Performance of a mixed filter to identify relevant studies for mixed studies reviews EC

Reem El Sherif, MBBCh; Pierre Pluye, MD, PhD; Genevieve Gore, MLIS; Vera Granikov, MLIS; Quan Nha Hong, PhD(c), MSc

See end of article for authors' affiliations.

Objective: Mixed studies reviews include empirical studies with diverse designs. Given that identifying relevant studies for such reviews is time consuming, a mixed filter was developed.

Methods: The filter was used for six journals from three disciplines. For each journal, database records were coded "empirical" (relevant) when they mentioned a research question or objective, data collection, analysis, and results. We measured precision (proportion of retrieved documents being relevant), sensitivity (proportion of relevant documents retrieved), and specificity (proportion of nonrelevant documents not retrieved).

Results: Records were coded with and without the filter, and descriptive statistics were performed, suggesting the mixed filter has high sensitivity.

Keywords and Medical Subject Headings (MeSH): Database, Bibliographic; Review, Systematic; Information Retrieval Systems; Empirical Research; Medical Informatics

An increasing number of researchers, practitioners, and policy makers are using systematic reviews to keep their knowledge up-to-date in the current context of managing a rapidly growing number of scientific publications [1]. In a mixed studies review, a team of reviewers reviews all types of empirical research (qualitative, quantitative, or mixed methods) concurrently to develop a breadth and depth of understanding of scientific knowledge [2]. Mixed studies reviews are a type of systematic review that is becoming popular in all health disciplines because they can address complex research questions [3]. Significant methodological advancement of mixed studies reviews have been seen in the past decade with the development of numerous synthesis methods for qualitative and quantitative evidence as well as frameworks for mixing evidence [2, 4–8]. Furthermore, guidance for researchers designing, conducting, and reporting systematic mixed studies reviews has been developed and is accessible in an open-access format [9].

EC A supplemental appendix is available with the online version of this journal.

The identification of potentially relevant studies in bibliographic databases constitutes a key stage of reviews. Because of the high number of scientific publications (e.g., estimated at more than 50 million), these bibliographic searches now yield thousands of records that require manual screening for relevance by the reviewers [10, 11]. Thus, the identification of potentially relevant studies is time consuming and labor intensive in systematic reviews.

Traditional search strategies in bibliographic databases have high specificity and sensitivity for finding randomized controlled trials but are limited for other types of research studies [12, 13].

In addition, mixed studies reviews search strategies yield a high number of irrelevant records, such as nonempirical work (e.g., commentaries, editorials, and opinion letters). The manual screening of thousands of irrelevant records is an extremely time- and resource-consuming process [14]. However, there is no research on the best strategies for identifying qualitative, quantitative, and mixed methods studies. Such research has the potential to greatly facilitate mixed studies reviews. To our knowledge, there is no database filter to retrieve

studies with diverse qualitative, quantitative, and mixed methods research designs.

The authors' research objective was to develop and evaluate a database filter to retrieve studies with diverse qualitative, quantitative, and mixed methods research designs. Our research question was: What are the precision, sensitivity, and specificity of this filter?

METHODS

We developed the mixed filter (Appendix, online only) with librarians and researchers who had expertise in systematic literature reviews using (a) validated filters for randomized controlled trials [15], non-randomized and descriptive quantitative studies [16–18], and qualitative research [19], and (b) a common filter for mixed methods research [20, 21]. The filter includes different keywords and subject headings for quantitative (e.g., cohort study), qualitative (e.g., focus group), and mixed methods (e.g., multimethod). It was developed for MEDLINE and can be adapted for different bibliographic databases. It is available free via the "Apply an extensive search strategy" page of a mixed studies reviews wiki [9].

