

Validated geographic search filters for bibliographic databases: a scoping review protocol

Catharina Munte^{1,2} • Alexander Pachanov^{1,2,3} • Julian Hirt^{4,5,6} • Lynda Ayiku⁷ • Dawid Pieper^{1,2,3}

¹Faculty of Health Sciences Brandenburg, Brandenburg Medical School (Theodor Fontane), Institute for Health Services and Health System Research, Rüdersdorf, Germany, ²Center for Health Services Research, Brandenburg Medical School (Theodor Fontane), Rüdersdorf, Germany, ³Evidence Based Practice in Brandenburg: A JBI Affiliated Group, Brandenburg Medical School (Theodor Fontane), Brandenburg, Germany, ⁴Pragmatic Evidence Lab, Research Center for Clinical Neuroimmunology and Neuroscience Basel (RC2NB), University Hospital Basel and University of Basel, Basel, Switzerland, ⁵Department of Clinical Research, University Hospital Basel and University of Basel, Basel, Switzerland, ⁶Institute of Nursing Science, Department of Health, Eastern Switzerland University of Applied Sciences, St. Gallen, Switzerland, and ⁷Information Services, National Institute for Health and Care Excellence, Manchester, United Kingdom

ABSTRACT

Objective: The purpose of this scoping review is to identify validated geographic search filters and report their methodology and performance measures.

Introduction: Data on specific geographic areas can be required for evidence syntheses topics, such as the investigation of regional inequalities in health care or to answer context-specific epidemiological questions. Search filters are useful tools for reviewers aiming to identify publications with common characteristics in bibliographic databases. Geographic search filters limit the literature search results to a specific geographic feature (eg, a country or region).

Inclusion criteria: We will include reports on validated geographic search filters that aim to identify research evidence about a defined geographic area (eg, a country/region or a group of countries/regions).

Methods: This review will be conducted in accordance with JBI methodology for scoping reviews. The literature search will be conducted in PubMed and Embase. The InterTASC Information Specialists' Sub-Group Search Filter resource and Google Scholar will also be searched. Reports published in any language, from database inception to the present, will be considered for inclusion. Two researchers will independently screen the title, abstract, and full text of the search results. A third reviewer will be consulted in the event of any disagreements. The data extraction will include study characteristics, basic characteristics of the geographical search filter (eg, country/region), and the methods used to develop and validate the search filter. The extracted data will be summarized narratively and presented in a table.

Review registration: Open Science Framework <https://osf.io/5czhs>

Keywords: database searching; evidence synthesis; geographic search filter; information retrieval

JBI Evid Synth 2024; 22(3):441–446.

Introduction

The increasing number of publications in bibliographic literature databases in recent years requires measures to limit the search and screening efforts. Search filters are a useful tool to identify evidence with a common characteristic, such as a study

design, an intervention, or a population group.¹ They may be applied to literature search strategies with Boolean operators (AND or NOT),² by choosing a search option,³ or by filtering the results after searching the database (database-integrated search filters). Search filters should be developed using a systematic process; a key attribute is the validation of the retrieval performance. Validation is important because it provides a generalizable indication of the retrieval performance of a search filter.¹ It is an independent assessment of the actual performance of the search filter outside the data set used for

Correspondence: Catharina Munte, catharina.munte@mhb-fontane.de
DP is a member of the JBI Evidence Synthesis editorial advisory board, but was not involved in the editorial processing of the manuscript.
The other authors declare no conflict of interest.

DOI: 10.1112/JBIES-23-00445

derivation and/or testing. These tests determine how well a search filter performs in the database and system interface for which it was developed.¹ A well-functioning search filter is important because it identifies relevant studies in a time-efficient manner and minimizes the effort in the screening process. In addition, if there are multiple search filters with the same target feature, validation and performance measures can be used to help select a search filter that is appropriate for the research question.

