

Protocol for the validation of four search strategies for retrieval of Clinical Practice Guidelines in MEDLINE, Embase and PubMed

Carole Lunny^{1*}, Douglas M Salzwedel¹, Tracy Liu¹, Cynthia Ramasubbu¹, Savannah Gerrish¹, Lorri Puil¹, Barbara Mintzes², James M. Wright¹

Affiliations:

1. Cochrane Hypertension Review Group, Therapeutics Initiative, Department of Anesthesiology, Pharmacology and Therapeutics, Faculty of Medicine, University of British Columbia, 2176 Health Science Mall, Vancouver, BC, Canada, V6T 1Z3
2. The University of Sydney, Charles Perkins Centre, and School of Pharmacy, 6W75, The Hub, Charles Perkins Centre D17, The University of Sydney, NSW, 2006

* Corresponding author: Carole Lunny, MPH, PhD(c), Cochrane Hypertension Review Group, Therapeutics Initiative, Department of Anesthesiology, Pharmacology and Therapeutics, Faculty of Medicine, University of British Columbia, 2176 Health Science Mall, Vancouver, BC, Canada, V6T 1Z3, carole.lunny@ti.ubc.ca

Abstract

Background

Guidelines are systematically developed recommendations to assist practitioner and patient decisions about treatments for clinical conditions. Researchers, healthcare professionals and policy makers need to be able to retrieve clinical practice guidelines (CPGs) efficiently and quickly from the literature. Despite the widespread use of CPGs in practice and policy formulation, no filter for retrieval of guidelines has been validated to date. The use of a validated search filter for CPGs would make their retrieval from major bibliographic databases more efficient.

Objectives

We aim to fill this gap by validating search filters for use in the systematic retrieval of CPGs and measure their performance according to sensitivity and precision.

Methods

We found four search filters for retrieval of CPGs (two CADTH, PubMed and University of Texas filters) which we will validate in three databases (MEDLINE, Embase and PubMed). We will derive a test set of CPGs from a search of the TRIP and Epistemonikos databases. The citations retrieved will be randomly sorted and screened sequentially by two reviewers until at least 100 CPGs are included. We will include CPGs that provide at least two explicit recommendations for the treatment of any clinical condition, and that are produced by a group or organization (i.e., not authored by one person). We will translate the filters into Ovid MEDLINE, Embase, and PubMed syntax as appropriate. Then, we will run the strategies and assess whether the filters retrieved the citations in our test set. We will calculate and compare the sensitivity and precision of the four filters in each database. The limitations of the CADTH, PubMed and University of Texas search filters for each database will be assessed by examining the keywords in the titles and abstracts of the citations not found by the search filters.

Discussion

Decision makers, healthcare providers and researcher will be able to choose the most precise and sensitive search strategy among the four available, which will enable them to more efficiently identify relevant clinical practice guidelines.

Background

Guidelines are defined as systematically developed recommendations to assist in practitioner and patient decision making about treatments for clinical conditions [1, 2]. Researchers, healthcare professionals and policy makers need to be able to retrieve clinical practice guidelines (CPGs) efficiently and quickly from the literature. While there are many websites dedicated to publishing CPGs such as TRIP, CPG Infobase and Epistemonikos, comprehensive search strategies include an additional search of bibliographic databases such as MEDLINE and Embase. The use of a validated search filter for CPGs would make their retrieval from major bibliographic databases more efficient. Despite the widespread use of CPGs in practice and policy formulation, no filter for retrieval of guidelines has been validated to date.

Search filters, or hedges, are a useful way of retrieving specific types of research, such as systematic reviews, randomised trials or articles on particular health conditions [3]. They aim to retrieve as many relevant citations as possible while minimising the number of irrelevant ones. Filters usually comprise a combination of free text and controlled vocabulary terms (e.g. Medical Subject Headings (MeSH)) and, depending on the purpose of the search filter, may be narrow or broad in scope.

The validation of a filter on a set of citations provides an independent indication of how well the search filter performs. The early search filters by Haynes for PubMed's Clinical Queries were validated against a single gold standard or test set of citations that had been split into a development set (a set of citations used to create the terms comprising the search filter) and a validation set (a set of citations used to test the measurement properties of the developed search strategy) [4].

