50.5-67.3)
Opposite conclusion	1	89	70	78.7	78.7 (N/A)	1	89	66	74.2	74.2 (N/A)	1	89	70	78.7	78.7 (N/A)
No conclusion possible	4	121	28	23.1	24.5 (13.6-41.8)	3	22	13	59.1	53.8 (26.9-64.4)	3	22	12	54.5	46.2 (23.1-60.6)
	M + E + C														
Same conclusion, same certainty	54			1,923				1,608			83.6			93.3 (85.0-100.0)	
Same conclusion, less certainty	2			46				33			71.7			67.6 (62.4-72.9)	
Opposite conclusion	1			89				70			78.7			78.7 (N/A)	
No conclusion possible	3			22				13			59.1			53.8 (26.9-64.4)	

Abbreviations: Incl., References included by reviews; IQR, interquartile range; N/A, not applicable (only 1 review).

Table 5. Characteristics of the 326 individual study references not found when limiting database searches to MEDLINE, Embase and/or CENTRAL

Characteristics	Of the 326:	Not indexed	Indexed	<1991	No abstract in database	Not English	Psychological/ Educational	Nutrition/ Physical activity	Pharma-cological	Complex	Other	Cerebro-vascular:	Chronic respi-ratory	CVD	Mental health	Osteo-arthritis
	Not indexed	*91 (28%)*	N/A	7	91	3	22	11	34	14	10	1	22	37	26	5
	Indexed	N/A	*235 (72%)*	73	71	2	4	145	51	14	21	2	30	174	23	6
	<1991	8%	31%	80 (25%)	39	0	1	43	16	1	19	0	20	54	5	1
	No abstract in database	100%	30%	49%	*162 (50%)*	4	22	53	50	23	14	2	28	92	34	6
	Not English	3%	1%	0%	2%	5 (2%)	1	1	2	0	1	0	1	2	1	1
Intervention type	Psychological/ Educational	2%	2%	1%	14%	20%	*26 (8%)*	N/A	N/A	N/A	N/A	0	0	7	18	1
	Nutrition/ Physical activity	12%	62%	54%	33%	20%	N/A	*156 (48%)*	N/A	N/A	N/A	0	8	140	6	2
	Pharmacological	37%	22%	20%	31%	40%	N/A	N/A	*85 (26%)*	N/A	N/A	3	16	50	12	4
	Complex	15%	6%	1%	14%	0%	N/A	N/A	N/A	*28 (9%)*	N/A	0	2	13	13	0
	Other	11%	9%	24%	9%	20%	N/A	N/A	N/A	N/A	*31 (10%)*	0	26	1	0	4
Topic	Cerebrovascular:	1%	1%	0%	1%	0%	0%	0%	4%	0%	0%	3 (1%)	N/A	N/A	N/A	N/A
	Chronic respiratory	24%	13%	25%	17%	20%	0%	5%	19%	7%	84%	N/A	52 (16%)	N/A	N/A	N/A
	CVD	41%	74%	68%	57%	40%	27%	90%	59%	46%	0%	N/A	N/A	211 (65%)	N/A	N/A
	Mental Health	29%	10%	6%	21%	20%	69%	4%	14%	46%	0%	N/A	N/A	N/A	49 (15%)	N/A
	Osteoarthritis	6%	3%	1%	4%	20%	4%	1%	5%	0%	13%	N/A	N/A	N/A	N/A	11 (3%)

Italics: points where row and column are identical, highlighted for readability. Absolute numbers of studies are shown in the upper right section, correlating percentages are shown in the lower left section.

To specifically assess the nature of these references, we removed 30 exact duplicates, i.e., references that were included in two different Cochrane reviews [18,19] or twice in the same Cochrane review [18,20]. Of the 326 individual references not found, 91 (28%) were not indexed in MEDLINE, Embase or CENTRAL. These were mainly grey literature (36/91 conference abstracts, 14/91 study registers, 12/91 theses, and 9/91 other types) and some journal articles (20/91, 22%). Of the remaining 235 individual, not found, indexed references, the majority had an abstract (70%), and was written in English (99%). The medical topics were mainly cardiovascular diseases (74%) and the intervention types were nutrition or physical activity (62%) (Table 5).