According to Jenkins,¹ there are 3 types of search filters:

- First-generation search filters are non-validated filters developed subjectively based on experience regarding the database structure, content, and the use of synonyms.
- Second-generation search filters are validated filters that are developed subjectively, but their retrieval performance is tested using a gold standard (GS). A GS is a set of known, relevant references on the topic of the filter used to identify search terms for the filter or to evaluate the retrieval performance of the filter.^{1,4}
- Third-generation search filters are validated filters that inform development by applying objective methods (eg, frequency analysis) on a GS.¹ The retrieval performance of these filters is calculated (ie, validated) against an independent GS that was not used for the development of the filter.¹

When preparing evidence syntheses, it can be helpful to use search filters to limit the literature search according to geography. Ayiku *et al.*⁵ define a geographic search filter as “a collection of search terms designed to retrieve research with a common geographic feature from bibliographic databases.”^(p.200) The benefit of geographical search filters can vary depending on the research question. For example, certain health issues may be confined to specific geographical regions, whether due to environmental factors, cultural differences, or socioeconomic conditions. A geographical search filter facilitates finding

and analyzing evidence in bibliographic databases related to local health issues.

In addition, a geographic search filter can be useful to identify epidemiological evidence and enable researchers to investigate health care disparities and variations in health care practices across different regions. This is vital for understanding and addressing inequalities in health care delivery. Researchers, clinicians, and decision-makers can also use geographical search filters to identify health care issues and gain an understanding of research and care gaps. This can help to ensure that research efforts are tailored to the specific health needs and challenges of a geographical region. A systematically developed and validated search filter facilitates the reliable retrieval of studies on a specific topic and ensures reproducibility.⁶

Our aim is to conduct a scoping review to identify existing validated geographic search filters and summarize the development and validation methods for each of them. In addition, we will report the retrieval performance of the filters in terms of sensitivity (also known as recall),⁷ precision, specificity, and number needed to read (NNR), as indicated in Table 1.^{1,4,8,9}

Review question

Which validated geographic search filters exist, and what are the methods that have been employed for their development and validation?

Inclusion criteria

Concept

We will consider validated geographic search filters (ie, a filter that aims to identify studies from a country, group of countries, region, city) for inclusion in the scoping review, as validation allows conclusions to be drawn about the actual performance of the search filter. To provide a comprehensive map of the available evidence, we will consider validated search filters, regardless of the number of references, used for the GS and validation. The geographic search filter must

Table 1: Geographic search filter performance measures

Sensitivity ^{1,4,8}	The number of relevant records retrieved by the search filter as a proportion of the total number of relevant records.
Precision ^{1,4,8}	The number of relevant records retrieved as a proportion of the total number of records retrieved.
Specificity ^{1,4,8}	The number of records that are not relevant and are not retrieved as a proportion of the total number that are not relevant.
Number needed to read ⁹	The number of records that need to be read to identify 1 relevant record.

be validated using 1 or more GS sets, and report 1 or more filter performance measures (eg, sensitivity, specificity, precision). If a source does not assess quality/performance measures during the development phase and/or report them in the publication, it will be excluded.

Types of sources

We will consider all full-text reports on the development and validation of a geographic search filter, and there will be no restrictions on the study design. Conference papers and abstracts will be considered as potentially relevant sources, and we will contact the authors to obtain a full-text report if it is available. Preprints will be excluded from the review, as they have not undergone a peer-review process, which may result in lower quality and reliability of the published information.

Methods

We will conduct the scoping review in accordance with the JBI methodology for scoping reviews^{10,11} and report it using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis extension for Scoping Reviews (PRISMA-ScR).¹² We will not perform a risk of bias assessment or critical appraisal, as this is not recommended for scoping reviews.¹¹

Search strategy

The search strategy will aim to locate published and unpublished full-text reports on validated geographic search filters. An initial search of PubMed, Embase, Google Scholar, and the InterTASC Information Specialists' Sub-Group (ISSG) Search Filter Resource¹³ was undertaken to identify reports on the topic. The search terms in the titles and abstracts of the relevant reports, as well as the index terms used to describe the reports, were used to develop a full search strategy for PubMed, Embase, Google Scholar, and ISSG following the Peer Review Electronic Search Strategies (PRESS) guidelines.¹⁴ The search strategy was developed by AP and peer reviewed by CM. In addition, we will perform forward and backward citation screening for all included studies. A sample search strategy for PubMed is presented in Appendix I.