Objective

Our aim is to validate search filters for use in the systematic retrieval of CPGs and measure their performance according to sensitivity and precision.

Methods

We have registered this protocol in the PROSPERO database (CRD42018105865), and in the Open Science Framework (osf.io/rju4f). The design is a methods study in the knowledge synthesis field, and the study follows systematic review methods guidance.

Search for Clinical Practice Guidelines filters

A Google search for filters to retrieve CPGs was conducted using the following terms: guidelines AND search filters, or search hedges, or databases. A search of the InterTASC Information Specialists' Sub-Group search filter registry [5] was also conducted for clinical practice guideline filters. Three search filters were identified online – CADTH [6], University of Texas [7], PubMed [8], and a published filter by Haase [9]. The CADTH filter was designed for use in Ovid MEDLINE, Embase, and PsycINFO [6]. In personal communication with the authors of the CADTH filter, two different strategies which were updated in 2016 were shared – one broad and one narrow strategy to retrieve CPGs. The University of Texas filter was designed for use in Ovid MEDLINE [7], and the Haase 2007 filter [9] was designed for use in SUMSEARCH and Google Scholar, and the PubMed filter was designed for use in the PubMed database [8].

We excluded the Haase filter as it is over 10 years old, and the two databases it was designed for (SumSearch and Google Scholar) are either obsolete, or not reproducible. the SumSearch database has been replaced by the SumSearch 2 database and is now housed at a different institution. In addition, we considered Google Scholar an inappropriate database for use in systematic searching as it has been

criticized for being up to one year out-of-date [10], not being transparent about the journals and grey literature that form the content of the database [11], offering limited search functionality (filtering, qualifiers, and MeSH terms) [12-14], and a database which is customised to an individuals searching preferences, making search strategies irreproducible by other users.

Therefore, the search filters that will be tested include two from CADTH (broad and narrow; **Supplementary file 1**), one from the University of Texas (**Supplementary file 2**), and one from PubMed **Supplementary file 3**). To our knowledge, these filters have not been tested for their sensitivity or precision against a test set of CPGs.

Given that MEDLINE has been the gold standard for structured searching for decades [14], and that Medline and Embase are recommended for systematic searches for medical studies, these databases were chosen as testing sites using the Ovid interface. The PubMed database was chosen with the rationale that it is free and accessible to all, whereas the other databases are paywalled. As a post-hoc decision, we excluded PsycINFO as a database to validate as the focus of the test set of CPGs was not on behavioral science and mental health.

Using the Open Science Framework

We will be using the Open Science Framework (OSF) as a tool to promote the principles of openness, transparency, and reproducibility in our research. OSF is an online platform to customise project workflows. A standard project layout includes a wiki (i.e. information about the project), a log of recent activity, and spaces to keep notes, upload files, collaborate, and share information [15]. In OSF we will document our project protocol in a section called wiki, which includes our purpose, objectives, methods, and procedures. We will use the OSF file storage area to upload files to keep track of progress on study screening and data extraction. If a co-author adds or renames a file on OSF it will be flagged in the recent activities tab. Finally, we aim to make our project open after study completion, as well as our data sets.

Developing the test set of citations

To ensure there is not an over-sampling of published guidelines from MEDLINE, PubMed, or Embase, we will retrieve a test set of citations from TRIP and Epistemonikos databases. These databases collect their content from various sources including bibliographic databases, journal publishers and the grey literature.

Epistemonikos includes citations retrieved from the following databases: Cochrane Database of Systematic Reviews; PubMed; Embase; CINAHL (The Cumulative Index to Nursing and Allied Health Literature); PsycINFO; LILACS (Literatura Latinoamericana y del Caribe en Ciencias de la Salud); DARE (Database of Abstracts of Reviews of Effects); the Campbell Collaboration's online library; JBI Database of Systematic Reviews and Implementation Reports; and the EPPI-Centre Evidence Library. The TRIP database retrieves guidelines from over 289 journal publications and has recently migrated all content from AHRQ's Clinical Guidelines Clearinghouse (www.guidelines.gov), which was shut down July 16, 2018 (John Brassey, personal communication, April 10, 2018).

