



University of Southern Denmark

**PubMed coverage varied across specialties and over time
a large-scale study of included studies in Cochrane reviews**

Frandsen, Tove Faber; Eriksen, Mette Brandt; Hammer, David Mortan Grøne; Christensen, Janne Buck

Published in:
Journal of Clinical Epidemiology

DOI:
10.1016/j.jclinepi.2019.04.015

Publication date:
2019

Document version:
Accepted manuscript

Document license:
CC BY-NC-ND

Citation for published version (APA):
Frandsen, T. F., Eriksen, M. B., Hammer, D. M. G., & Christensen, J. B. (2019). PubMed coverage varied across specialties and over time: a large-scale study of included studies in Cochrane reviews. *Journal of Clinical Epidemiology*, 112, 59-66. <https://doi.org/10.1016/j.jclinepi.2019.04.015>

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use

This work is brought to you by the University of Southern Denmark.
Unless otherwise specified it has been shared according to the terms for self-archiving.
If no other license is stated, these terms apply:

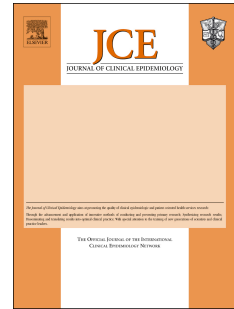
- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.
Please direct all enquiries to puresupport@bib.sdu.dk

Accepted Manuscript

PubMed coverage varied across specialties and over time: a largescale study of included studies in Cochrane reviews

Tove Faber Frandsen, Mette Brandt Eriksen, David Mortan Grøne Hammer, Janne Buck Christensen



PII: S0895-4356(19)30083-6

DOI: <https://doi.org/10.1016/j.jclinepi.2019.04.015>

Reference: JCE 9866

To appear in: *Journal of Clinical Epidemiology*

Received Date: 30 January 2019

Revised Date: 10 April 2019

Accepted Date: 24 April 2019

Please cite this article as: Frandsen TF, Eriksen MB, Grøne Hammer DM, Christensen JB, PubMed coverage varied across specialties and over time: a largescale study of included studies in Cochrane reviews, *Journal of Clinical Epidemiology* (2019), doi: <https://doi.org/10.1016/j.jclinepi.2019.04.015>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

PubMed coverage varied across specialties and over time: a largescale study of included studies in Cochrane reviews

Tove Faber Frandsen¹, Mette Brandt Eriksen², David Mortan Grøne Hammer^{3,4}, Janne Buck Christensen⁵

¹ University of Southern Denmark, Department of Design and Communication, Universitetsparken 1, DK-6000 Kolding, Denmark

² The University Library of Southern Denmark, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark.

³ University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

⁴ Goethe-Universität, Robert-Mayer-Straße 11-15, 60325 Frankfurt am Main, Germany

⁵ South-West Jutland Hospital, Department of Quality and Improvement, Finsensgade 35, 6700 Esbjerg

Abstract

Objective: PubMed is one of the most commonly used search tools in Biomedical and Life Sciences. Existing studies on database coverage generally conclude that searching PubMed may not be sufficient although some find that the contributions from other databases are modest at best. However, generalizability of the studies of the coverage of PubMed is typically restricted. The objective of this study is to analyze the coverage of PubMed across specialties and over time.

Study Design and Setting: We use the more than 50,000 included studies in all Cochrane reviews published from 2012 to 2016 as our population and examine if the studies as well as resulting publications can be identified in PubMed.

Results: The results show that PubMed has a coverage of 70.9, 95% CI [68.40,73.30] of all the included publications and 82.8%, 95% CI [80.9, 84.7]. There are huge differences in coverage across as well as within specialties. In addition, coverage varies within groups over time.

Conclusion: Databases used for searching topics within the groups with highly varying or low coverage should be chosen with care as PubMed may have a relatively low coverage.

Keywords: PubMed, bibliographic databases, information storage and retrieval, systematic reviews, Cochrane,

Running title: PubMed coverage varies across specialties and over time

1. Introduction

Systematic reviews (SR) and meta-analyses are essential to summarize evidence on health care interventions on which conclusions can be drawn and decisions made. One of the key characteristics of a systematic review is that the researchers perform a systematic search, which seeks to identify all studies that would meet the eligibility criteria for the research questions examined. Identifying the relevant studies typically involves searching in bibliographic databases, although not all studies will be identifiable in traditional databases and some need to be located using alternative strategies (1, 2). Failing to identify all studies may introduce bias.

Several factors play a role in the quality of searches for primary studies to be included in SRs (3). Involvement of an information specialist or librarian in the search process seems to play a role for the quality of searches (4, 5). Choice of database and searching multiple databases must be considered intertwined and can be formulated as a question of which databases to search and how many (6-12).

