Development and validation of search filters to identify articles on family medicine in online medical databases

David H.J. Pols, Wichor M. Bramer, Patrick J.E. Bindels, Floris A. van de Laar, Arthur M. Bohnen

Ann Fam Med. 2015 Jul-Aug;13(4):364-6. (published as 'Research Brief')

Abstract

Purpose: Physicians and researchers in the field of family medicine often need to find relevant articles in online medical databases. Because a search filter may help improve the efficiency of such searches, we aimed to develop and validate search filters to identify studies in the field of family medicine/general practice.
Method: To develop a search filter for family medicine, a precise definition was obtained which allows to classify articles as 'relevant' or 'irrelevant' to family medicine. This definition allowed to create a reference standard set of articles. Using specialized software, filter candidate terms and phrases were derived from this reference standard. Using these candidate terms and phrases, an optimal sensitive and an optimal specific filter were created. Finally, two filters were validated on two external validation sets.
Results: The sensitive filter has a sensitivity of 96.8% with an adequate specificity of 74.9%. The specific filter has a specificity of 97.4% with an adequate sensitivity of 90.3%
Conclusions: Two well-validated search filters were developed for family medicine with good sensitivity and specificity. Both filters can be applied in daily practice by family physicians and researchers. The quality of these filters is good when compared with other search filters applied in different scientific fields.

Background

Although many physicians use online medical databases to obtain biomedical information for clinical practice (1-3), the enormous volume and diversity of the available literature makes this a challenging process. Lack of time and skills, as well as a clear preference for asking an expert colleague or consulting a print source, are considered as barriers to the use of online databases (4, 5). A specific search filter might enhance the retrieval of relevant articles at the point of care by the physician. On the other hand, researchers in the field of family medicine completing a systematic review, will need a 'sensitive' search tool in order not to miss relevant articles. Electronic search filters, both sensitive and specific, can be used to improve the overall efficiency of a literature search. Search filters are strings of keywords and/or text words connected with Boolean operators (e.g. AND, OR, NOT). These topic-specific keywords can be found in the title or abstract of an article, or in the subject headings assigned to it. However, the indexing of these articles with subject headings is often inconsistent. The area of family medicine is particularly broad and difficult to define, mainly due to the different terminologies used. For example, the terms 'family medicine', 'general practice' and 'primary care' (amongst others), can be used to describe basically the same field.

Therefore, there is a need for a validated 'family medicine' search filter to support both family physicians and researchers. These filters should apply to the most frequently-used databases, e.g. PubMed, Ovid Medline, Embase and Cochrane.

To our knowledge, four filters have been developed for family medicine. Although the PHC Search Filter (6) can be considered specific, it was not designed to be comprehensive regarding what it retrieves. Jelercic et al. (7), Glanville et al. (8) and Gill et al. (9) also created search filters, but they also lack good sensitivity.

The present study was conducted to develop and validate objective search filters, applicable in frequently-used databases, to identify studies that are conducted in, or apply to, or refer to family medicine and general practice settings.

Method

Definition of family medicine

To develop an efficient and objective search filter, a clear definition of relevance to family medicine/general practice (FM/GP) is needed. WONCA Europe provides a consensus statement in which they define the discipline of FM/GP (10). Based on a short questionnaire that was sent to colleagues worldwide using the e-mail list of

the Cochrane Primary Healthcare Field, we learned that this definition was shared by many. However, the respondents indicated that two additional aspects should be taken into account. First, an inpatient hospital setting should explicitly be excluded. Second, one should be aware of the difference between 'primary care' and 'FM/GP'. Primary care was often regarded by the respondents as an umbrella term, that includes FM/GP, but could also include (amongst others) midwives, psychologists and physiotherapists. Based on an analysis of the submitted answers to the questionnaire, the WONCA definition was shortened as followed:
General practice/family medicine is the frontline of health care. It is a place where a patient can go without referral. This specifically trained physician can be consulted for acute and chronic health-related matters. Family medicine is considered to be out-of-hospital (together with the emergency department) care.
Relevance to general practice referred to any research article that explicitly indicated it was completed in a FM/GP setting as defined by WONCA, excluding inpatient hospital care and focusing specifically on FM/GP. Research articles that have 'FM/GP' as their research domain were also considered relevant (e.g. research on the efficiency of GPs).

Development of reference standard

Using Scopus, a list of 160 journals (in order of relevance for family medicine) was compiled. Five journals with a high rating (top 20) and five journals with a low rating

Table 1. Journal titles randomly selected from Scopus.