We have chosen very recent but narrow search dates (between January 1, 2017 and March 30, 2018), to identify CPGs which will limit the number of CPGs being screened.

As the TRIP database only contains CPGs, we will limit our search by the dates listed above and download all records. For retrieval of CPGs in the Epistemonikos, we will select the filter for guidelines as well as the date limitations.

Random selection

After deduplication in EndNote, the retrieved citations will be randomly sorted using Excel's RAND function. Two authors will independently screen the randomly selected titles, abstracts and full text against predefined eligibility criteria. They will start with the lowest random number and work up until at least 100 eligible guidelines are identified. Ten citations will first be screened independently by two authors in Excel as a calibration exercise to establish agreement on definitions of eligibility criteria. Discrepant decisions will be resolved by discussion or by arbitration with a third author.

Eligibility criteria

Guidelines are defined as systematically developed statements to assist in practitioner and patient decision making about treatments for any clinical condition [1, 2]. We will include CPGs that provide recommendations for treatment of any clinical condition, and that are produced by a group or organization.

CPGs must contain at least two explicit recommendations. We will include CPGs published between January 1, 2017 and March 30, 2018, with an explicit description of their methodology (for example, definition of search strategy, methods used to create recommendations, and quality assessment)—either within the CPG or in supporting documents.

Conversely, documents lacking recommendations or only focused on screening or diagnosis will be excluded. We will also exclude CPGs without a comprehensive reference list. CPGs will also be excluded for the following reasons: absence of treatment recommendations; designed for local use (for example, in a single health facility or single regional health service); designed for use with only hospitalized patients or patients in long-term care facilities; or if the full text is unavailable or only a summary of the recommendations is available [16]. Documents that only address the use of medications, such as guidance about adherence to medications, will be also excluded.

Sample size

The size and content of the test set of citations may have an impact on the performance measures recorded for a specific filter. Sampson 2011 recommend that at least 100 citations should be included so that the sensitivity can be accurately estimated [17]. For example, assuming an underlying sensitivity of 90%, a sample of 100 citations would yield a 95% confidence interval of width $\pm 6\%$.

Translation of the search strategies into each database

We will translate the CADTH, PubMed and University of Texas filters into Ovid MEDLINE, Ovid Embase, and PubMed syntax (e.g. MeSH terms, adjacency operators) as needed. We will first confirm that all the lines of the search algorithms do not produce any syntax errors when entered into each of the databases line by line. If any errors do occur, one reviewer will translate the strategies, and a second expert librarian (DS) will correct and edit the translations line by line into the database syntax. We will copy the Ovid MEDLINE translated strategies into the software MEDLINE Transpose to translate to PubMed format (<https://medlinetranspose.github.io/#results>) [18]. MEDLINE Transpose is a freely available, open source tool for researchers that automatically converts PubMed and Ovid interfaces.

Testing the strategies

Citations without PubMed IDs (PMIDs) or DOIs, or that are not indexed in the respective database will be excluded.

We will run the strategies and assess whether the filters retrieved the test set of citations. We will compare the sensitivity and precision of the four filters to each other. The limitations of the CADTH, PubMed and University of Texas search filters for the three databases will be assessed by examining the titles and abstracts of the citations not retrieved for words that could have identified them as guidelines.

Calculation of sensitivity and precision

We will calculate the sensitivity and precision of the test set of citations in the four search strategies in each of the three databases, as well as the confidence intervals (**Table 1**). Confidence intervals will be calculated using Stata version 13. The Stata command for calculating confidence intervals of proportions is `cii` with option `binomial`. Data will be displayed in tables.

Table 1: Formula for calculating sensitivity and precision

	Eligible articles	Ineligible articles	Total articles
Retrieved by search filter	a	b	a+b
Not retrieved by search filter	c	d	c+d
Total	a+c	b+d	N

$$\text{Sensitivity: } \frac{\text{Number of eligible articles retrieved by the search filter}}{\text{Total number of eligible articles in the gold standard set}} = \frac{a}{a + c}$$

$$\text{Precision: } \frac{\text{Number of eligible articles retrieved by the search filter}}{\text{Total number of articles retrieved}} = \frac{a}{a + b}$$

Risk of bias assessment of the search strategies

Only the search algorithms are reported without any accompanying manuscript; therefore, risk of bias cannot be assessed.

