

Using the full PICO model as a search tool for systematic reviews resulted in lower recall for some PICO elements

Frandsen, Tove Faber; Nielsen, Michael Friberg Bruun; Lindhardt, Christina Louise; Eriksen, Mette Brandt

Published in:
Journal of Clinical Epidemiology

DOI:
10.1016/j.jclinepi.2020.07.005

Publication date:
2020

Document version:
Accepted manuscript

Document license:
CC BY-NC-ND

Citation for published version (APA):

Frandsen, T. F., Nielsen, M. F. B., Lindhardt, C. L., & Eriksen, M. B. (2020). Using the full PICO model as a search tool for systematic reviews resulted in lower recall for some PICO elements. *Journal of Clinical Epidemiology*, 127, 69-75. <https://doi.org/10.1016/j.jclinepi.2020.07.005>

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

Using the full PICO model as a search tool for systematic reviews will result in lower recall for some PICO elements

Tove Faber Frandsen¹, University of Southern Denmark, Department of Design and Communication, t.faber@sdu.dk

Michael Friberg Bruun Nielsen, UCL University College, Odense, Denmark

Christina Louise Lindhardt, Department of Clinical Research, University of Southern Denmark, Odense, Denmark.

Mette Brandt Eriksen, University of Southern Denmark, University Library

Abstract

Objective: The use of the four-part PICO model to facilitate search strategy development for a precise answer is recommended for structuring searches for systematic reviews. Existing guidelines generally recommend that a search strategy should include the population, intervention(s), and types of study design. Consequently, comparison and outcome are not recommended as a part of the search strategy. There is evidence that comparison and particularly outcome is not represented in enough detail, but this needs to be confirmed.

Study Design: The present study examines the presence of PICO elements in the records in two commonly used databases for health sciences research: Embase and PubMed. We examine the field of upper GI and pancreatic diseases as well as the field of pregnancy and childbirth by extracting the included studies as well as the related PICO elements from a random selection of Cochrane reviews within these two areas.

Results: We find that the PICO elements C and O had a lower retrieval potential across the two Cochrane groups and databases also when combining text words and subject headings. In particular we find a lower retrieval when searching for both primary and secondary outcomes.

Conclusion: Our results support the existing recommendation not to search for outcomes.

Keywords: PICO, systematic searches, search tools, systematic reviews, search strategies, Embase, PubMed

¹ Corresponding author, address: University of Southern Denmark, Universitetsparken 1, DK-6000 Kolding, Denmark.
E-mail: t.faber@sdu.dk.

Introduction

The literature search forms the underlying basis of systematic reviews and thus the quality of the literature search is of crucial importance to the overall quality of the systematic review (1). The use of the four-part PICO model to facilitate searching for a precise answer is recommended (2) stating that a clinical question must be focused and well-articulated for all four parts: the patient or problem (P); the intervention or exposure (I); the comparison intervention or exposure (C), if relevant; and the clinical outcome of interest (O). Consequently, in addition to acting as a conceptualizing tool for asking clinical and research questions, the PICO model can be used as a tool for developing search strategies. According to Considine et al., “the PICO Framework should also be used to develop the search terms that are informed by the PICO question, Medical Subject Headings (MeSH) and any other terms deemed to be relevant” (3). It has been shown that using a PICO-structured form can elicit more detail, result in more complex and specific search strategies, and thus improve precision of retrieval (4). However, a recent systematic review finds only three studies assessing the effect of using the PICO model versus other available models or unguided searching on the quality of literature search results, and no solid conclusions about the effect of using the PICO model on the quality of the literature search can be drawn (5).

Existing guidelines generally recommend that a search strategy may include the population, intervention(s), and types of study design (6, 7). The rationale is that some elements may tend to be mentioned to a lesser degree in the abstract or subject headings (8, 9). The PICO model has been compared to the PIC model with respect to the median recall and precision (10). Adding study type to the search strategy is recommended by Cochrane and several studies analyses the recall of various study type search filters (11-13).

