

Limits of search filter development*

Nancy L. Wilczynski, PhD; Cynthia Lokker, PhD; Kathleen Ann McKibbin, PhD, FMLA;
Nicholas Hobson, R. Brian Haynes, MD, PhD

See end of article for authors' affiliations.

DOI: <http://dx.doi.org/10.3163/1536-5050.104.1.006>

Objective: The research attempted to develop search filters for biomedical literature databases that improve retrieval of studies of clinical relevance for the nursing and rehabilitation professions.

Methods: Diagnostic testing framework compared machine-culled and practitioner-nominated search terms with a hand-tagged clinical literature database.

Results: We were unable to: (1) develop filters for nursing, likely because of the overlapping and expanding scope of practice for nurses in comparison with medical professionals, or (2) develop filters for rehabilitation, because of its broad scope and the profession's multifaceted understanding of "health and ability."

Conclusions: We found limitations on search filter development for these health professions: nursing and rehabilitation.

Keywords: Nursing; Occupational Therapy; Physical Therapy Specialty; Databases, Bibliographic; Search Engine; Terminology as Topic; Information Storage and Retrieval

It is a challenge for health professionals to search large biomedical databases such as MEDLINE to find studies of high quality and relevance for their clinical practice and to avoid becoming overwhelmed by articles that are irrelevant or of low quality. Search filters ("hedges") have been successfully developed and validated for use in such large electronic databases to retrieve studies that are scientifically sound and clinically relevant (e.g., studies designed to answer questions relating to the effectiveness of a therapy or the accuracy of a diagnostic test [1, 2]). Hedges are also useful for detecting literature relevant to a specific disease or medical discipline (e.g., chronic kidney disease, nephrology, mental health [3–5]).

The use of these search filters can increase the proportional retrieval of relevant articles ("sensitivity") and reduce the proportional retrieval of off-target articles ("specificity") [1–5]. Many of these search filters have been empirically derived by

our research team at McMaster University and are available for use in PubMed on the Clinical Queries <<http://www.ncbi.nlm.nih.gov/pubmed/clinical>> and Health Services Research (HSR) PubMed Queries <<http://www.nlm.nih.gov/nichsr/hedges/search.html>> pages, as well as on the Ovid and EBSCO platforms for MEDLINE, EMBASE, PsycINFO, and CINAHL <http://hiru.mcmaster.ca/hiru/HIRU_Hedges_home.aspx>. These filters have been derived by applying a set of terms and phrases to a dataset of articles that were tagged as relevant or not relevant to the specific article types [1–5]. This approach allows the sensitivity, specificity, and accuracy of search filters to be determined.

The filters developed to date were calibrated for retrieving articles known to be relevant to medical practice by physicians. Search filters might also be useful for retrieving articles that are relevant to other professional groups, such as nurses and rehabilitation specialists (occupational therapists [OTs] and physical therapists [PTs]), but the authors expected that these professionals' information needs and interests might be different than physicians' needs and interests and require different search filters. For example, nurses might be more interested in studies

* This project was supported by Canadian Institutes of Health Research grant no. 230790.



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

that tested innovations in nursing care than in those testing a surgical technique delivered by a surgeon.

Further, OTs and PTs might have more use for studies of nondrug treatments than studies of drug therapies. Search filter development for the nursing profession has been limited and has focused on specific aspects of nursing care [6–8], and, to our knowledge, research on empirically derived search filters for rehabilitation professionals is nonexistent.

The authors set out to address the following two research questions: (1) Can search filters be developed for bibliographic databases (such as MEDLINE) to retrieve the clinical care articles that are most relevant for nurses? (2) Can search filters be developed to retrieve the most relevant clinical articles for rehabilitation professionals (OTs and PTs)?

METHODS

Nurses

We assembled an expert panel of nurse leaders from Canada, the United States, Australia, New Zealand, and the United Kingdom (see “Acknowledgments”). The panel members were asked to recommend nursing textbooks and websites with content “that nurses consider to be the heart of their discipline (roles and issues related to nursing practice that reflect the scope and breadth of nursing practice).” Programmers in the Health Information Research Unit (HiRU) at McMaster University used software called “Text Miner” to mine the text of these nominated sources and created spreadsheets listing the frequency of individual, two adjacent, and three adjacent text words. To test their ability to retrieve articles of relevance to nurses, these terms were used in search filter development and applied to the McMaster PLUS database <http://hiru.mcmaster.ca/hiru/HIRU_McMaster_PLUS_projects.aspx>.