Reports published in any language and from database inception to the present will be considered for inclusion. Any reports that are not written in English or German (the languages the authors are proficient in) will be translated.

Source of evidence selection

After completing the literature search, we will upload all identified citations into EndNote v.2020.3 (Clarivate Analytics, PA, USA) and remove any duplicates. Potentially relevant sources will be transferred to Rayyan (Qatar Computing Research Institute, Doha, Qatar).¹⁵ As recommended, the screening process will be piloted on a subset of potentially relevant sources.¹⁶ We decided to conduct the pilot test with 5% of potentially relevant sources at the title/abstract level and one-third at the full-text level to ensure the consistency of the selection process. Any disagreements that arise between the reviewers at each stage of the selection process will be resolved through discussion or with additional reviewers. In the final scoping review, we will present the results of the literature search and the study inclusion process in a PRISMA flow diagram.¹⁷ Additionally, we will provide a reference list of both included and excluded reports, with reasons for exclusion for the latter.

Data extraction

A data extraction tool (Appendix II) developed by the reviewers will be used. The components of data extraction refer to the filter development and validation methodology as described by Jenkins¹ and the UK InterTASC ISSG Search Filter Appraisal Checklist.⁴ This checklist contains categorical and descriptive responses and includes a summary of the most important quality characteristics of search filters. The draft data extraction form will capture information on the database(s) for which the filter was developed; geographic features; search terms (eg, city names, federal states, universities); search fields and techniques (eg, wildcards, quotations); and methodological details on the filter development according to the ISSG Search Filter Appraisal Checklist.⁴ It should be noted that there may be multiple publications for a geographic search filter (eg, 1 publication for development and 1 for validation); therefore, multiple publications may be included for 1 geographic search filter.

Two reviewers will independently extract data from the reports included in the scoping review to reduce the chance of error and bias. Prior to this, the 2 reviewers will independently test 3 reports using the draft data extraction form to ensure consistency in the way they perform the extraction.¹¹ Subsequently, they will discuss the results and, in the event of a disagreement, consult a third reviewer. We will

start data extraction when substantial agreement has been achieved.

We will modify and revise the draft data extraction tool as necessary during the process. The changes will be documented in the scoping review. We will contact authors of reports once to request missing or additional data, where required.

Data analysis and presentation

Our aim is to provide an overview of existing geographic search filters and how they have been developed and validated. We also want to obtain information on the performance measures of geographic search filters (eg, sensitivity, precision, specificity, and NNR). This will be achieved using the appraisal checklist by Glanville *et al.*⁴ to identify and evaluate relevant aspects of search filter development. We will compile the results in a table accompanied by a narrative summary describing how the results relate to the goal of identifying methods for developing and validating geographic search filters. The narrative summary will refer to core aspects of the Glanville *et al.* checklist.⁴ We will use these methodological aspects to provide a descriptive analysis of the methods applied to develop and validate the geographic filters, as well as a narrative, thematic summary of the data collected. For this purpose, the search terms used (eg, city names, federal states, universities), search fields/limitations, and techniques will be documented. We will categorize, summarize, and discuss quantitative aspects (eg, frequencies) of the methods, as well as similarities and differences between the geographical search filters.

Author contributions

CM, AP, JH, and DP conceived the protocol. CM drafted the manuscript. All authors have critically reviewed, provided feedback on, and approved the final version of the protocol. DP is a guarantor of the scoping review protocol.