There is an impressive body of research on extracting PICO elements from abstracts or full text of clinical documents (14-26). The presence of PICO elements in clinical abstracts has been examined and PICO elements are identified in nearly all abstracts (16). A recent study of 20 abstracts identifies the four PICO elements in 90-100% in these 20 abstracts (27). A number of studies examine identification methods for PICO elements in clinical abstracts and report recall ranging from 37 to 84% for P and 26-80% for I (14, 17, 20, 21, 24). Using a named entity recognition model on 191 biomedical abstracts recall for P is 62%, I is 47% and outcomes 75% (28). However, the above-mentioned studies are focused on the presence of PICO elements in some form, not whether a thorough search would be able to retrieve the relevant studies.

There is evidence that outcome is not represented in enough detail but needs to be confirmed. The aim is to identify the PICO elements in title, abstracts and controlled vocabulary of included studies from Cochrane reviews with the aim of evaluating the retrieval potential of each PICO element.

Methods

This study is taking the perspective of the systematic review team. We sought to determine whether carefully crafted, exhaustive searches for the P, I, C, or O elements could retrieve the EMBASE or PubMed records of studies meeting the inclusion criteria of a set of previously published systematic

reviews. We examine the performance of each PICO element against the articles judged relevant for inclusion in systematic reviews. We analyze the relative recall which is the proportion of articles retrieved in a specific system in relation to the total or a pool of relevant documents. Relative recall has been suggested for validation of search filters (29). We use the included studies in several reviews as our starting point and we limit our analysis to the included studies that were indexed in PubMed and Embase. We determine the relative recall for each of the PICO elements (as defined by each of the reviews) by identifying the PICO elements in the title / abstract and subjects headings of the included studies. Hence, the question is whether is it possible to identify the PICO elements in the title / abstract or subject terms of the included studies? Consequently, there are two steps in the collection of data for this study:

1. Searching the Cochrane Library (Wiley) for eligible systematic reviews (or overview of reviews if such existed).
2. Searching Embase (Ovid) as well as PubMed for the included studies and extracting the records from both databases.

Identification of systematic reviews

The average annual number of included publications in Cochrane reviews exceed 10,000 (30) and a representative sample is therefore not possible and therefore reviews were selected until a minimum of 100 eligible included publications were identified. Eligible systematic reviews (or overview of reviews if such existed) were located within the two Cochrane Groups, Cochrane Pregnancy and Childbirth Group (CPC) and Cochrane Upper GI and Pancreatic Diseases Group (CUGIPD). These two groups were selected as the authors of the present study have the relevant expertise to assess the PICO elements within these two areas. Fig. 1 shows how Cochrane systematic reviews (data material) were identified. To analyze the most recent Cochrane overview of reviews or reviews, only reviews published within the last five years (from January 1, 2014 up until final search June 13, 2019) were eligible in this study.

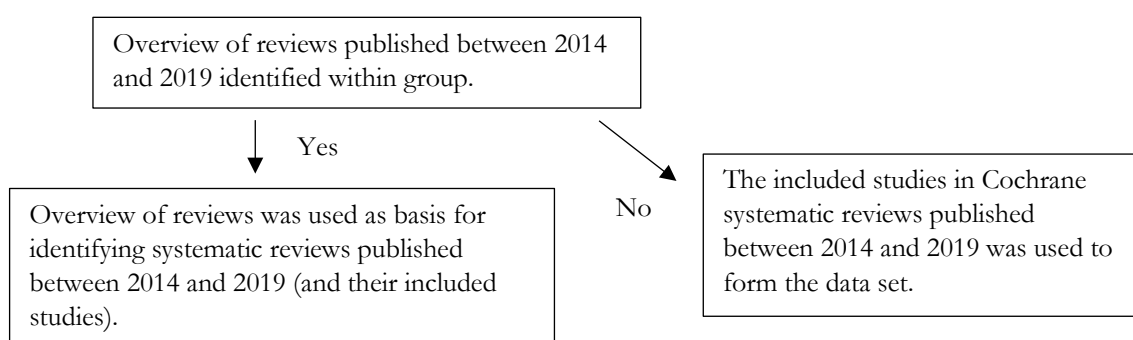


Fig. 1. Overview of the identification of systematic reviews.

Selection of Cochrane reviews from CPC group took place as follows: From the three identified overview of reviews (31-33) one was selected based on the clinical expert's knowledge within the area (32). This overview of reviews included ten Cochrane reviews (Table 2 in (32)), one which was published before 2014 and thus excluded from this study. The remaining nine Cochrane reviews published between 2014 – 2019 were numbered in alphabetical order by the first author's last name (Biesty # 1 and Raman # 9). The Random.org (using *true Random Number Generator*) website was used to select three reviews (34-36) by generating a random number between 1 - 9.