PubMed is one of the most commonly used search tools in Biomedical and Life Sciences. However, a search of Medline or PubMed alone is generally not considered adequate, and the number of searched sources is generally increasing (13). However, being identifiable is not the same as being retrieved as relevant publications in a database may fail to be found due to e.g. poor searching or poor indexing. This also supports the need for several databases. Existing recommendations regarding choice of databases are typically either very limited (14) or tied to specific topics (15-17), although exceptions do exist (18).

Database coverage can be investigated using different approaches and there a large body of relevant literature exist dealing with the consequences of the specific choice of searched databases. Typically, such studies use a specific topic or medical specialty as example (19, 20). This means that the results of the searches depend on the coverage of the specific topic or specialty in the searched databases. Consequently, choice of database is closely related to coverage which in turn can be described both in terms of width (number of indexed journals) and depth (number of indexed articles from a journal) (21). Another approach is to use document types as a starting point (12) or a sample collected using handsearching (22). Finally, several of the recent studies use a gold standard to form the basis for an examination of database coverage (23-26) which ensures that relevance has been assessed for a specific purpose being either for systematic reviews, meta-analyses or guidelines.

The literature on database coverage generally concludes that searching PubMed may not be sufficient and that other databases can contribute with relevant publications (6, 27-29). However, some studies argue that the contributions from other databases are modest at best (15, 23). An effect of adding another database to the already searched databases in a systematic review is found to be greater in the case of meta-analyses with 10+ included studies (20). Being widely used in health research a number of studies focus on the specific coverage of PubMed or Medline. A systematic search retrieved a number of relevant studies regarding coverage of PubMed or Medline (search histories are available in appendix 1). The references in relevant studies were also scrutinized and the studies formed the basis for a supplementing citation search. Only studies reporting coverage of PubMed or Medline is reviewed whereas studies only reporting merged coverage is excluded (e.g. 30). In the existing literature we find a great number of case studies within a wide variety of disciplines and topics: maternal morbidity and mortality (31), family medicine (32), cardiovascular disease (15), telemedicine (33), adverse effects (34, 35), hypertension (19), injury prevention and safety promotion (36), frozen shoulder (10), medical imaging (37), renal literature (38), orthopedic surgery (28), exercise therapy (39), rehabilitation (40),

clinical neurology (3) and diagnostic accuracy of depression screening tools (41). These case studies of coverage find highly varying coverage ranging from less than 50% (34, 39, 40) to over 90% (15, 28, 37, 38, 41). Similarly, a systematic review finds estimates of recall in Medline varying substantially (9).

Furthermore, studies on coverage of PubMed or Medline can focus on language or geographic issues such as coverage of Chinese (8) or German literature (42). In both cases coverage is found to be low (i.e. less than 50%). The coverage of PubMed has also been investigated across different topics or fields using a convenience sample (18, 23, 43), a selection based on purposive sampling (25, 26) or a random sample of reviews as starting point (20). The samples sizes in the aforementioned studies range from 21 reviews (25) to 129 reviews (26) and the resulting number of included publications range from 541 (25) to 4795 (43). The results of these studies of coverage indicate that considerable variation is found depending on topic (23) and coverage of Medline ranges from 73 to 85%.

Summing up, the existing studies typically restrict to one or a few topics which seems to limit generalizability. As stated in a study of 129 meta-analyses: “Future research across different topics will provide additional evidence upon which to base recommendations for searching in evidence reviews” (26, p 9 of 13). The aim of the present study is thus to provide insight in to the coverage of PubMed across several subfields to answer the following question: What is the share of included publications and studies in Cochrane reviews available in PubMed?

2. Methods

The population in this study is all the included studies in all Cochrane reviews from 2012 to 2016. The included studies are extracted from the Cochrane Library separated into Cochrane groups. In some cases, there are several publications generated from a study. For the present analysis, both units of analysis (i.e. study and publication) are of interest, as we examine (i) whether all publications are available in PubMed, and (ii) whether at least one publication per study is available in PubMed. Consequently, we measure the availability at both levels. Studies and the publications stemming from these studies are extracted from the Cochrane group overview page (44) using the Python programming language. The data extracted was subsequently filtered by inclusion criteria (such as publication year).

The present study examines how many of the included publications and studies in these reviews we can identify in PubMed. For the present study, a distinction between Medline and PubMed is essential. PubMed contains over 29 million references coming from various sources but the vast majority of the references (> 25 million) are from Medline. In addition to the content from Medline, PubMed also includes books and book chapters available in NCBI Bookshelf, as well as PubMed Central® (PMC), an electronic archive of freely available biomedical articles. PMC contains freely available articles from journals that are included in the PMC based on an assessment by National Library of Medicine (NLM). In addition, the PMC also contains articles from the National Institute of Health (NIH) funded researchers, as the PMC acts as an archive of manuscripts published in connection with NIH funded research (45). Cochrane recommends Medline and thus Medline coverage is of great importance, however, PubMed has a higher sensitivity than Medline (46) and many researchers as well as clinicians use PubMed. Consequently, in this study the coverage of PubMed is examined, not Medline.