Rank in Scopus	Journal title	Hits on FM/GP* in Scopus	Hits in 2009 in PubMed	With an abstract	Included in the reference standard
2	British Journal of Family Medicine	5309	246	97 (39%)	63
3	Journal of Family Practice	3712	170	78 (46%)	44
5	American Family Physician	3404	260	104 (40%)	73
10	Canadian Family Physician	2669	264	89 (34%)	58
12	Family Practice	2288	119	112 (94%)	77
108	Age and Ageing	391	213	117 (55%)	79
121	Journal of Clinical Psychiatry	371	373	272 (73%)	188
128	Palliative Medicine	363	129	109 (84%)	73
144	Emergency Medicine Journal	305	367	217 (59%)	146
148	Intensive Care Medicine	280	415	303 (73%)	199
Total			**2556**	**1498**	**1000**

* FM: family medicine; GP: general practice

were randomly taken from this list (Table 1). From the obtained list of journals, 1000 articles published in the randomly selected year 2009, with abstracts and MeSH terms, were randomly selected. These articles were imported in EndNote X5 and anonymized, showing only the titles, abstracts and keywords to the reviewers. Two independent reviewers (DP and FvdL) classified the articles as being relevant or irrelevant to family medicine using the shortened definition based on the WONCA definition. If the anonymized information was not sufficient for a classification, all bibliographic data or even the full text was provided. Articles that refer to family medicine were tagged 'positives'. From this reference standard, two random sets were derived: a term identification set containing 1/3 of the reference standard and a development set containing 2/3 of the reference standard.

Generating a list of potentially useful terms

Using specialized software (PubReMiner (11) and AntConc (12)) candidate filter terms and phrases were derived in the term identification set from the bibliographic information of positive articles based on frequency of occurrence. Each retrieved term (MeSH term, text word or text phrase) was subsequently combined with various PubMed field codes ([mh]; [mh:noexp]; [mj]; [mj:noexp]; [sh]; [all fields]; [ad]; [tw]; [tiab]; [ti]). We included candidate filter terms for further analysis if that term retrieved at least 5% of the positive articles. Furthermore, the ratio between the percentage of positive articles containing the term and the percentage of negative articles containing the term had to be ≥ 1, and this ratio had to be significant (Chi-square test: $p<0.05$).

Creating and validating a sensitive and specific filter

With a list of candidate terms and phrases retrieved during the process described above, optimal search filters were created in the development set. The sensitive filter was created by sorting the search terms by accuracy. One by one, the items were meticulously added to the filter, whilst monitoring its performance. When an added term did not contribute to the overall accuracy of the filter, the item was excluded. The specific filter, with a target specificity of at least 95%, was created by discarding all search terms that had a specificity of ≤ 95%. Search terms that scored a specificity of 100% formed the basis of the filter. The remaining search terms were then sorted by accuracy, and were added one by one to the existing filter. When an added term did not contribute to the overall accuracy of the filter, the item was excluded.
The obtained filters were then validated in different validation sets (see below), calculating sensitivity and specificity. Finally, all the false negatives from different validation sets, missed by the sensitive filter, were manually screened by two

independent reviewers (DP and AB) to identify unique extra terms that could be added to the sensitive filter in order to improve its performance. These terms were then tested on both development and validation sets and included if they improved the overall accuracy of the sensitive filter.

Development of validation sets

In addition to the reference standard, two external validation sets were created. The first was created during the screening process of a family medicine relevant systematic review on atopic disorders in children (review standard) (13). The search for this review was not limited to family medicine, but all the references found for this review were also scored by two independent reviewers (DP and E. van Alphen) to classify articles as being relevant or irrelevant to family medicine. Relevance to general practice referred to any research article that explicitly indicated it was completed in a FM/GP setting as defined for the reference standard.
The second validation set was created by sending an e-mail to the list of the Cochrane Primary Healthcare Field (questionnaire standard). In this e-mail the participant was asked to send a reference of an article that they considered to be relevant for 'primary care', in particular for family medicine. These 500 references are considered to be positives. The negatives were collected from a random sample of articles from PubMed that were manually reviewed by two independent reviewers (FvdL and D. Al Rashad), creating 1,000 negatives.

Results

Creating and validating the filters

A total of 126 terms and phrases were considered as candidate filter terms. The original sensitive filter that was constructed missed a total of 35 'positives' in both the reference set and in the two validation sets. Manual evaluation of these 35 false negative references led to our decision to add three more terms to the sensitive filter to increase its performance, i.e. 'GP' 'GPs' and 'general pract*' were added; this substantially improved the filter.
Table 2 shows the strings of the sensitive and specific filters that were constructed using this methodology, including the translation for use in different search engines. Table 3 presents the results of a comparison between the performance of our filters and that of other published search filters (6-9). In the validation process the sensitive filter had an overall sensitivity of 96.8% (range 95.4-100%), with an adequate overall specificity of 74.9% (range 69.2-89.5%). For the specific filter

Table 2. The filters translated for different interfaces.