Selection of Cochrane reviews from CUGIPD Group took place as follows: No overviews of reviews were published in this Group, but a total of 65 reviews were published between January 1, 2014 up until final search June 13, 2019. From the 65 reviews, seven were randomly selected by numbering the reviews from 1 – 65 (with the oldest review numbered 1 and the most recently published review numbered 65) using Random.org (*true Random Number Generator*) (37-43). One review was excluded due to vaguely described secondary outcomes (38).

In this study we only use the references of the included studies from the Cochrane reviews that are indexed in both PubMed (Medline) and Embase to be able to compare the relative recall across two different databases. Furthermore, the data is limited to publications published after 1990 as the databases and the indexing have changed over time.

Identification of the gold standard

The included studies in the reviews were searched in the two databases using title words and author names as typical for known item searches (44). If a publication could not be retrieved in PubMed or Embase at least three different search term combinations were tried.

Our study aims to determine whether each of the included studies could be retrieved if a particular PICO element is used in the development of thorough search strategies (in either title/abstract/keywords or subject headings). Consequently, we examine titles/abstracts/keywords and subject headings of the included studies to identify the PICO elements. From each of the included studies in the selected reviews the following information was extracted (**Supplementary material – data extraction template**): date for extraction, publication characteristics (title, authors, journal and year), Cochrane review characteristics (title, PMID and Embase accession no.), reported findings summary table, data extraction and validation information, review characteristics (P, I, C, O elements as reported in the Cochrane review), PubMed bibliographic information (publication title, PMID, abstract and author keywords, Publication types, MeSH terms, Substances), Embase bibliographic information (publication title, accession no., abstract and author keywords, subject headings, drug index terms & other index terms).

Calculation of relative recall

Relative recall is determined on the basis of whether each of the included studies could be retrieved if a particular PICO element is used in the development of thorough search strategies. A thorough

search for relevant studies is in this study defined as using synonyms as well as narrower terms. Related terms are defined in this study by using the “entry terms” in PubMed, the “used for” terms in Embase as well as narrower terms in both thesauri. Consequently, the synonyms and related terms used in thorough search are determined using the thesauri. An example from the CUGIPD group can serve as an example. In one of the included reviews “postoperative mortality” is listed as a primary outcome. In this case “Surgical mortality” is accepted as a synonym for “postoperative mortality” as “postoperative mortality” is a “used for”-term for Embase SH “surgical mortality”.

If at least one of the search terms for a PICO element is identified in the bibliographic record of a specific publication, we noted it in the data as “1”. In the case of not being able to identify one of the terms it is marked as “0” in the data. An example of the calculation of relative recall is provided in appendix D.

Data was extracted and analyzed by one author which was validated by an author with research experience in each of the two disciplines the Cochrane groups represent. CLL validated data extraction from the CPC group, and MFBN validated data extraction from the CUGIPD Group.

Results

We analyzed a total of 104 publications included in Cochrane reviews published by two Cochrane groups. From CPC group 56 publications are used in the analysis (appendix A) and from CUGIPD group 48 publications are used in the analysis (appendix B). An overview of the reviews and their included studies is available in appendix C.

Only in different 3 publications out of 104 the identification of one PICO element was changed by an author with relevant research experience (CLL or MFBN). Consequently, agreement on the identification of PICO elements is high.

Table 1 provides an overview of the retrieval potential in percentages of the PICO elements in publications from the CPG Group.

PICO element	PubMed Text words (%)	Embase Text words (%)	PubMed MeSH-terms (%)	Embase Subject headings (%)
P	100	100	94.6	98.2
I	98.2	96.4	91.1	94.6
C	100	100	94.0	96.0
O (primary)	37.5	39.3	26.8	33.9
O (secondary)	69.6	71.4	64.3	82.1

Table 1. Recall in percentages of the PICO elements in publication from the CPC Group

In table 2 we see that relative recall is high for P, I and C in text words and only slightly lower using subject headings. We can see that the relative recall is somewhat lower for secondary outcome and considerably lower for primary outcome. The subject headings describe P, I and C rather well whereas O is present in the subject headings to a much lesser extent. Generally, PICO elements are identified more often in text words than in subject heading although the secondary outcomes are present more often in subject headings in Embase than in textwords.