The Cochrane Library does include direct links to PubMed if the article is available in PubMed. The metadata was used to determine whether a given publication is indexed in PubMed. For some

publications, the reference in the Cochrane review will have a direct link to PubMed with the relevant PMID. Direct links are available in Cochrane for 49.1% of all included publications for the relevant reviews in this study. However, as the links may not necessarily direct the reader to the correct publication in PubMed all links were verified matching title and journal name. This revealed matching problems in 1.2% of the included publications which amounts to 2.4% of the links. In many of the cases the link is indeed correct, and the problem is caused by spelling variations, omitted subtitles or use of journal abbreviations instead of full title. Furthermore, some of the matching problems are caused by publications that are indexed in PubMed but not linked correctly. Finally, some of the incorrect links are due to incorrect links with no correct match available in PubMed. This latter group of links, covering 335 publications (0.4%), was subsequently manually corrected in the dataset.

A preliminary examination of the links also revealed that some of the studies without a link in the Cochrane Library to PubMed are nevertheless available in the database. Consequently, the existence of a link is not sufficient to determine availability in PubMed. Consequently, to lower the number of false negatives due to missing links, we created queries based on titles and years automatically for the remaining publications, allowing for manual verification. In the cases, where the search resulted in several results, we narrowed down the search until we identified the correct reference if present.

The data collection took place from January 2018 to December 2018.

3. Results

The overall coverage from 2012 to 2016 in PubMed is 61,052 publications out of 86,171 corresponding to a coverage rate of 70.9, 95% CI [68.40,73.30]. Figure 1 provides an overview of the coverage rate by publication year from 1970 to 2015 (but we include all publication years are in the analyses below). The figure shows that coverage is more is slightly higher for older publications, although the coverage is relatively stable for publications from the last two decades.

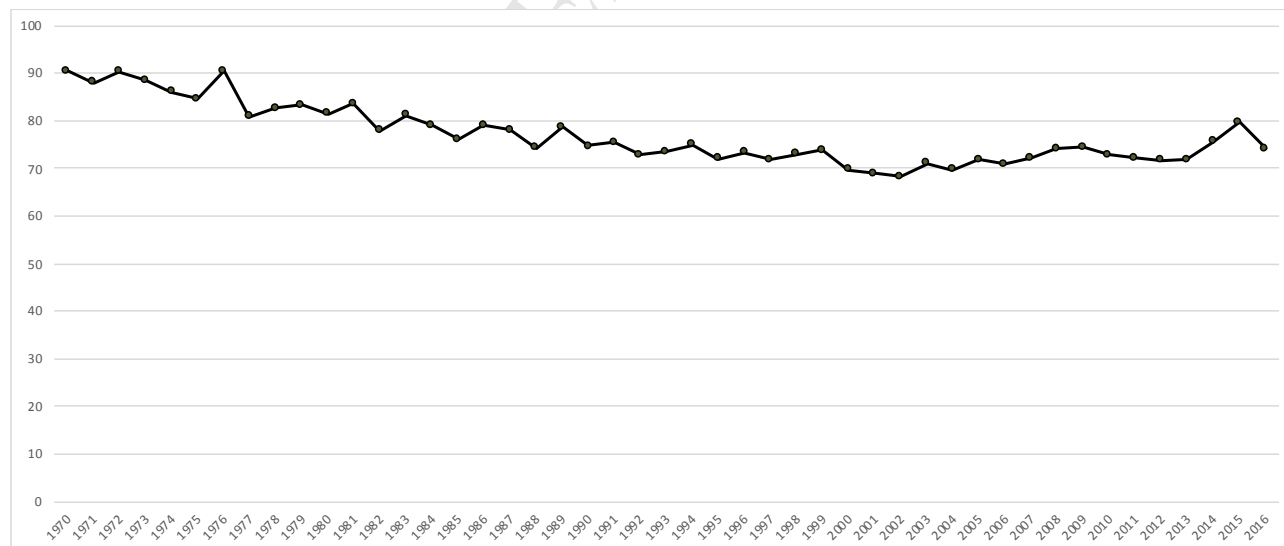


Figure 1. Publication year of included publications in Cochrane reviews 2012-2016 and the average coverage of PubMed.

Figure 1 excludes the publications which have no publication year, and as PubMed has a lower coverage of these (e.g. data sets and grey publications) the average coverage in this figure is higher than in the entire dataset (73.0% instead of 70.9%).