The McMaster PLUS database has been created by research staff in HiRU who critically appraise the content of more than 120 clinically relevant journals <<http://hiru.mcmaster.ca/hiru/journalslist.asp>> on an ongoing basis. Articles that meet explicit criteria <<http://hiru.mcmaster.ca/hiru/InclusionCriteria.html>> are included in the database after they have been rated for clinical relevance and newsworthiness by practicing clinicians <http://hiru.mcmaster.ca/more/physicians/sample_rating_form.htm>, including physicians, nurses, and rehabilitation specialists. An article passing our criteria is included

in the database if at least three members of a given clinical discipline rate the article four or above on seven-point rating scales (seven high) for both relevance and clinical interest. All articles of interest to nurses are available through Nursing+ <<http://plus.mcmaster.ca/NP/Default.aspx>>. This tagging is checked by a faculty member from the School of Nursing at McMaster University and confirmed by at least three nurses in current clinical practice.

In addition to applying terms and phrases identified from nursing texts, we used Text Miner on the contents of the McMaster PLUS database to determine if there were indexing terms and text words that would differentiate articles of interest to nurses from those that were of interest to physicians, but not nurses, that could be used to create search filters, and we calculated the amount of overlap in articles of interest to both nurses and physicians.

Rehabilitation professionals

To help define articles of interest to rehabilitation professionals, specifically OTs and PTs, we recruited three faculty members (two OTs and one PT), two from the School of Rehabilitation Sciences at McMaster University and one from the Department of Occupational Science and Occupational Therapy at the University of Toronto (see “Acknowledgments”). These faculty members worked as a team to define the content of interest to OTs and PTs. After several iterations, the definitions were finalized (Appendix, online only). Four graduate students (two OTs and two PTs) were then recruited to determine the inter-rater reliability when applying these definitions to the research literature. These students worked independently and in duplicate, reviewing two sets of thirty articles, and indicated if the content was of interest to OTs, PTs, or both. We used a mix of articles, those tagged of interest to rehabilitation professionals in the McMaster PLUS database and available through Rehab+ <<http://plus.mcmaster.ca/rehab/Default.aspx>> and a random sample of articles that were not included in Rehab+. Each set of thirty articles included eighteen articles from Rehab+ and twelve articles not included in Rehab+.

RESULTS

Nurses

Mining the text from identified nursing textbooks and websites yielded 164,953 unique terms or

Leading 3-term combinations	Sensitivity	Specificity	Precision	Accuracy
Objectives OR behaviors OR health research	69.1%	55.2%	73.1%	64.1%
Objectives OR primary health care OR health research	66.6%	55.8%	72.7%	62.7%
Objectives OR health research OR population health	66.6%	55.8%	72.7%	62.7%
Objectives OR health organization OR health research	66.5%	55.6%	72.5%	62.5%
Care OR primary care OR reviews	66.1%	55.4%	72.3%	62.2%

* Term combinations maximized sensitivity. Single terms had to have sensitivity of at least 60% and specificity of at least 55% to be considered for term combinations. Sensitivity is the proportion of nurse-relevant articles retrieved; specificity is the proportion of nurse-nonrelevant articles not retrieved; precision is the proportion of articles retrieved that are relevant to nurses; accuracy is the proportion of all articles in the database that are correctly identified.

Table 1

Ability of top performing 3-term search strings, compiled from 164,953 candidate terms, in separating articles relevant to nurses from articles relevant to physicians but not nurses in the McMaster PLUS database*

phrases (up to 3 adjacent text words). The majority (over 155,000) of the terms or phrases had a frequency of less than 9 occurrences. The term or phrase with the highest frequency, “health,” occurred over 27,000 times. “Community” was the second most frequently occurring term or phrase with almost 9,000 occurrences, followed by “nursing” with just over 7,000 occurrences. Due to the common occurrence of these terms in the medical research literature as a whole, that is, whether of direct interest to nursing practice or not, it did not appear that these terms or phrases would be helpful in differentiating nursing clinical care articles from other articles published in large biomedical electronic databases such as MEDLINE. For example, when reviewing the first page of results when searching in PubMed using the term “health” (search conducted April 13, 2015), 9 of the first 20 retrieved articles were not relevant to clinical care (e.g., basic science articles), regardless of profession.

We tested the retrieval performance of 1-, 2- and 3-term strings from the 164,953 candidate terms from text mining in the McMaster PLUS database, comparing articles judged by nurses to be relevant to nursing practice with articles judged by nurses to not be relevant. The best performing 3-term string, “objectives OR behaviors OR health research,” had a sensitivity of 69% and specificity of 55% (Table 1). Thus, the best search filter that we could find would fail to retrieve more than 30% of the articles that were relevant to nursing, while retrieving 45% of articles that were not relevant.