Table 2 provides an overview of the retrieval potential in percentages of the PICO elements in publications from the CUGIPD Group.

PICO element	PubMed Text words (%)	Embase Text words (%)	PubMed MeSH-terms (%)	Embase Subject headings (%)
P	97.9	97.9	100	97.9
I	100	97.9	93.8	95.8
C	97.9	97.9	93.8	91.6
O (primary)	88.9	88.9	73.3	86.7
O (secondary)	67.6	67.6	43.2	48.6

Table 2. Recall in percentages of the PICO elements in publications from the CUGIPD.

In table 2 we can clearly see that the relative recall is higher for P, I and C, whereas relative recall is somewhat lower for primary outcome and considerably lower for secondary outcome. We also notice that the retrieval potential of P is remarkably similar for searches in text words and subject headings. For I, C and O, the retrieval potential differs between text words and subject headings.

Tables 1 and 2 clearly shows us that the retrieval potential of PICO elements is varying considerably in text words as well as subject headings. We then combine searches in text words as well as subject headings and table 3 provides an overview of the retrieval potential of each element.

PICO element	Cochrane Pregnancy and Childbirth Group		Cochrane Upper GI and Pancreatic Diseases Group	
	PubMed Text words- MeSH (%)	Embase Text words- Subject headings (%)	PubMed Text words- MeSH (%)	Embase Text words- Subject headings (%)
P	100	100	100	100
I	98.2	96.4	100	100
C	100	100	97.9	97.9

O (primary as well as secondary)	80.4	89.3	93.3	93.3
----------------------------------	------	------	------	------

Table 3. Recall in percentages of the PICO elements in publications from the CPC Group and the CUGIPD Group.

Table 3 shows the retrieval potential of a combined search of both text words and subject headings. A combination of text words and subject heading P results in a recall of 100% across groups and databases. For the CUGIPD Group, I also has a recall of 100% in both Embase and PubMed, but for CPC Group this is slightly lower (96.4 – 98.2%). C has a relatively high retrieval potential in both groups (97.9-100%). The retrieval potential of O is the lowest in both groups compared to the other PICO elements (80.4-93.3%).

Discussion and conclusion

This study aims to substantiate the assumption that when using PICO as a search strategy tool, searching on all of the PICO elements will result in a lower recall. The current recommendations point out that appropriate elements from PICO (P, I and type of study design) should be included in the search strategy (7). Our results support this statement (table 3), since the PICO elements C and O had a lower recall across the two Cochrane groups and databases also when combining text words and subject headings, as recommended for higher recall (7). It is relevant to note that for both CUGIPD and CPC Group recall of all the PICO elements were higher among Embase subject headings than PubMed MeSH.

Furthermore, our findings show that searches text words and searches using subject headings results in different results although they are of course overlapping. Current recommendations suggest using both (7, 45) which is supported by our results.

There are limitations to this study that affect the generalizability of the results. First and foremost, retrieval of PICO elements is most likely overestimated in this study as all potential synonyms and narrower terms are included when defining a thorough search strategy. In reality, the information specialist would have to balance recall and precision (29) and that would most likely lower the recall.

Despite the convincing similarity of the high recall of P and I in both the investigated groups this is a case study, which only investigates reviews and included studies / publications from these two Cochrane groups. Consequently, we cannot generalize to all topics across all groups. Furthermore, a change in the type of studies assessed might also have an effect on the results. In our study, only RCTs are included and we cannot determine if other study types would have an effect on the results. Finally, our investigations focused on quantitative studies; qualitative research may be indexed differently (46, 47).

Furthermore, we found a lower recall of especially the outcome element, however, we need to keep in mind that we investigate recall of the PICO elements from the perspective of the Cochrane review team and thus, the PICO elements in the original study may be different. The outcomes may be prioritized differently by authors of the original study (publication) than by the Cochrane

reviewer. Therefore, outcomes may not necessarily be identified through a combined search of text words and subject headings.

The method used to assess retrievability is another limitation of the study. Retrievability assessments were made comparing search terms to record contents, thus human errors were possible. However, our use of independent assessment by the two, literature search experts (TFF and MBE) followed by verification of the judgement by an author with relevant research experience (CLL or MFBN), decreased the chance of error.