Splitting the total coverage of PubMed into review years and sub-disciplinary groups reveals that there are huge differences in coverage among the groups as well as within the groups. Figure 2 depicts the mean coverage of publications of each group as well as the minimum and maximum coverage over the years 2012-2016. The circle marks the average whereas the length of the line marks the 95 percent confidence intervals (CI). The overall mean coverage is marked with a red line and confidence intervals (95 percent) with a dotted line.

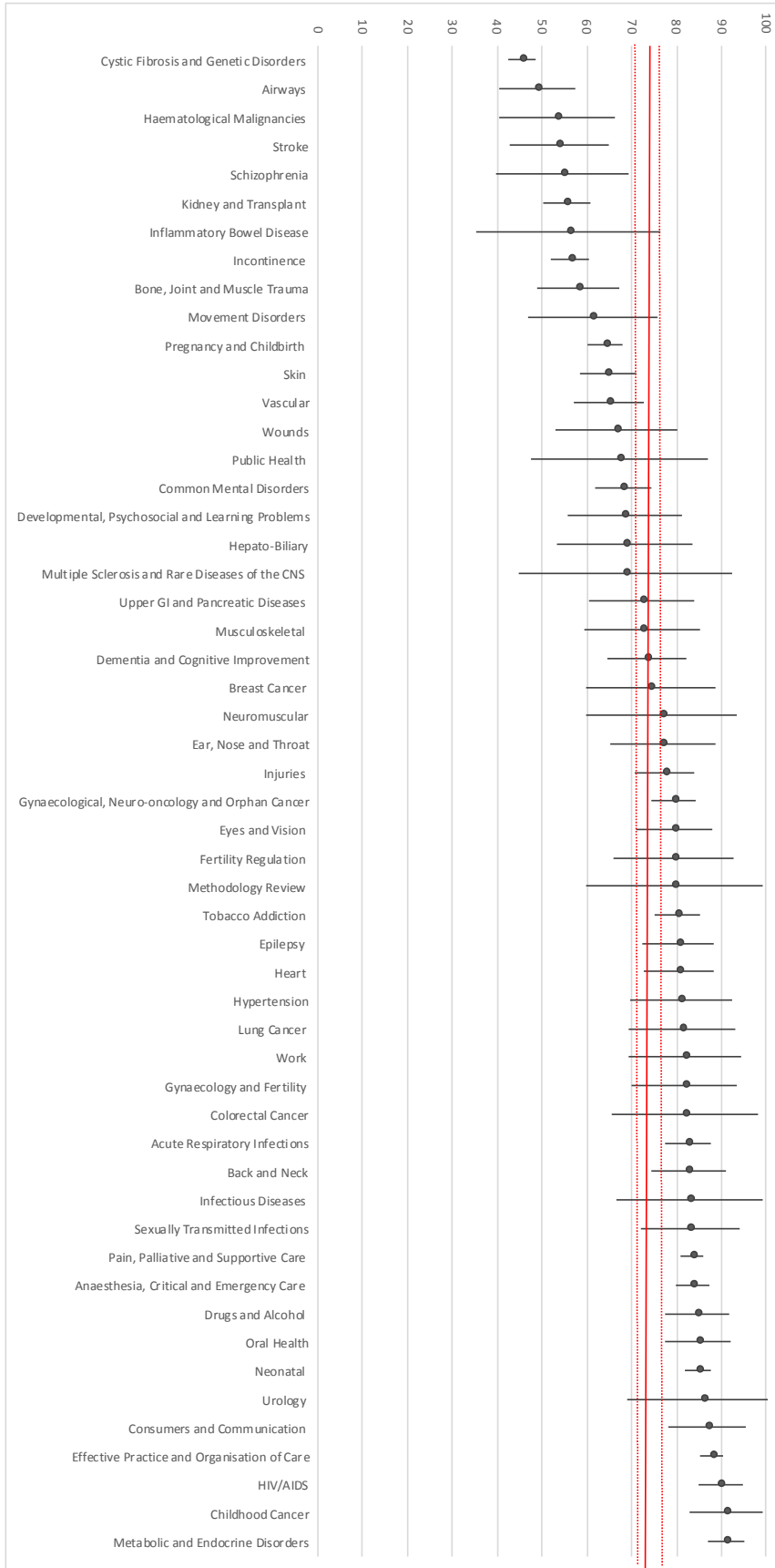


Figure 2. PubMed coverage of publications

From Figure 2, we can see that the mean coverage of the included publications in PubMed ranges from less than 50 percent for Cystic Fibrosis and Genetic Disorders and Airways to more than 90 percent for Metabolic and Endocrine Disorders and Childhood Cancer. If we look at individual years mean coverage exhibits even wider ranges from less than 40 percent to 100 percent. The variations within groups are also noticeable. Some groups vary very little over time (e.g. less than 5 percentage points) whereas other groups can range considerably (e.g. over 50 percentage points).