We reviewed the contents of the McMaster PLUS database to determine if any indexing and/or text

words differentiated articles tagged of interest to nurses from those tagged of interest to physicians but not nurses. We found no terms specific to nursing practice, and the content that was tagged of interest to nurses was also of interest to physicians. Indeed, when reviewing the contents of the McMaster PLUS database, we found that a high percentage (46%–74%) of articles of interest to physicians were also of interest to the nurses and that the content of these articles spanned most general and speciality areas of clinical practice.

We additionally explored developing a filter that discarded content that was not of interest to nurses, (e.g., adding unique terms identified from articles that were tagged as “physician and not nurse” to search phrases with the Boolean NOT) but had no success in developing such a search filter.

Rehabilitation professionals

The levels of agreement in determining if published articles were of interest to OTs and PTs are shown in Table 2. Although we were able to achieve agreement among the expert panel members on definitions of the content areas of interest to OTs and PTs (Appendix, online only), we found low levels of agreement concerning articles of interest when these definitions were applied independently by the graduate students to the literature, even within each group (i.e., separately for OTs and PTs) and after additional training sessions. For example, in the second round of reviewing articles, the chance-corrected level of agreement for OTs was only 46% (95% confidence interval [CI] 10% to 81%) and 47% (12% to 81%) for PTs.

Training session, reviewer type, response categories	Number of disagreements (n=30)	Chance-adjusted level of agreement, unweighted kappa (95% CI)
First round, set of 30 articles, response categories Yes/No/Unsure		
Within discipline review:		
OT pair rating for "of interest to OT"	4	0.67 (0.33 to 1.00)
PT pair rating for "of interest to PT"	16	0.13 (-0.02 to 0.28)
Across discipline review:		
OT pair rating for "of interest to PT"	7	0.57 (0.26 to 0.86)
PT pair rating for "of interest to OT"	20	0.13 (0.004 to 0.26)
Second round, new set of 30 articles, response categories Yes/No		
Within discipline review only:		
OT pair rating for "of interest to OT"	8	0.46 (0.10 to 0.81)
PT pair rating for "of interest to PT"	8	0.47 (0.12 to 0.81)
CI=confidence interval.		

Table 2

Level of agreement in determining if published articles are "of interest to occupational therapists (OTs)" and/or "of interest to physiotherapists (PTs)"

DISCUSSION

Nurses

We found that developing search filters to optimally retrieve clinical care articles that are relevant for nurses was not possible, because we were not able to define this content area in a manner that allowed the accurate retrieval of articles that are of interest to this professional group. The accuracy of retrieval of the top terms was less than 64%, with high error rates for missing relevant articles and retrieving irrelevant articles. By contrast, a "sensitive" filter for retrieving studies of the treatments of health disorders had a sensitivity of 99% and specificity of 70% [1], while the terms and phrases most commonly found in nursing texts and curricula achieved a highest sensitivity of only 69%, with a corresponding specificity of only 55%.

This poor performance may be due in part to the lack of distinct boundaries of nursing practice, notably overlapping with medicine; the rapid expansion currently taking place in the scope of professional practice for nurses; and differences in the scope of practice within and across jurisdictions. Indeed, we do not think that creating search filters for detecting the clinical care articles that are most relevant for nurses is feasible or even plausible: the field is a broad, expanding, and changing target. Such filters may also be unnecessary to the extent that nurses are interested in much the same health research literature as medical practitioners, for which

the existing filters have both high sensitivity and specificity.

We limited our search for high-performing filters to 1, 2, and 3 terms because of diminishing returns in sensitivity for additional terms, offset by reductions for specificity. For example, the best single-term filter ("well-being") had a sensitivity of 60% with a specificity of 57%, compared with a sensitivity of 69% and specificity of 55% for the best 3-term filter. Adding terms could have increased the sensitivity to a little over 70%, but that is still not acceptable performance, and the specificity would have continued to fall.

Other researchers [6–8] have had some success in developing search filters for detecting the nursing literature but the focus of their work was different than ours. Our work was focused on retrieving articles that are of interest to practicing nurses in general, whereas others focused on a specific article type (e.g., retrieving diagnostic articles that are relevant to nurses [6, 7]) or a specific area of nursing practice (e.g., nurse staffing research [8]).

Rehabilitation professionals

Search filter development to retrieve articles that were most relevant for rehabilitation professionals was not possible in our study because of the low levels of agreement among OTs and PTs when applying the definitions of interest to the research literature. As such, we were unable to compile a dataset of relevant and not relevant articles as testing

ground for search terms. The low levels of agreement might reflect the broad scope and multifaceted understanding of “health and ability” of the rehabilitation professions.