References

- Jenkins M. Evaluation of methodological search filters—a review. *Health Info Libr J* 2004;21(3):148–63.
- Wilczynski NL, Marks S, Haynes RB. Search strategies for identifying qualitative studies in CINAHL. *Qual Health Res* 2007;17(5):705–10.
- National Library of Medicine. PubMed clinical queries [internet]. NLM. n.d. [cited 2023 Oct 14]. Available from: <https://pubmed.ncbi.nlm.nih.gov/clinical/>.
- Glanville J, Bayliss S, Booth A, Dundar Y, Fernandes H, Fleeman ND, *et al.* So many filters, so little time: the development of a search filter appraisal checklist. *J Med Libr Assoc* 2008;96(4):356–61.
- Ayiku L, Levay P, Hudson T, Craven J, Barrett E, Finnegan A, *et al.* The medline UK filter: development and validation of a geographic search filter to retrieve research about the UK from OVID medline. *Health Info Libr J* 2017;34(3):200–16.
- Ayiku L, Craven J, Hudson T, Levay P. How to develop a validated geographic search filter: five key steps. *Evid Based Lib Inform Pract* 2020;15(1):170–8.
- 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;6:33.
- Terwee CB, Jansma EP, Riphagen II, Vet HCWd. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009;18(8):1115–23.
- Bachmann LM, Coray R, Estermann P, Riet Gt. Identifying diagnostic studies in MEDLINE: reducing the number needed to read. *J Am Med Inform Assoc* 2002;9(6):653–8.
- Munn Z, Pollock D, Khalil H, Alexander L, McInerney P, Godfrey CM, *et al.* What are scoping reviews? Providing a formal definition of scoping reviews as a type of evidence synthesis. *JBIM Evid Synth* 2022;20(4):950–2.
- Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, *et al.* Updated methodological guidance for the conduct of scoping reviews. *JBIM Evid Implement* 2021;19(1):3–10.
- Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, *et al.* PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018;169(7):467–73.
- Glanville JLC, Manson P, Robinson S, Shaw N. eTISS-G ISSG Search Filter Resource [internet]. ISSG; 2006 [cited 2023 Oct 14]. Available from: <https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home>.
- 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.
- Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Syst Rev* 2016;5(1):210.
- Peters MDJ, Godfrey C, McInerney P, Khalil H, Larsen P, Marnie C, *et al.* Best practice guidance and reporting items for the development of scoping review protocols. *JBIM Evid Synth* 2022;20(4):953–68.
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.

Appendix I: Search strategy

PubMed
Search conducted: August 31, 2023.

Search	Query	Records retrieved
#1	("geographic filter"[Title/Abstract:~3] OR "geographic filters"[Title/Abstract:~3] OR (("geographic*"[Title/Abstract] OR "terror*"[Title/Abstract] OR "provinc*"[Title/Abstract] OR "region*"[Title/Abstract]) AND ("search filter"[Title/Abstract] OR "search filters"[Title/Abstract]))) AND ("retriev*"[Title/Abstract] OR "recall"[Title/Abstract] OR "reliab*"[Title/Abstract] OR "perform*"[Title/Abstract] OR "evaluat*"[Title/Abstract] OR "develop*"[Title/Abstract] OR "validat*"[Title/Abstract] OR "sensitiv*"[Title/Abstract] OR "specifi*"[Title/Abstract] OR "precis*"[Title/Abstract] OR "number needed to read"[Title/Abstract] OR "NNR"[Title/Abstract] OR "positive predictive value*"[Title/Abstract] OR "PPV"[Title/Abstract] OR "negative predictive value*"[Title/Abstract] OR "NPV"[Title/Abstract] OR "sensitivity and specificity"[MeSH Terms])	31

Appendix II: Draft data extraction instrument

No.	Author	Year	Database(s) for which the filter was developed	Geographic feature	Search terms (eg, city names, federal states, universities)	Search fields and techniques	Methodological details on the filter development according to the ISSG Search Filter Appraisal Checklist ⁴					
							(A) Objective	(B) Identification of a GS of known relevant records (methodology and number of GS)	(C) Identification of search terms (ie, objective or subjective approach)	(D) Internal validity (ie, number of filters tested, characteristics and number of GS, values of reported performance measures)	(E) External validity (ie, number of filters tested, characteristics and number of GS, values of reported performance measures)	(F) Limitations and comparisons (reported study limitations)
1												
2												
3												
4												
5												
6												

GS, gold standard; ISSG, Information Specialists' Sub-Group.