Figure 3 depicts the mean coverage rate of studies as well as the minimum and maximum coverage over the years 2012-2016. PubMed covers 45,483 out of 54,903 studies and thus has a coverage rate of 82.8%, 95% CI [80.9, 84.7]. Again, the coverage of studies depicts large differences across groups (with the coverage rate ranging from a little over 40 percent to more than 95 percent).

Since only one publication from each study needs to be available in PubMed to count in this analysis, coverage for studies is higher than for all publications. Furthermore, we see that the review group on Cystic Fibrosis and Genetic Disorders, which had the lowest coverage at the publication level, is much closer to the average on study level – and similar changes in the rank of groups can be seen for other groups as well. Consequently, the coverage of PubMed within each group differs also depending on the unit of analysis (publications or studies).

Another repeated result is that PubMed coverage varies greatly within some groups whereas other groups have a much more stable coverage by PubMed. The Tobacco Addiction Group has a PubMed coverage of included studies of more than 90 percent and coverage seems to be stable over time. Within this group, PubMed thus covers the large majority of the included studies. It is another matter altogether with the Musculoskeletal Group, where the coverage rate varies from a little over 40 percent to 90 percent indicating that coverage for this group can be difficult to predict.

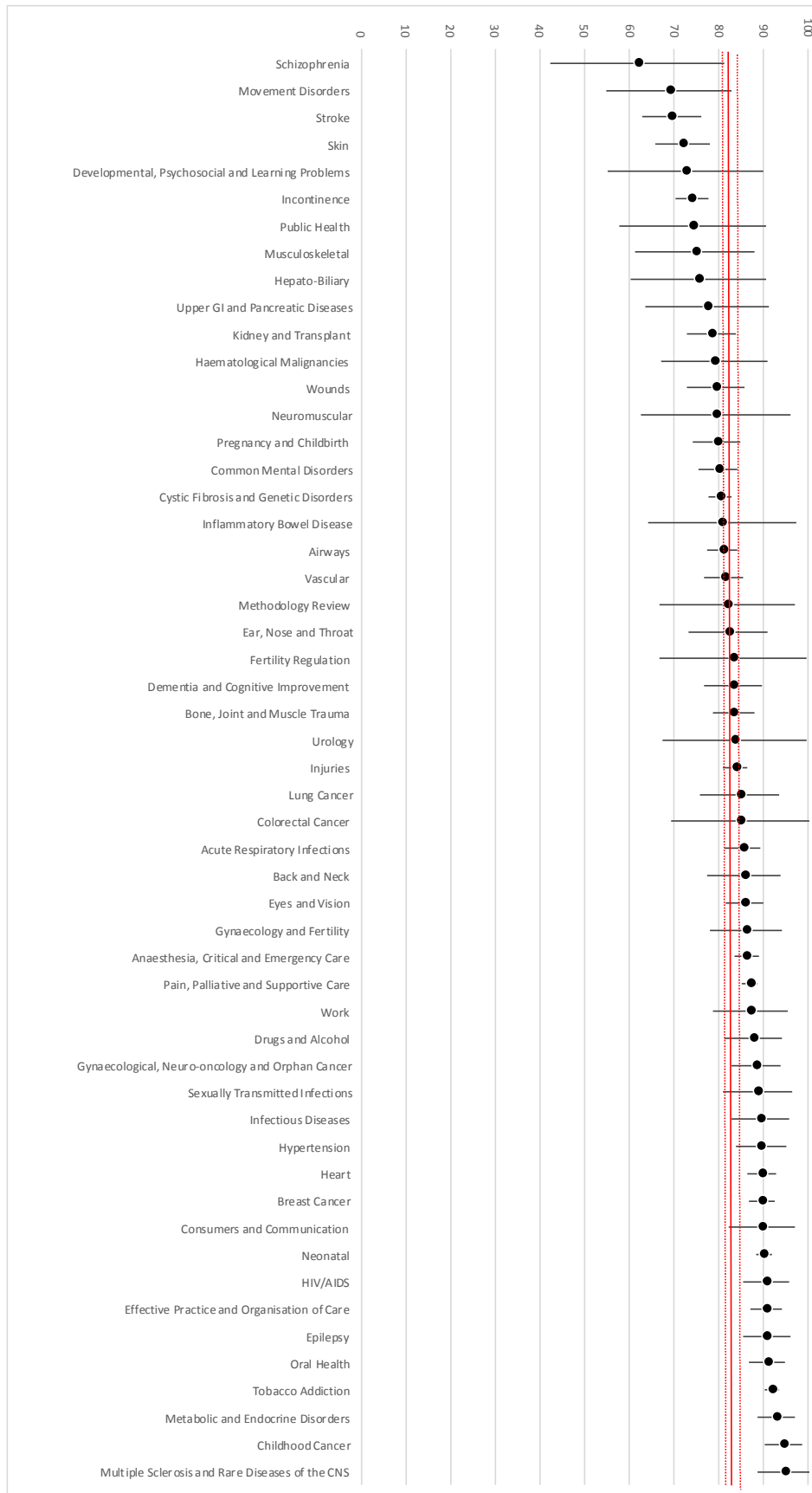


Figure 3. PubMed coverage of studies

4. Discussion and conclusion

The results show that PubMed has a coverage of 71 percent of all the included publications and 83 percent of all included studies in Cochrane reviews from 2012 to 2016. Few of the existing case studies can be directly compared to our results as the topics and fields are defined differently. A few examples are worth mentioning: comparing the results found in this study regarding coverage of PubMed to the existing literature, we find a higher share of the publications within the area of hypertension than an earlier study (19) but as the inclusion criteria differ the results are not directly comparable. Within Kidney and Transplant group we find a somewhat lower coverage, however, as coverage within this areas differs in terms of publication types this may explain the differences (38).

Splitting the total coverage of PubMed into review years and groups confirms that there are huge differences among the groups as well as within the groups. The mean coverage in PubMed in this study ranges from less than 40 percent to more than 95 percent of the included publications, and from a little over 40 to more than 95 percent for the included studies. Variations across reviews have been shown to be considerable in 50 systematic reviews of therapeutic interventions (20) as well as in 129 reviews from three Cochrane groups (26). However, being based on small samples, their results cannot be generalized which they both acknowledge. The existing studies of coverage across different fields use samples consisting of 541 to 4795 included publications and find a somewhat higher coverage of Medline and PubMed. However, as there are considerable variations across groups, the specific choice of topics in these smaller studies will affect the results. This is confirmed in the studies indicating confidence intervals where we can see a considerable variation among reviews (18, 20, 43).