In comparison to the references that could be found with at least one of the limited database search approaches, the references that were not found were on average older (28% vs. 17% published before 1991) and more often had no abstract available (30% vs. 11%) (Appendix 5).

5. Discussion

5.1. Overall results

The majority of references included in a random sample of 60 Cochrane reviews were indexed in at least one of the databases MEDLINE, Embase, and CENTRAL. We found a median coverage of 94%-99% for every combination of at least two databases. As expected, searching in more databases improved recall, but not all indexed references could be found. The median recall was 85%-90% for every combination of at least two databases. In comparison, median coverage of a single database ranged between 83% and 91%, and median recall between 71% and 76%. The overall database coverage and recall observed in our study are similar to findings of previous research [7−9].

We could not conclusively answer the question on how many databases should be searched. Although searching two databases may be enough for some systematic reviews, others may require a more thorough approach like MECIR. However, we identified some factors that seem to play a role: When we stratified according to pharmacological and nonpharmacological interventions, we found that although median coverage of studies with pharmacological interventions was slightly higher than with nonpharmacological interventions in Embase, and CENTRAL, all databases had a higher mean and median recall of studies with pharmacological interventions than of studies with nonpharmacological interventions. This could indicate that successfully finding studies may depend on the intervention-type or topic: pharmacological interventions can be well-described and delimited by using the drugs' generic names, trade names, and codes. Nonpharmacological interventions are often more complex (e.g., psychological interventions) and the multiplicity of possible terms may be more difficult to capture in a sensitivity-specificity balanced search strategy. Depending on the topic, this could be especially aggravated by missing granularity or even gaps in subject headings [21], which gives reason to carefully pick suitable electronic databases and subject headings [22] and/or consult an information specialist [1,23,24].

We found that references of Cochrane reviews where the conclusion was no longer possible tended to have lower median database coverage than the others, although Cochrane reviews where the conclusion and its certainty remained the same had a markedly higher median coverage than those where changes occurred. However, on an individual level, there is no consistent pattern of low coverage and conclusion (Appendix 4). In two instances [14,25], all references that were indexed were also found (coverage = recall), but the conclusion was still not possible. In such cases it would have been necessary to find the nonindexed references. As our previous study has shown, combining the database search with reference list checking would already have helped to accurately determine conclusions [11]. Hence, researchers should consider using supplementary search techniques that are less dependent on indexation, such as citation tracking or contacting experts in the field [7,26,27].

Almost a fifth of all references were not found when using any of the seven search approaches in a limited number of databases. It may not always be necessary to find all the relevant references do draw a ''correct'' conclusion. However, we do not know when the necessary studies are found and measures should be taken to optimize the search and tailor the search approach to the research topic. A third of the references in our study that were indexed in at least one database but were not found had no abstract. As most databases only allow searching in the metadata of a study (especially in title, abstract, author keywords and subject headings) rather than the full text, a reference without an abstract is harder to find and more dependent on the use of subject headings such as MeSH and Emtree. To find such articles, one would have to improve the performance of searches. This includes database choice and supplementary search techniques. Another aspect is the quality and adequacy of the searches. Amongst others, these can be improved by using text analysis to inform search strategies [28], using systematic approaches to build comprehensive searches [29,30], and peer review of search strategies [31].

5.2. Limitations

This study has some limitations. First, the relationship between recall and conclusions is not straightforward. As Cochrane reviews list all included study publications, it is possible that references not found by the searches belonged to studies included in the narrative syntheses, not the meta-analyses. In that case, they affected the recall of the searches but not the conclusions of the review. A future analysis could stratify results by a Cochrane review's ''primary included study''. In addition, we calculated the recall independently of the database coverage. This means that even a search strategy that found all indexed references

would have a low recall if the review contained many non-indexed references.

Second, when determining if records were indexed in the databases, we limited results to MEDLINE-only and Embase-only, to ensure the results were independent of the search interface. However, this might differ from the actual search experience as some interfaces index additional documents, e.g., PubMed includes MEDLINE-in-Process materials and nonMEDLINE records, and Embase.com includes most MeSH-indexed MEDLINE records. Hence, we may have effectively underestimated the coverage and recall of searching in a limited number of databases. Conversely, by using references from Cochrane reviews, we might overestimate the database coverage of CENTRAL because references to RCTs included in Cochrane reviews are regularly added to this database. We sought to counter this issue by verifying that the references were available before the date of the latest search. However, this does not account for updated reviews, where some references might have been added after the publication of the original review and then picked up by the update searches.