We assessed this filter in Ovid MEDLINE. We needed a mix of qualitative, quantitative, and mixed method research studies, along with nonempirical studies. We looked for two "extreme case" types of journals: one with a high proportion of empirical research and one with a low proportion. Assuming that the former is frequently cited by researchers and the latter rarely cited, we used the impact factor as a proxy of this proportion (Journal Citation Reports in ISI Web of Knowledge 2013). We selected two journals (one with a high impact factor and one with a low impact factor) from three disciplines: Primary Care, Medical Informatics, and Public Health and Epidemiology. The rationale for this approach was twofold: (1) in accordance with the literature on multiple case studies and extreme case analysis, a pattern that crosses over extreme cases is deemed to be theoretically transferable (but not statistically generalizable) to all cases (replication logic) [22, 23], and (2) these three disciplines face complex research questions and are known to publish studies with diverse designs.

In Primary Care, the journals we chose were *Annals of Family Medicine* and the *Journal of Family Practice* (impact factor being 4.57 and 0.74,

respectively). In Public Health and Epidemiology, they were the *International Journal of Epidemiology* and the *Journal of Palliative Medicine* (impact factor being 4.67 and 1.23, respectively). In Medical Informatics, they were the *Journal of Medical Internet Research* and *Biomedical Engineering* (impact factor being 9.19 and 2.85, respectively).

There is no consensus on the minimum number of records to calculate the performance of filters [24]. To obtain a manageable sample of database records with at least 250 records per journal, we focused on articles published between 2008 and 2013. We measured the performance of the mixed filter in terms of (a) precision: the proportion of retrieved documents being relevant, (b) sensitivity: the proportion of relevant documents being retrieved, and (c) specificity: the proportion of nonrelevant documents not retrieved [24]. High performance refers to precision, sensibility, and specificity above 50%, 80%, and 50%, respectively. Equations are:

- Precision=(Relevant Retrieved)/(All Retrieved)
- Sensitivity=(Relevant Retrieved)/(All relevant)
- Specificity=(Not relevant Not retrieved)/(All not relevant)

For each journal, the identification and coding for relevance of the database records were conducted with and without the filter. First, all database records from each journal were identified using MEDLINE without the filter and imported into specialized review software (Distiller SR). "Relevant" documents were those that contained empirical research (i.e., they mentioned at least a research question/objective, a qualitative or quantitative data collection/analysis, and results). A researcher coded these documents as relevant or not based on those criteria. If there was any uncertainty, the document was referred to a second researcher until a consensus was reached.

Full-text papers were assessed when records had no abstracts or the abstract was unclear. In the second step, the mixed filter was combined with the search results of each journal, and the retrieved studies were imported and coded as empirical (relevant) or not.

The descriptive statistics measured for each journal were: precision, sensitivity, and specificity. The overall precision, sensitivity, and specificity were also calculated by combining the total number of records in all six journals, the total number of records

Journal	Number of relevant records retrieved (a*)	Number of relevant records not retrieved by the filter (b*)	Number of nonrelevant records retrieved by the filter (c*)	Number of nonrelevant records not retrieved by the filter (d*)	Precision (%)	Sensitivity (%)	Specificity (%)
Discipline 1: Primary Care							
High IF:							
Ann Fam Med	294	21	78	254	79.0	93.3	76.5
Low IF:							
J Fam Pract	23	2	492	402	4.5	92.0	45.0
Discipline 2: Public Health and Epidemiology							
High IF:							
Int J Epi	488	9	317	495	60.6	98.2	61.0
Low IF:							
Palliative Med	291	26	162	142	64.2	91.8	46.2
Discipline 3: Medical Informatics							
High IF:							
J Med Int R	577	64	68	52	89.5	90.0	43.3
Low IF:							
Biomed Eng	103	87	47	53	68.7	54.2	53.0
Overall							
6 included journals	1,176	208	1,164	1,396	60.4%	89.5%	54.5%

IF=impact factor.
* Legend:
High precision ≥50%=a:(a+c)
High sensitivity ≥80%=a:(a+b)
High specificity ≥50%=d:(c+d)
a=relevant records retrieved; b=relevant records not retrieved; c=nonrelevant records retrieved; d=nonrelevant records not retrieved.