Furthermore, the results of this study show that the coverage of PubMed within each group differs depending on the unit of analysis being publications or studies. The existing studies using reviews as a means of sampling distinguish between reviews and included references (18, 20, 43).

Finally, this study shows remarkable differences within groups over time. Few studies of coverage of PubMed or Medline includes variations over time. A study using handsearching to sample references does includes variations in coverage over time (22) but only in three aggregate groups covering almost three decades.

Limitations of this study should also be considered. First, we examined whether the included studies could be found in PubMed and our results are thus based on whether the studies were present in the databases, but do not reflect the ability of searchers to find those references. Poor searchers may find a great deal less than what is available, so the coverage rates we report can be seen as an upper limit of the share of studies, researchers may find. Lacking subject headings or abstracts can also make it extremely difficult to retrieve a particular publication.

Second, we focus exclusively on PubMed, and there are many other relevant databases depending on the topic (26), although it is fair to say that PubMed is the most widely used and known database. Nevertheless, whether a study or a publication is available in PubMed may be less important if the majority of these studies or publications can be retrieved in other key databases. We are also not considering the difference between PubMed and Medline in this study. Cochrane recommends Medline and thus Medline coverage is of great importance, however, many researchers as well as clinicians use PubMed.

Finally, if a Cochrane group includes a large number of publications not indexed by Pubmed (e.g. grey publications) we are likely to find lower shares of retrievable publications in PubMed. This would largely be due to thorough searches performed in grey sources by the review team and is likely less related to PubMed coverage per se.

5. Conclusion

This large-scale study finds huge differences across groups in terms of coverage in PubMed. PubMed coverage differs from group to group but also within groups. Changes in coverage from one year to another is likely caused by the different review topics. Some groups vary very little over time (e.g. less than 5 percentage points) whereas other groups can range considerably (e.g. over 50 percentage points), possibly due to the fact that the specific subjects of the Cochrane reviews may vary a lot, even though they come from the same group. Coverage also depends on whether we use publications or studies as unit of analysis.

Based on this study we conclude that coverage can be very difficult for some groups to predict and thus researchers should choose databases for searches within these areas with care. Furthermore, future studies of database coverage need to consider the variation in coverage as well as account for differences in coverage depending on whether coverage is measured on study or publication level.