Third, most Cochrane reviews' search strategies are designed by trained information specialists who follow the Cochrane handbook [1] and MECIR standards [2]. We assumed that these would be comprehensive and of high quality to serve as gold-standard [30]. Hence, we only checked the search strategies for formal errors (i.e., syntax, spelling, block building), not for comprehensiveness of search terms or overall quality. Although a low recall could relate to suboptimal searches [32], it is also possible that a good quality search was designed to be specific and hence, a lower recall would have been a deliberate choice.

6. Conclusion

Database coverage alone cannot predict whether relevant references will be found (=recall) by a search strategy. Cumulative search results of two or more databases improve coverage and recall and decrease the chances of making an inappropriate conclusion in a review. Yet, even the combination of the three most frequently used databases Medline, Embase, and CENTRAL did not suffice to achieve total recall or avoid any change in the conclusions of systematic reviews. For some research topics a small number of database searches combined with supplementary search methods may be more useful than searching many bibliographic databases. However, this approach needs to be empirically tested in future research.

CRediT authorship contribution statement

Hannah Ewald: Writing − original draft, Writing − review & editing, Project administration. **Irma Klerings:** Conceptualization, Investigation, Writing − original draft, Writing − review & editing, Project administration. **Gernot Wagner:** Writing − review & editing. **Thomas L. Heise:** Writing − review & editing. **Jan M. Stratil:** Writing − review & editing. **Stefan K. Lhachimi:** Writing − review & editing. **Lars G. Hemkens:** Writing − review & editing. **Gerald Gartlehner:** Conceptualization, Writing − review & editing. **Susan Armijo-Olivo:** Writing − review & editing. **Barbara Nussbaumer-Streit:** Conceptualization, Writing − review & editing.

Acknowledgments

We thank everyone who was involved in the main project for their contributions indirectly help us realize this project.

Appendix A

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2022.05.022.

References

[1] Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane handbook for systematic reviews of interventions version 6.2 (updated February 2021). 2021. Available at www.training.cochrane.org/handbook.2021. Accessed January 20, 2022.

[2] Higgins JPT, Lasserson T, Chandler J, Tovey D, Thomas J, Flemyng E, et al. Methodological expectations of Cochrane intervention reviews. London: Cochrane; 2022.

[3] Hartling L, Featherstone R, Nuspl M, Shave K, Dryden DM, Vandermeer B. The contribution of databases to the results of systematic reviews: a cross-sectional study. BMC Med Res Methodol 2016; 16:127.

[4] Halladay CW, Trikalinos TA, Schmid IT, Schmid CH, Dahabreh IJ. Using data sources beyond PubMed has a modest impact on the results of systematic reviews of therapeutic interventions. J Clin Epidemiol 2015;68:1076−84.

[5] Frandsen TF, Eriksen MB, Hammer DMG, Christensen JB. PubMed coverage varied across specialties and over time: a large-scale study of included studies in Cochrane reviews. J Clin Epidemiol 2019;112:59−66.

[6] Marshall IJ, Marshall R, Wallace BC, Brassey J, Thomas J. Rapid reviews may produce different results to systematic reviews: a meta-epidemiological study. J Clin Epidemiol 2019;109:30−41.

[7] Aagaard T, Lund H, Juhl C. Optimizing literature search in systematic reviews - are MEDLINE, EMBASE and CENTRAL enough for identifying effect studies within the area of musculoskeletal disorders? BMC Med Res Methodol 2016;16:161.

[8] Bramer WM, Giustini D, Kramer BM. Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and Google Scholar: a prospective study. Syst Rev 2016;5:39.

[9] Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. Syst Rev 2017;6:245.

[10] Levay P, Raynor M, Tuvey D. The contributions of MEDLINE, other bibliographic databases and various search techniques to NICE public health guidance. Evid Based Libr Inf Pract 2015;10:50−68.