Table 1
Performance of the mixed filter

retrieved with the filter, and the total number of relevant records.

RESULTS

Detailed results are presented in Table 1. In total, 4,547 records were identified in the 6 journals. Of these, 2,940 were retrieved by the filter. The journal with the most records was the *International Journal of Epidemiology* (1,309 records) and that with the fewest records was *Biomedical Engineering* (290 records). The overall precision, sensitivity, and specificity (calculated by grouping the results for each journal) were high for all 6 journals (60.4%, 89.5%, and 54.5%, respectively).

Results by discipline show that all measures for the Primary Care journals were generally high for the high impact factor journal, while precision was extremely low for the low impact factor journal. For the Public Health journals, the measures were once again high for the high impact factor journal, while only specificity was low for the low impact factor journal. Finally, for the Medical Informatics journal, the precision and sensitivity were high for the high

impact factor journal, while sensitivity was low for the low impact factor journal.

DISCUSSION

Overall, the mixed filter showed high performance, though the performance varied by journal. Sensitivity is important for systematic mixed studies reviews because researchers aim to achieve a comprehensive and exhaustive retrieval of database records: a high sensitivity means that almost all the relevant records are retrieved. The sensitivity of the mixed filter was high (≥90%) across the 3 disciplines for journals with a high impact factor. Only 1 low impact factor journal had a sensitivity level lower than 90% (54.2% for *Biomedical Engineering*). This may be due to the small number of articles published by the journal and abstracts that did not include typical research-related terms.

In journals with a high impact factor, the mixed filter showed high specificity and precision (≥60%), except for the *Journal of Medical Internet Research*, which had moderate specificity. We believe this may be due to the larger number of empirical studies

published in those journals, compared to those with a low impact factor (Table 1). Regarding low impact factor journals, specificity and precision varied. For example, the precision for the low impact factor journal in Primary Care (*Journal of Family Practice*) was 4.5%. This may be due to the low number of empirical studies published in this professional journal over the last 6 years (Table 1) and the large number of nonempirical records that were reported as empirical studies, with abstract, methods, and results sections.

Given the filter's sensitivity was not 100%, some relevant papers might not be captured; thus, a complementary citation search (screening backward and forward citations of included studies) may overcome the limitations of the filter for systematic reviews where the search must be comprehensive. Because citation searches are recommended by guidelines on systematic reviews anyway, this does not increase the workload [25].

To our knowledge, no other similar mixed filter has been tested, so we are unable to compare our results to the performance results of other filters. We obtained promising results with high sensitivity. The mixed filter can be useful to reduce the workload of identifying potential papers for inclusion in mixed studies reviews. Future research can explore filtering using text mining and semi-automated text classification methods.

LIMITATIONS

The mixed filter was tested in a limited number of journals. Journals and disciplines other than the ones we used may yield differing results. Notably, we sought to mitigate issues associated with our sample through the use of the extreme case analysis approach.

REFERENCES

1. Larsen PO, von Ins M. The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. Scientometrics. 2010 Sep;84(3): 575–603.

2. Pluye P, Hong QN. Combining the power of stories and the power of numbers: mixed methods research and mixed studies reviews. Annu Rev Public Health. 2014 35: 29–45.

3. Shaw RL, Larkin M, Flowers P. Expanding the evidence within evidence-based healthcare: thinking about the

context, acceptability and feasibility of interventions. Evid Based Med. Dec;19(6):201–3.

4. Kastner M, Tricco AC, Soobiah C, Lillie E, Perrier L, Horsley T, Welch V, Cogo E, Antony J, Straus SE. What is the most appropriate knowledge synthesis method to conduct a review? protocol for a scoping review. BMC Med Res Methodol. 2012 Aug 3;12:114.