6. References

1. McManus RJ, Wilson S, Delaney BC, Fitzmaurice DA, Hyde CJ, Tobias RS, et al. Review of the usefulness of contacting other experts when conducting a literature search for systematic reviews. *British Medical Journal*. 1998;317(7172):1562-3.
2. Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. *Cochrane database of systematic reviews (Online)*. 2011(8):MR000026.
3. Vassar M, Yerokhin V, Sinnott PM, Weiher M, Muckelrath H, Carr B, et al. Database selection in systematic reviews: an insight through clinical neurology. *Health information and libraries journal*. 2017;34(2):156-64.
4. Rethlefsen ML, Farrell AM, Osterhaus Trzasko LC, Brigham TJ. Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. *J Clin Epidemiol*. 2015;68(6):617-26.
5. Koffel JB. Use of recommended search strategies in systematic reviews and the impact of librarian involvement: a cross-sectional survey of recent authors. *PLoS One*. 2015;10(5):e0125931.
6. Sampson M, Barrowman NJ, Moher D, Klassen TP, Pham B, Platt R, et al. Should meta-analysts search Embase in addition to Medline? *J Clin Epidemiol*. 2003;56(10):943-55.
7. Jones CW, Keil LG, Weaver MA, Platts-Mills TF. Clinical trials registries are under-utilized in the conduct of systematic reviews: A cross-sectional analysis. *Systematic reviews*. 2014;3(1).
8. Cohen JF, Korevaar DA, Wang J, Spijker R, Bossuyt PM. Should we search Chinese biomedical databases when performing systematic reviews? *Systematic reviews*. 2015;4(1).
9. Crumley ET, Wiebe N, Cramer K, Klassen TP, Hartling L. Which resources should be used to identify RCT/CCTs for systematic reviews: a systematic review. *BMC Med Res Methodol*. 2005;5:24.
10. Beyer FR, Wright K. Can we prioritise which databases to search? A case study using a systematic review of frozen shoulder management. *Health information and libraries journal*. 2013;30(1):49-58.
11. Kwon Y, Powelson SE, Wong H, Ghali WA, Conly JM. An assessment of the efficacy of searching in biomedical databases beyond MEDLINE in identifying studies for a systematic review on ward closures as an infection control intervention to control outbreaks. *Systematic reviews*. 2014;3:135.

12. Suarez-Almazor ME, Belseck E, Homik J, Dorgan M, Ramos-Remus C. Identifying clinical trials in the medical literature with electronic databases: MEDLINE alone is not enough. *Controlled clinical trials*. 2000;21(5):476-87.
13. Lam MT, McDiarmid M. Increasing number of databases searched in systematic reviews and meta-analyses between 1994 and 2014. *Journal of the Medical Library Association : JMLA*. 2016;104(4):284-9.
14. Higgins JP, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*: Cochrane Book Series 2008. 1-649 p.
15. Bayliss SE, Davenport CF, Pennant ME. Where and how to search for information on the effectiveness of public health interventions - a case study for prevention of cardiovascular disease. *Health information and libraries journal*. 2014;31(4):303-13.
16. Rollin L, Darmoni S, Caillard JF, Gehanno JF. Searching for high-quality articles about intervention studies in occupational health - What is really missed when using only the Medline database? *Scand J Work Environ Health*. 2010;36(6):484-7.
17. Whiting P, Westwood M, Burke M, Sterne J, Glanville J. Systematic reviews of test accuracy should search a range of databases to identify primary studies. *J Clin Epidemiol*. 2008;61(4):357-64.
18. Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Systematic reviews*. 2017;6:245.
19. Rathbone J, Carter M, Hoffmann T, Glasziou P. A comparison of the performance of seven key bibliographic databases in identifying all relevant systematic reviews of interventions for hypertension. *Systematic reviews*. 2016;5:27.
20. 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(9):1076-84.
21. Jacso P. Database source coverage: hypes, vital signs and reality checks. *Online Inf Rev*. 2009;33(5):997-1007.
22. Hopewell S, Clarke M, Lusher A, Lefebvre C, Westby M. A comparison of handsearching versus MEDLINE searching to identify reports of randomized controlled trials. *Stat Med*. 2002;21(11):1625-34.
23. 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(1):50-68.
24. Beckles Z, Glover S, Ashe J, Stockton S, Boynton J, Lai R, et al. Searching CINAHL did not add value to clinical questions posed in NICE guidelines. *J Clin Epidemiol*. 2013;66(9):1051-7.
25. Bramer WM, Giustini D, Kramer BM, Anderson P. The comparative recall of Google Scholar versus PubMed in identical searches for biomedical systematic reviews: a review of searches used in systematic reviews. *Systematic reviews*. 2013;2:115.
26. 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(1):127.
27. Bahaadinbeigy K, Yogesan K, Wootton R. MEDLINE versus EMBASE and CINAHL for telemedicine searches. *Telemed J E Health*. 2010;16(8):916-9.
28. Slobogean GP, Verma A, Giustini D, Slobogean BL, Mulpuri K. MEDLINE, EMBASE, and Cochrane index most primary studies but not abstracts included in orthopedic meta-analyses. *J Clin Epidemiol*. 2009;62(12):1261-7.
29. Royle PL, Bain L, Waugh NR. Sources of evidence for systematic reviews of interventions in diabetes. *Diabet Med*. 2005;22(10):1386-93.
30. 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(1):1-11.
31. Betran AP, Say L, Gulmezoglu AM, Allen T, Hampson L. Effectiveness of different databases in identifying studies for systematic reviews: experience from the WHO systematic review of maternal morbidity and mortality. *BMC Med Res Methodol*. 2005;5(1):6.
32. Wilkins T, Gillies RA, Davies K. EMBASE versus MEDLINE for family medicine searches: can MEDLINE searches find the forest or a tree? *Canadian family physician Medecin de famille canadien*. 2005;51:848-9.
33. Ahmadi M, Sarabi RE, Orak RJ, Bahaadinbeigy K. Information retrieval in telemedicine: A comparative study on bibliographic databases. *Acta Informatica Medica*. 2015;23(3):172-6.