Discussion

The main objective of this study is to assess and compare the sensitivity and precision of four search strategies in three databases. Despite the widespread use of clinical practice guidelines in practice and policy formulation, to our knowledge, no filter for retrieval of guidelines has been validated. Several organisations have published search strategies that have not been validated against test datasets, and so their ability to capture all published guidelines is unknown. Researchers, practitioners and policy makers need to be able to access guidelines more efficiently than is currently the case.

Recommended methods for the validation of search filters have been published [19, 20]. Ideally, the search filter should be tested based on its ability to retrieve all guidelines in existence. As this is not a feasible approach, testing the strategy on a set of citations which is representative of the population of all available guidelines is common practice. Studies have used multiple methods to develop test sets for various filters: studies retrieved by handsearch of journals, studies retrieved by using reference lists of

systematic reviews, or using a database search to identify records to include in their test set [20]. We chose to use this last method.

Different metrics are used to assess the measurement properties of filters including sensitivity, specificity, accuracy, precision, and number needed to read (NNR), with accompanying confidence intervals [21]. Ideally, a sensitivity-maximising strategy should be considered, where resources are available, to screen a larger number of citations and there is need for comprehensiveness, or when the strategy will be appended to search terms for specific conditions or interventions. In contrast, a precision-maximising strategy is likely to result in a smaller number of citations to screen, with a large number of relevant citations. An optimal search filter will have a high sensitivity/recall (how much of the relevant literature is located) and high precision (how much of the retrieved literature is relevant) with a low NNR. Sensitivity of over 90% is considered optimal in systematic review searches with an accompanying precision of over 10% (typical ranges for precision in searches of systematic reviews are between 3 and 6% [17]).

Healthcare decision-makers and professionals need efficient access to the best evidence [22]. Search filters for retrieval of guidelines expedite this process and are used by librarians, researchers, health technology agencies and policy makers. Data are lacking on how well published search filters perform in finding CPGs. Comparison and validation of search filters to identify CPGs in bibliographic databases are needed to identify the most sensitive and precise strategies for use by healthcare professionals [22].

We aim to present the results of this research at the annual Cochrane Colloquium and the Guidelines International Network (GIN) conference. The Cochrane Colloquium is an international gathering to promote methods in the production of high-quality, relevant, accessible systematic reviews and other synthesized research [23]. The GIN conference is an international symposium for those who work with guidelines; from development and methodology through to implementation and evaluation [24]. The results will also be disseminated through social media (twitter, Facebook, Researchgate), author affiliated websites, university workshops, and sent to the InterTASC Information Specialists' Sub-Group for inclusion on their search filter registry [5].

The strength of our methods include the adoption of systematic and transparent methods, specific and explicit eligibility criteria, broad search strategies using multiple sources, randomised screening, and duplicate and independent processes for study selection. The main limitations of our study is the narrow search dates of the test set of CPGs and not validating the strategies in the PsycINFO database. We have also not performed a systematic search for search strategies for CPGs, and therefore may have missed some that are used in-house or that are not available online. If our readers are aware of any other search filters for retrieval of CPGs, please email the principal investigator.

Our study will enable decision makers, healthcare providers and researcher to choose the most precise and sensitive search strategy among the four available, which will allow for more efficiently retrieval of relevant clinical practice guidelines.