This research was limited by a small sample size of articles for the rehabilitation science reliability exercises. This gave somewhat broad confidence intervals around estimates of reliability. Nevertheless, these estimates showed unreliable agreement, with no improvement in the second round.

ACKNOWLEDGMENTS

We acknowledge, with thanks, the contributions of the following colleagues:

- Nursing Expert Panel: Donna Ciliska, McMaster University, School of Nursing, Hamilton, Ontario, Canada; Ann Mohide, McMaster University, School of Nursing, Hamilton, Ontario, Canada; Marlene Cohen, University of Nebraska Medical Center, College of Nursing, Omaha, Nebraska, United States; Peter Griffiths, University of Southampton, School of Health Sciences, Southampton, United Kingdom; and Andrew Jull, University of Auckland, School of Nursing, Auckland, New Zealand
- Rehabilitation Science Expert Panel: Joy MacDermid, McMaster University, School of Rehabilitation Sciences, Hamilton, Ontario, Canada; Brenda Vrkljan, McMaster University, School of Rehabilitation Sciences, Hamilton, Ontario, Canada; and Heather Colquhoun, University of Toronto, Department of Occupational Science and Occupational Therapy, Toronto, Ontario, Canada
- Rehabilitation graduate students: Joshua Vincent, St. Joseph’s Healthcare London, Roth-McFarlane Hand and Upper Limb Center, London, Ontario, Canada; Neha Chugh-Gupta, McMaster University, School of Rehabilitation Sciences, Hamilton, Ontario, Canada; Michael Cammarata, McMaster University, School of Rehabilitation Sciences, Hamilton, Ontario, Canada; and Folarin Babatunde, McMaster University, School of Rehabilitation Sciences, Hamilton, Ontario, Canada.

REFERENCES

1. Haynes RB, McKibbin KA, Wilczynski NL, Walter SD, Werre SR; Hedges Team. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: an analytic survey. *BMJ*. 2005 May 21;330(7501):1179.
2. Haynes RB, Wilczynski NL. Optimal search strategies for retrieving scientifically strong studies of diagnosis from MEDLINE: an analytic survey. *BMJ*. 2004 May 1;328(7447):104.
3. Iansavichus AV, Hildebrand AM, Haynes RB, Wilczynski NL, Levin A, Hemmelgarn BR, Tu K, Nesrallah GE, Nash DM, Garg AX. High-performance information search filters for CKD content in PubMed, Ovid MEDLINE, and MEDLINE. *Am J Kidney Dis*. 2015 Jan;65(1):26–32.
4. Garg AX, Iansavichus AV, Wilczynski NL, Kastner M, Baier LA, Shariff SZ, Rehman F, Weir M, McKibbin KA, Haynes RB. Filtering Medline for a clinical discipline: diagnostic test assessment framework. *BMJ*. 2009 Sep 18;339:b3435.
5. Wilczynski NL, Haynes RB; Team Hedges. Optimal search strategies for identifying mental health content in MEDLINE: an analytic survey. *Ann Gen Psychiatry*. 2006 Mar 23;5:4.
6. Berg A, Fleischer S, Behrens J. Development of two search strategies for literature in MEDLINE-PubMed: nursing diagnoses in the context of evidence-based nursing. *Int J Nurs Terminol Classif*. 2005 Apr–Jun;16(2):26–32.
7. Lavin MA, Krieger MM, Meyer GA, Spasser MA, Cvitan T, Reese CG, Carlson JH, Perry AG, McNary P. Development and evaluation of evidence-based nursing (EBN) filters and related databases. *J Med Lib Assoc*. 2005 Jan;93(1):104–15.
8. Simon M, Hausner E, Klaus SF, Dunton NE. Identifying nurse staffing research in Medline: development and testing of empirically derived search strategies with the PubMed interface. *BMC Med Res Methodol*. 2010 Aug 23;10:76.

AUTHORS’ AFFILIATIONS



Nancy L. Wilczynski, PhD, wilczyn@mcmaster.ca, Assistant Professor and Research Manager (retired); **Cynthia Lokker, PhD**, lokker@mcmaster.ca, Assistant Professor (part-time) and Research Associate; **Kathleen Ann McKibbin, PhD, FMLA**, mckib@mcmaster.ca Professor; **Nicholas Hobson**, hobson@mcmaster.ca, Computer Programmer; **R. Brian Haynes, MD, PhD** (corresponding author), bhaynes@mcmaster.ca, Professor; Health Information Research Unit, CRL 125, McMaster University, 1280 Main Street West, Hamilton, ON, Canada, L8S 4K1

Received April 2015; accepted August 2015