[11] Nussbaumer-Streit B, Klerings I, Wagner G, Heise TL, Dobrescu AI, Armijo-Olivo S, et al. Abbreviated literature searches were viable

alternatives to comprehensive searches: a meta-epidemiological study. J Clin Epidemiol 2018;102:1−11.
[12] Nussbaumer-Streit B, Klerings I, Wagner G, Titscher V, Gartlehner G. Assessing the validity of abbreviated literature searches for rapid reviews: protocol of a non-inferiority and meta-epidemiologic study. Syst Rev 2016;5:197.
[13] Ewald H, Klerings I, Wagner G, Heise TL, Dobrescu AI, Armijo-Olivo S, et al. Abbreviated and comprehensive literature searches led to identical or very similar effect estimates: a meta-epidemiological study. J Clin Epidemiol 2020;128:1−12.
[14] Geretsegger M, Elefant C, Mössler Karin A, Gold C. Music therapy for people with autism spectrum disorder. Cochrane Database of Systematic Reviews. Cochrane Database Syst Rev 2014;2014: CD004381.
[15] McNamara DA, Goldberger JJ, Berendsen MA, Huffman MD. Implantable defibrillators versus medical therapy for cardiac channelopathies. Cochrane Database Syst Rev 2015;2015: CD011168.
[16] Wong GW, Wright JM. Blood pressure lowering efficacy of nonselective beta-blockers for primary hypertension. Cochrane Database Syst Rev 2014;28:CD007452.
[17] Blessberger H, Kammler J, Domanovits H, Schlager O, Wildner B, Azar D, et al. Perioperative beta-blockers for preventing surgery-related mortality and morbidity. Cochrane Database Syst Rev 2014; 2014:CD004476.
[18] Hooper L, Martin N, Abdelhamid A, Davey Smith G. Reduction in saturated fat intake for cardiovascular disease. Cochrane Database Syst Rev 2015;5:CD011737.
[19] Hooper L, Abdelhamid A, Bunn D, Brown T, Summerbell CD, Skeaff CM. Effects of total fat intake on body weight. Cochrane Database Syst Rev 2015:CD011834.
[20] Birks Jacqueline S, Chong Lee Y, Grimley Evans J. Rivastigmine for Alzheimer's disease. Cochrane Database Syst Rev 2015;2015: CD001191.
[21] Minguet F, Van Den Boogerd L, Salgado TM, Correr CJ, Fernandez-Llimos F. Characterization of the medical subject headings thesaurus for pharmacy. Am J Health Syst Pharm 2014;71:1965−72.
[22] Barroso J, Gollop CJ, Sandelowski M, Meynell J, Pearce PF, Collins LJ. The challenges of searching for and retrieving qualitative studies. West J Nurs Res 2003;25:153−78.
[23] Kugley S, Wade A, Thomas J, Mahood Q, Jørgensen A-MK, Hammerstrøm K, et al. Searching for studies: a guide to information retrieval for Campbell systematic reviews. Campbell Syst Rev 2017; 13:1−73.
[24] Rethlefsen ML, Murad MH, Livingston EH. Engaging medical librarians to improve the quality of review articles. JAMA 2014;312: 999−1000.
[25] Wong GW, Laugerotte A, Wright JM. Blood pressure lowering efficacy of dual alpha and beta blockers for primary hypertension. Cochrane Database Syst Rev 2015;2015:CD007449.
[26] Cooper C, Booth A, Britten N, Garside R. A comparison of results of empirical studies of supplementary search techniques and recommendations in review methodology handbooks: a methodological review. Syst Rev 2017;6:234.
[27] Vassar M, Yerokhin V, Sinnett PM, Weiher M, Muckelrath H, Carr B, et al. Database selection in systematic reviews: an insight through clinical neurology. Health Inf Libr J 2017;34:156−64.
[28] Hausner E, Guddat C, Hermanns T, Lampert U, Waffenschmidt S. Prospective comparison of search strategies for systematic reviews: an objective approach yielded higher sensitivity than a conceptual one. J Clin Epidemiol 2016;77:118−24.
[29] Bramer WM, de Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to searching: an efficient and complete method to develop literature searches. J Med Libr Assoc 2018;106:531−41.
[30] Hausner E, Guddat C, Hermanns T, Lampert U, Waffenschmidt S. Development of search strategies for systematic reviews: validation showed the noninferiority of the objective approach. J Clin Epidemiol 2015;68:191−9.
[31] 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.
[32] Franco JVA, Garrote VL, Escobar Liquitay CM, Vietto V. Identification of problems in search strategies in Cochrane Reviews. Res Synth Methods 2018;9:408−16.