References

1. Burgers JS, Fervers B, Haugh M, Brouwers M, Browman G, Philip T, Cluzeau FA: **International assessment of the quality of clinical practice guidelines in oncology using the Appraisal of Guidelines and Research and Evaluation Instrument.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2004, **22**(10):2000-2007.
2. Field MJ, Lohr, K.N.: **Clinical Practice Guidelines: Directions for a New Program.** Institute of Medicine. Washington, DC: National Academy Press; 1990.
3. 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.
4. Haynes RB, McKibbin K, Wilczynski NL, Walter SD, Werre SR: **Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey.** *BMJ (Clinical research ed)* 2005, **330**(7501):1179.
5. InterTASC Information Specialists' Sub-Group: **ISSG Search Filter Resource.** Available from: <https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home>. In: Edited by Glanville J, Lefebvre, C., Wright, K. York, UK; 2008 [updated 2018 February 26].
6. **Strings Attached: CADTH's Database Search Filters: Guidelines — OVID Medline, Embase, PsycINFO.** Available at: <https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters>
7. University of Texas: **Search Filters for Various Databases: Ovid Medline: Guidelines/Recommendations (revised 12/3/2015).** Available at: http://libguides.sph.uth.tmc.edu/search_filters/ovid_medline_filters. 2015.
8. Geng, Y. **PubMed Guidelines filter. 2017.** Available at: https://mdanderson.libguides.com/ld.php?content_id=34393248
9. Haase A, Follmann M, Skipka G, Kirchner H: **Developing search strategies for clinical practice guidelines in SUMSearch and Google Scholar and assessing their retrieval performance.** *BMC medical research methodology* 2007, **7**:28.
10. Vine R: **Google Scholar.** *Journal of the Medical Library Association* 2006, **94**(1):97-99.
11. Gray JE, Hamilton MC, Hauser A, Janz MM, Peters JP, Taggart F: **Scholarish: Google Scholar and its value to the sciences.** *Issues in Science and Technology Librarianship* 2012, **70**(Summer).
12. Boeker M, Vach W, Motschall E: **Google Scholar as replacement for systematic literature searches: good relative recall and precision are not enough.** *BMC medical research methodology* 2013, **13**:131.
13. Giustini D: **How Google is changing medicine.** *BMJ (Clinical research ed)* 2005, **331**(7531):1487-1488.
14. Giustini D, Boulos MN: **Google Scholar is not enough to be used alone for systematic reviews.** *Online journal of public health informatics* 2013, **5**(2):214.
15. Foster ED, Deardorff A: **Open Science Framework (OSF).** *Journal of the Medical Library Association : JMLA* 2017, **105**(2):203-206.
16. Molino CdGRC, Ribeiro E, Romano-Lieber NS, Stein AT, Melo DOd: **Methodological quality and transparency of clinical practice guidelines for the pharmacological treatment of non-communicable diseases using the AGREE II instrument: a systematic review protocol.** *Systematic reviews* 2017, **6**:220.
17. Sampson M, Tetzlaff J, Urquhart C: **Precision of healthcare systematic review searches in a cross-sectional sample.** *Res Synth Methods* 2011, **2**(2):119-125.
18. Wanner A, Baumann, N.: **Medline Transpose: Conversion Tool for Pubmed and Ovid Medline Interfaces.** In: *Virtual Projects: Facing reality: the growth of virtual reality and health sciences*

- libraries Available at: <http://jmlamlanetorg/ojs/jmla/rt/prINTERfriendly/329/503>. Volume 105, edn.: Journal of the Medical Library Association; 2017.
19. Harbour J, Fraser C, Lefebvre C, Glanville J, Beale S, Boachie C, Duffy S, McCool R, Smith L: **Reporting methodological search filter performance comparisons: a literature review**. *Health information and libraries journal* 2014, **31**(3):176-194.
 20. Jenkins M: **Evaluation of methodological search filters - a review**. In: *Health Information & Libraries Journal*. vol. 21; 2004: 148-163.
 21. Cooper C, Varley-Campbell J, Booth A, Britten N, Garside R: **Systematic review identifies six metrics and one method for assessing literature search effectiveness but no consensus on appropriate use**. *Journal of clinical epidemiology* 2018, **99**:53-63.
 22. Glanville J, Fleetwood, K., Yellowlees, A., Kaunelis, D., Mensinkai, S.: **Development and Testing of Search Filters to Identify Economic Evaluations in MEDLINE and EMBASE** In. Ottawa, Canada: Canadian Agency for Drugs and Technologies in Health 2009.
 23. **Cochrane Colloquium**. Available at: <https://colloquium.cochrane.org/news/what-colloquium>
 24. **G-I-N Conference**. Available at: <https://www.ginconference.net/homepage>