We studied possible retrieval of P, I, C, and O elements using searches containing relevant headings and text words derived from specified sources. Recall of the P, I, C, and O elements during an actual systematic review will, of course, depend on the skill of the searcher and the sources or methods s/he uses to gather search terms. This study does not repeat existing searches and we can determine if the publications are retrievable but not if they were retrieved by a specific review team.

Summing up, the results seem to support the recommendation not to search for outcome, as the retrieval potential of outcome is lower than for other PICO elements. The consequences of searching for outcome may depend on the specific search topic. Obviously, more studies are needed to draw final conclusions and ideally these should address different areas as well as review types. The present study has outlined a methodology suitable for the analyses of retrieval potential and the findings indicate that outcome is the element with the smallest retrieval potential of the four PICO elements. Preferably, the outcome element should be left out when developing the search strategy for a systematic review.

Data associated with this article are available in the Open Science Framework at www.sdu.dk

Literature

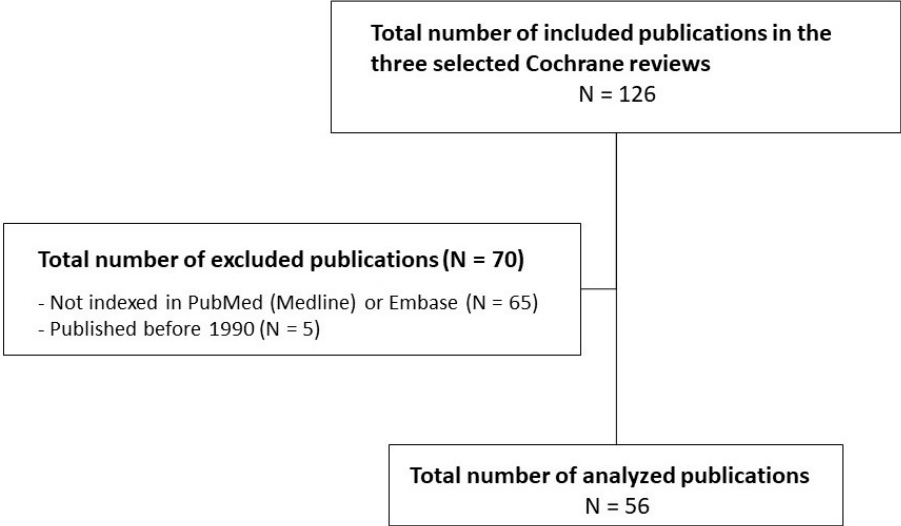
1. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol.* 2016;75:40-6.
2. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J Club.* 1995;123(3):A12-3.
3. Considine J, Shaban RZ, Fry M, Curtis K. Evidence based emergency nursing: Designing a research question and searching the literature. *Int Emerg Nurs.* 2017;32:78-82.
4. Booth A, O'Rourke AJ, Ford NJ. Structuring the pre-search reference interview: a useful technique for handling clinical questions. *Bull Med Libr Assoc.* 2000;88(3):239-46.
5. Eriksen MB, Frandsen TF. The impact of patient, intervention, comparison, outcome (Pico) as a search strategy tool on literature search quality: A systematic review. *Journal of the Medical Library Association.* 2018;106(4):420-31.

6. 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.
7. 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 60 (updated July 2019): Cochrane; 2019.
8. 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.
9. 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. *Systematic reviews*. 2019;8(1):174.
10. 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. *Journal of medical Internet research*. 2012;14(3):e85.
11. McKibbin KA, Wilczynski NL, Haynes RB. Retrieving randomized controlled trials from medline: a comparison of 38 published search filters. *Health information and libraries journal*. 2009;26(3):187-202.
12. Fraser C, Murray A, Burr J. Identifying observational studies of surgical interventions in MEDLINE and EMBASE. *BMC Med Res Methodol*. 2006;6.
13. Beynon R, Leeftang MMG, McDonald S, Eisinga A, Mitchell RL, Whiting P, et al. Search strategies to identify diagnostic accuracy studies in MEDLINE and EMBASE. *Cochrane Database of Systematic Reviews*. 2013(9).
14. Boudin F, Nie JY, Bartlett JC, Grad R, Pluye P, Dawes M. Combining classifiers for robust PICO element detection. *BMC Med Inform Decis Mak*. 2010;10:29.
15. Egan M, MacLean A, Sweeting H, Hunt K. Comparing the effectiveness of using generic and specific search terms in electronic databases to identify health outcomes for a systematic review: A prospective comparative study of literature search methods. *BMJ open*. 2012;2(3).
16. Dawes M, Pluye P, Shea L, Grad R, Greenberg A, Nie JY. The identification of clinically important elements within medical journal abstracts: Patient-Population-Problem, Exposure-Intervention, Comparison, Outcome, Duration and Results (PECODR). *Inform Prim Care*. 2007;15(1):9-16.
17. Chabou S, Iglewski M. Combination of conditional random field with a rule based method in the extraction of PICO elements. *BMC Med Inform Decis Mak*. 2018;18(1):128.
18. Kim SN, Martinez D, Cavedon L, Yencken L. Automatic classification of sentences to support Evidence Based Medicine. *BMC Bioinformatics*. 2011;12 Suppl 2:S5.
19. Wallace BC, Kuiper J, Sharma A, Zhu MB, Marshall IJ. Extracting PICO Sentences from Clinical Trial Reports using Supervised Distant Supervision. *Journal of machine learning research : JMLR*. 2016;17.