34. Golder S, Loke YK. The contribution of different information sources for adverse effects data. *Int J Technol Assess Health Care*. 2012;28(2):133-7.
35. Golder S, Wright K, Rodgers M. The contribution of different information sources to identify adverse effects of a medical device: a case study using a systematic review of spinal fusion. *Int J Technol Assess Health Care*. 2014;30(4):423-9.
36. Lawrence DW. What is lost when searching only one literature database for articles relevant to injury prevention and safety promotion? *Inj Prev*. 2008;14(6):401-4.
37. Berry E, Kelly S, Hutton J, Harris KM, Smith MA. Identifying studies for systematic reviews. An example from medical imaging. *Int J Technol Assess Health Care*. 2000;16(2):668-72.
38. Shariff SZ, Sontrop JM, Iansavichus AV, Haynes RB, Weir MA, Gandhi S, et al. Availability of renal literature in six bibliographic databases. *Clinical Kidney Journal*. 2012;5(6):610-7.
39. Stevinson C, Lawlor DA. Searching multiple databases for systematic reviews: Added value or diminishing returns? *Complement Ther Med*. 2004;12(4):228-32.
40. Minozzi S, Pistotti V, Forni M. Searching for rehabilitation articles on MEDLINE and EMBASE. An example with cross-over design. *Arch Phys Med Rehabil*. 2000;81(6):720-2.
41. Rice DB, Kloda LA, Levis B, Qi B, Kingsland E, Thombs BD. Are MEDLINE searches sufficient for systematic reviews and meta-analyses of the diagnostic accuracy of depression screening tools? A review of meta-analyses. *J Psychosom Res*. 2016;87:7-13.
42. Turp JC, Schulte J-M, Antes G. Nearly half of dental randomized controlled trials published in German are not included in Medline. *Eur J Oral Sci*. 2002;110(6):405-11.
43. 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. *Systematic reviews*. 2016;5:39.
44. Frandsen TF, Eriksen MB, Hammer DMG, Christensen JB. Fragmented publishing: a large-scale study of health science. Forthcoming in *Scientometrics*. 2019.
45. MEDLINE, PubMed, and PMC (PubMed Central): How are they different? 2019 [Available from: <https://www.nlm.nih.gov/bsd/difference.html>].
46. Katchamart W, Faulkner A, Feldman B, Tomlinson G, Bombardier C. PubMed had a higher sensitivity than Ovid-MEDLINE in the search for systematic reviews. *J Clin Epidemiol*. 2011;64(7):805-7.

Appendix 1. Search histories

Searches performed March 13, 2019.

Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily 1946 to March 12, 2019

1	exp PubMed/	6814
2	MEDLINE/	4763
3	exp MEDLARS/	5750
4	1 or 2 or 3	7800
5	exp "Review Literature as Topic"/	12007
6	exp "Information Storage and Retrieval"/	161237
7	4 and 5 and 6	259
8	((coverage* or recall* or performance*) adj5 (medline or Pubmed or medlars)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating	158

	sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	
9	7 or 8	409

Embase Classic+Embase 1947 to 2019 March 12

1	Medline/	85672
2	Review Literature as Topic.mp.	150
3	"systematic review (topic)"/	23091
4	consensus development/	22772
5	2 or 3 or 4	45843
6	information storage/	1582
7	information retrieval/	34022
8	(information storage or information retrieval).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]	37430
9	6 or 7 or 8	37430
10	1 and 5 and 9	98
11	((coverage* or recall* or performance*) adj5 (medline or Pubmed or medlars)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]	259
12	10 or 11	353

Highlights

- This study presents the results of an analysis of more than 85.000 publications from the 53 Cochrane groups
- PubMed covers more than 80 percent of all studies and 71 percent of publications included in Cochrane reviews.
- Coverage varies across specialties and within specialties over time.