5. Mays N, Pope C, Popay J. Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. J Health Serv Res Policy. 2005 Jul;10(suppl 1):6–20.

6. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantitative evidence: a review of possible methods. J Health Serv Res Policy. 2005 Jan;10(1):45–53.

7. Heyvaert M, Hannes K, Maes B, Onghena P. Critical appraisal of mixed methods studies. J Mixed Methods Res. 2013 Oct;7:302–27.

8. Sandelowski M, Voils CI, Barroso J. Defining and designing mixed research synthesis studies. Res Sch. 2006 Spring;13(1): 29.

9. Pluye P, Hong QN, VedelI. Toolkit for mixed studies reviews [Internet]. 2013 [cited 3 Dec 2014]. <http://toolkit4mixedstudiesreviews.pbworks.com>.

10. Bjork BC, Roos A, Lauri M. Scientific journal publishing: yearly volume and open access availability. Inf Res Int Electronic J. 2009 Mar;14(1).

11. Jinha AE. Article 50 million: an estimate of the number of scholarly articles in existence. Learned Publishing. 2010 Jul;23(3):258–63.

12. Lefebvre C, Manheimer E, Glanville J. Searching for studies. In: Cochrane handbook for systematic reviews of interventions. New York, NY: Wiley; 2008. p. 95–150.

13. McKibbon KA, Lokker C, Wilczynski NL, Ciliska D, Dobbins M, Davis DA, Haynes RB, Straus SE. A cross-sectional study of the number and frequency of terms used to refer to knowledge translation in a body of health literature in 2006: a Tower of Babel. Implement Sci. 2010 Feb 12;5(1):16.

14. Gough D, Oliver S, Thomas J. An introduction to systematic reviews. Thousand Oaks, CA: Sage; 2012.

15. Robinson KA, Dickersin K. Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. Int J Epidemiol. 2002 Feb; 31(1):150–53.

16. Wilczynski NL, Haynes RB, Team H. Developing optimal search strategies for detecting clinically sound causation studies in MEDLINE. In: 2003 AMIA Annual Symposium Proceedings. American Medical Informatics Association.

17. Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey. BMC Med. 2004 Jun 9;2(1):23.

18. Furlan A, Irvin E, Bombardier C. Limited search strategies were effective in finding relevant nonrandomized studies. J Clin Epidemiol. 2006 Dec; 59(12):1303–11.

19. Grant MJ. How does your searching grow? a survey of search preferences and the use of optimal search strategies in the identification of qualitative research. Health Info Lib J. 2004 Mar;21(1):21–2.

20. Bryman A. Integrating quantitative and qualitative research: how is it done? Qualitative Res. 2006 Feb;6(1): 97–113.

21. Creswell JW, Plano Clark VL. Designing and conducting mixed methods research. Thousand Oaks, CA: Sage; 2011.

22. Patton M. Qualitative evaluation and research methods. Thousand Oaks, CA: Sage; 2005.

23. Yin RK. Case study research: design and methods. Thousand Oaks, CA: Sage; 2009.

24. Hersh WR. Information retrieval: a health and biomedical perspective. New York, NY: Springer; 2008.

25. Furlan AD, Pennick V, Bombardier C, van Tulder M; Editorial Board, Cochrane Back Review Group. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. Spine (Phila Pa 1976). 2009 Aug 15;34(18):1929–41.

AUTHORS' AFFILIATIONS

Reem El Sherif, MBBCh, reem.elsherif@mail.mcgill.ca, Family Medicine; **Pierre Pluye,** MD, PhD, pierre.pluye@mcgill.ca, Family Medicine; **Genevieve Gore,** MLIS, genevieve.gore@mcgill.ca, Life Sciences Library; **Vera Granikov,** MLIS, veragranikov@gmail.com, Family Medicine; **Quan Nha Hong,** PhD(c), MSc, quan.nha.hong@mail.mcgill.ca, Family Medicine; McGill University, Montreal, QC, Canada