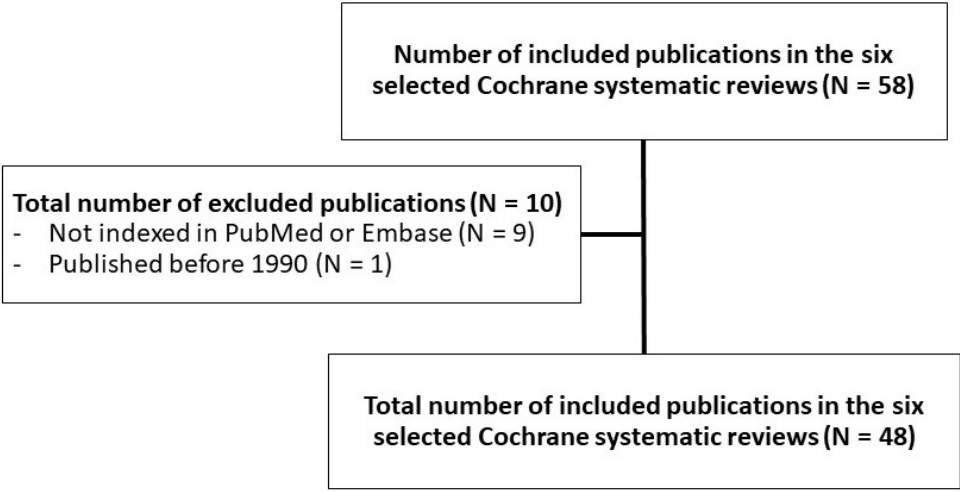
20. Chung GY. Sentence retrieval for abstracts of randomized controlled trials. *BMC Med Inform Decis Mak.* 2009;9:10.
21. Hansen MJ, Rasmussen NO, Chung G. A method of extracting the number of trial participants from abstracts describing randomized controlled trials. *J Telemed Telecare.* 2008;14(7):354-8.
22. Hassanzadeh H, Groza T, Hunter J. Identifying scientific artefacts in biomedical literature: the Evidence Based Medicine use case. *J Biomed Inform.* 2014;49:159-70.
23. Groza T, Hassanzadeh H, Hunter J. Recognizing scientific artifacts in biomedical literature. *Biomedical informatics insights.* 2013;6:15-27.
24. Chung GY. Towards identifying intervention arms in randomized controlled trials: extracting coordinating constructions. *J Biomed Inform.* 2009;42(5):790-800.
25. Yuan X, Xiaoli L, Shilei L, Qinwen S, Ke L, editors. Extracting PICO elements from RCT abstracts using 1-2gram analysis and multitask classification. 2019: Association for Computing Machinery.
26. Zlabinger M, Andersson L, Brassey J, Hanbury A. Extracting the population, intervention, comparison and sentiment from randomized controlled trials. In: Klein GO, Karlsson D, Moen A, Ugon A, editors.: IOS Press; 2018. p. 146-50.
27. Eldawlatly A, Alshehri H, Alqahtani A, Ahmad A, Al-Dammas F, Marzouk A. Appearance of Population, Intervention, Comparison, and Outcome as research question in the title of articles of three different anesthesia journals: A pilot study. *Saudi J Anaesth.* 2018;12(2):283-6.
28. Brockmeier AJ, Ju M, Przybyla P, Ananiadou S. Improving reference prioritisation with PICO recognition. *BMC Med Inform Decis Mak.* 2019;19(1):256.
29. Sampson M, Zhang L, Morrison A, Barrowman NJ, Clifford TJ, Platt RW, et al. An alternative to the hand searching gold standard: Validating methodological search filters using relative recall. *BMC Med Res Methodol.* 2006;6.
30. Frandsen TF, Eriksen MB, Hammer DMG, Christensen JB. Fragmented publishing: a large-scale study of health science. *Scientometrics.* 2019;119(3):1729-43.
31. Shepherd E, Salam RA, Middleton P, Makrides M, McIntyre S, Badawi N, et al. Antenatal and intrapartum interventions for preventing cerebral palsy: an overview of Cochrane systematic reviews. *Cochrane database of systematic reviews (Online).* 2017;8:CD012077.
32. Martis R, Crowther CA, Shepherd E, Alsweiler J, Downie MR, Brown J. Treatments for women with gestational diabetes mellitus: an overview of Cochrane systematic reviews. *Cochrane database of systematic reviews (Online).* 2018;8:CD012327.
33. Medley N, Vogel JP, Care A, Alfirevic Z. Interventions during pregnancy to prevent preterm birth: an overview of Cochrane systematic reviews. *Cochrane database of systematic reviews (Online).* 2018;11:CD012505.
34. Brown J, Martis R, Hughes B, Rowan J, Crowther CA. Oral anti-diabetic pharmacological therapies for the treatment of women with gestational diabetes. *Cochrane database of systematic reviews (Online).* 2017;1:CD011967.