Supplementary file 1

Name of the filter: CADTH guidelines filter – broad

Last edited by CADTH: March 2, 2016

1. exp clinical pathway/
2. exp clinical protocol/
3. exp consensus/
4. exp consensus development conference/
5. exp consensus development conferences as topic/
6. critical pathways/
7. exp guideline/
8. guidelines as topic/
9. exp practice guideline/
10. practice guidelines as topic/
11. health planning guidelines/
12. exp treatment guidelines/
13. (guideline or practice guideline or consensus development conference or consensus development conference, NIH).pt.
14. (position statement* or policy statement* or practice parameter* or best practice*).ti,ab,kf,kw.
15. (standards or guideline or guidelines).ti,kf,kw.
16. ((practice or treatment* or clinical) adj guideline*).ab.
17. (CPG or CPGs).ti.
18. consensus*.ti,kf,kw.
19. consensus*.ab. /freq=2
20. ((critical or clinical or practice) adj2 (path or paths or pathway or pathways or protocol*)).ti,ab,kf,kw.
21. recommendat*.ti,kf,kw.
22. (care adj2 (standard or path or paths or pathway or pathways or map or maps or plan or plans)).ti,ab,kf,kw.
23. (algorithm* adj2 (screening or examination or test or tested or testing or assessment* or diagnosis or diagnoses or diagnosed or diagnosing)).ti,ab,kf,kw.
24. (algorithm* adj2 (pharmacotherap* or chemotherap* or chemotreatment* or therap* or treatment* or intervention*)).ti,ab,kf,kw.
25. or/1-24

Name of the filter: CADTH guidelines filter – narrow

Last edited by CADTH: August 27, 2010

1. (guideline or practice guideline or consensus development conference or consensus development conference, NIH).pt.
2. (guideline* or standards or consensus* or recommendat*).ti.
3. (practice parameter* or position statement* or policy statement* or CPG or CPGs or best practice*).ti.
4. (care adj2 (path or paths or pathway or pathways or map or maps or plan or plans or standard)).ti.

5. ((critical or clinical or practice) adj2 (path or paths or pathway or pathways or protocol*)).ti.
6. (algorithm* and (pharmacotherap* or chemotherap* or chemotreatment* or therap* or treatment* or intervention*)).ti.
7. (algorithm* and (screening or examination or test or tested or testing or assessment* or diagnosis or diagnoses or diagnosed or diagnosing)).ti.
8. or/1-7

Supplementary file 2

Name of the filter: University of Texas Guidelines/Recommendations

Last edited: March 3, 2015

clinical adj3 pathway).ti,ab,kw. or (clinical adj3 pathways).ti,ab,kw. or (practice adj3 parameter).ti,ab,kw. or (practice adj3 parameters).ti,ab,kw. or algorithms/ or care pathway.ti,ab,kw. or care pathways.ti,ab,kw. or clinical protocols/ or Consensus/ or Consensus Development Conference.pt. or Consensus Development Conference, NIH.pt. or Consensus Development Conferences as Topic/ or Consensus Development Conferences, NIH as Topic/or critical pathway/ or guidance.ti,ab. or guideline*.ti. or guidelines as topic/ or practice guidelines as topic/ or Health Planning Guidelines/ or practice guideline/

Supplementary file 3

Name of the filter: PubMed - Guidelines Filter

Last edited: 2017

(((((("Guideline" [Publication Type]) OR "Practice Guideline" [Publication Type]) OR "Consensus"[Mesh]) OR ("Consensus Development Conference, NIH" [Publication Type] OR "Consensus Development Conference" [Publication Type])) OR (consensuses[ti] or consensus[ti])) OR "position statement"[ti] OR "position statements"[ti] OR "practice parameter"[ti] OR "practice parameters"[ti]) OR "appropriate use criteria" [ti] OR "appropriateness criteria" [ti] OR ("guidance statement"[ti] OR "guidance statements"[ti])) OR (guideline[ti] or guidelines[ti])