35. Brown J, Grzeskowiak L, Williamson K, Downie MR, Crowther CA. Insulin for the treatment of women with gestational diabetes. *Cochrane database of systematic reviews (Online)*. 2017;11(11):CD012037.
36. Biesty LM, Egan AM, Dunne F, Smith V, Meskell P, Dempsey E, et al. Planned birth at or near term for improving health outcomes for pregnant women with pre-existing diabetes and their infants. *Cochrane Database Syst Rev (online)*. 2018;2:CD012948.
37. Cheng Y, Briarava M, Lai M, Wang X, Tu B, Cheng N, et al. Pancreaticojejunostomy versus pancreaticogastrostomy reconstruction for the prevention of postoperative pancreatic fistula following pancreaticoduodenectomy. *Cochrane database of systematic reviews (Online)*. 2017;9:CD012257.
38. Gurusamy KS, Belgaumkar AP, Haswell A, Pereira SP, Davidson BR. Interventions for necrotising pancreatitis. *Cochrane database of systematic reviews (Online)*. 2016;4:CD011383.
39. Gurusamy KS, Kumar S, Davidson BR, Fusai G. Resection versus other treatments for locally advanced pancreatic cancer. *Cochrane database of systematic reviews (Online)*. 2014(2):CD010244.
40. Huttner FJ, Fitzmaurice C, Schwarzer G, Seiler CM, Antes G, Buchler MW, et al. Pylorus-preserving pancreaticoduodenectomy (pp Whipple) versus pancreaticoduodenectomy (classic Whipple) for surgical treatment of periampullary and pancreatic carcinoma. *Cochrane database of systematic reviews (Online)*. 2016;2:CD006053.
41. Kidane B, Coughlin S, Vogt K, Malthaner R. Preoperative chemotherapy for resectable thoracic esophageal cancer. *Cochrane database of systematic reviews (Online)*. 2015(5):CD001556.
42. Song H, Zhu J, Lu D. Long-term proton pump inhibitor (PPI) use and the development of gastric pre-malignant lesions. *Cochrane Database of Systematic Reviews*. 2014(12):CD010623.
43. Yuan Y, Zeng X, Hu Y, Xie T, Zhao Y. Omentoplasty for oesophagogastronomy after oesophagectomy. *Cochrane database of systematic reviews (Online)*. 2014(10):CD008446.
44. Lee JH, Renear A, Smith LC. Known-Item Search: Variations on a Concept. *Proceedings of the American Society for Information Science and Technology*. 2006;43(1):1-17.
45. Booth A, Papaioannou D, Sutton A. *Systematic approaches to a successful literature review*: Sage; 2011.
46. Booth A. Searching for qualitative research for inclusion in systematic reviews: a structured methodological review. *Systematic reviews*. 2016;5:74.
47. Wagner M, Rosumeck S, Kuffmeier C, Doring K, Euler U. A validation study revealed differences in design and performance of MEDLINE search filters for qualitative research. *J Clin Epidemiol*. 2020;120(April):17-24.

Appendix A. Total number of publications analyzed from Cochrane Pregnancy and Childbirth Group.



Appendix B. Total number of publications analyzed from CUGIPD Group.



Appendix C. Characteristics of included Cochrane reviews.

Cochrane group	Cochrane review title	Topic	Included study type	No. of included studies
Upper GI and Pancreatic Group	Pancreaticojejunostomy versus pancreaticogastrostomy reconstruction for the prevention of postoperative pancreatic fistula following pancreaticoduodenectomy.	Prevention of postoperative pancreatic fistula	Randomized controlled trials	10 studies (10 publications / 10 included in this study)
	Resection versus other treatments for locally advanced pancreatic.	Pancreatic cancer	Randomized controlled trials	2 studies (5 publications / 4 included in this study)
	Pylorus-preserving pancreaticoduodenectomy (pp Whipple) versus pancreaticoduodenectomy (classic Whipple) for surgical treatment of periampullary and pancreatic carcinoma.	Pancreatic carcinoma	Randomized controlled trials	8 studies (9 publications / 7 included in this study)
	Preoperative chemotherapy for resectable thoracic esophageal cancer.	Esophagus cancer	Randomized controlled trials	13 studies (21 publications / 14 included in this study)
	Long-term proton pump inhibitor (PPI) use and the development of gastric pre-malignant lesions.	Gastric premalignant lesions	Randomized controlled trials	8 studies (10 publications / 10 included in this study)
	Omentoplasty for oesophagogastronomy after oesophagectomy.		Randomized controlled trials	3 studies (3 publications / 3 included in this study)
Cochrane Pregnancy and Childbirth Group (CPC)	Planned birth at or near term for improving health outcomes for pregnant women with gestational diabetes and their infants.	Gestational diabetes	Randomized controlled trial	1 study (3 publications / 2 included in this study)
	Oral anti-diabetic pharmacological therapies for the treatment of women with gestational diabetes.	Gestational diabetes	Randomized controlled trials	11 studies (20 publications / 6 included in this study)
	Insulin for the treatment of women with gestational diabetes.	Gestational diabetes	Randomized controlled trials	52 studies (103 publications / 48 included in this study)

Appendix D. An example of the calculation of retrieval potential for a review

The review titled “Planned birth at or near term for improving health outcomes for pregnant women with gestational diabetes and their infants” includes one study published in three publications. Out of these three publications two fulfil our eligibility criteria. Table 1 provides an example of the collected data to calculate retrieval potential of these two publications.

Table 1. Example of collected data

Publication	PICO element	PubMed Text words	Embase Text words	PubMed MeSH-terms	Embase Subject headings
#1	P	1	1	1	0
#1	I	1	1	0	0
#1	C	1	1	0	0
#1	O (primary)	1	1	1	0
#1	O (secondary)	1	1	1	0
#2	P	1	1	1	1
#2	I	1	1	1	1
#2	C	1	1	1	0
#2	O (primary)	0	0	0	0
#2	O (secondary)	0	0	0	0

These two publications could both be retrieved in Embase and PubMed when searching for P (population), I (intervention) and C (comparison) using title, abstract and keywords. If only using subject headings for the search, retrieval is lower as one of the publications would not be retrieved using search terms for the following PICO elements: P (only Embase), I (both PubMed and Embase) and C (both Embase and PubMed). Retrieval potential for outcome (O) is much lower. Searching for either primary or secondary outcomes, one of the included studies could not be retrieved. The other publication would be retrieved if searchers use title, abstract and keywords. If only using subject headings it would be retrieved in PubMed but not in Embase. Looking at these two publications exclusively the retrieval potential of P, I as well as C is consequently 100% in both databases if the searches are made using title, abstract, keywords as well as subject headings. Looking at these two publications exclusively the retrieval potential of C (primary and secondary combined) is 50% in both databases if the searches are made using title, abstract, keywords as well as subject headings. Searches in either title, abstract, keywords or subject headings have a lower retrieval potential.