	Pubmed	**Ovid (Medline/ Embase)**	**Embase.com**	**Cochrane**
Sensitive filter	("family"[all fields] OR physician*[all fields] OR practice*[tw] OR "primary care"[all fields] OR "Primary Health Care"[mh] OR primary[tw] OR general pract*[tiab] OR gp[tiab] OR gps[tiab])	(family.af. OR physician$. af. OR practice$.mp. OR primary care.af. OR exp Primary Health Care/ OR primary.mp. OR general pract$.af. OR gp.tw. OR gps.tw.)	(family OR physician* OR practice*:de,it,lnk,ab,ti OR 'primary care' OR 'Primary Health Care'/exp OR primary:de,it,lnk,ab,ti OR (general NEXT/1 pract*) OR gp:ab,ti OR gps:ab,ti)	("family" OR physician* OR practice*:ti,ab,kw,pt OR "primary care" OR [mh "Primary Health Care"] OR "primary":ti,ab,kw,pt OR general pract*:ab,ti OR "gp":ab,ti OR "gps":ab,ti)
Specific filter	("Primary Health Care"[mh] OR "primary care"[all fields] OR "Physicians, Family"[mh] OR general pract*[all fields] OR "family"[ad] OR family pract*[all fields] OR family physician*[tw])	(exp Primary Health Care/ OR primary care. af. OR exp Physicians, Family/ OR general pract$.af. OR family.in. OR family pract$.af. OR family physician$.mp.)	('Primary Health Care'/ exp OR 'primary care' OR (general NEXT/1 pract*) OR family:ad OR (family NEXT/1 pract*) OR (family NEXT/1 physician*): de,it,Ink,ab,ti)	([mh "Primary Health Care"] OR "primary care" OR [mh "Physicians, Family"] OR general pract* OR family pract* OR family physician*:ti,ab,kw,pt)

the overall specificity was 97.4% (range 94.8-99.3%), with an adequate overall sensitivity of 90.3% (83.9-96.0%). Both the sensitive and the specific filters perform better compared to other recently published filters on the same topic (6-9). In table 4 the performance of our filters is compared to a combination of relevant Mesh terms (General Practice[Mesh] OR General Practitioners [Mesh] OR Physicians, Family [Mesh] OR physicians, primary care [mh]), i.e. a strategy used by many physicians in daily practice. Furthermore, the filter was tested against five search strategies used for general practice relevant Cochrane Reviews (14-18).

Discussion

Two well-validated search filters were created for family medicine, both with good sensitivity and specificity.

Our specific filter was developed to help family physicians find answers to clinical questions at the point of care when time is limited. The specific filter provides the physician with references that are relevant, but with a small risk of missing articles. If an answer to the question is not found using the specific filter, use of the sensitive filter could be the next step.

Our sensitive filter can also be used by researchers conducting a systematic review. The sensitive filter provides considerable efficiency. For example, we constructed a search string for a systematic review on atopic disorders in children through which 3,972 publications were found. If our sensitive filter had been applied, the number

Table 3. Performance of our search filters compared with that of other published filters

Used standard		Sensitive filter	Specific filter	PHCRIS (6)	Jeleric (7)	Glanville (8)	Gill (9) (high sens)	Gill (9) (balanced)	Gill (9) (high spec)
Review	**Sensitivity**	100%	90.7%	81.4%	46.5%	95.3%	95.3%	93.0%	88.4%
	Specificity	69.2%	97.9%	99.0%	94.4%	77.0%	61.0%	99.1%	99.5%
Reference	**Sensitivity**	95.4%	83.9%	65.9%	78.0%	84.3%	85.9%	68.9%	57.0%
	Specificity	69.5%	94.8%	96.3%	89.4%	89.4%	47.6%	96.7%	98.6%
Questionnaire	**Sensitivity**	97.4%	96.0%	81.4%	87.4%	97.0%	96.2%	92.2%	78.0%
	Specificity	89.5%	99.3%	99.0%	97.8%	96.1%	84.9%	99.4%	99.8%
Overall	**Sensitivity**	96.8%	90.3%	80.8%	81.9%	92.3%	91.9%	83.7%	70.9%
	Specificity	74.9%	97.4%	98.3%	94.3%	83.6%	65.0%	98.7%	99.4%

Table 4 Performance of our search filters compared with that of other search strategies

Used standard		Sensitive filter	Specific filter	Relevant Mesh	Cochrane 1 (14)	Cochrane 2 (15)	Cochrane 3 (16)	Cochrane 4 (17)	Cochrane 5 (18)
Review	**Sensitivity**	100%	90.7%	44.2%	91.1%	88.9%	88.9%	88.9%	71.1%
	Specificity	69.2%	97.9%	99.8%	99.0%	95.7%	93.1%	98.5%	96.4%
Reference	**Sensitivity**	95.4%	83.9%	56.7%	68.9%	70.2%	71.1%	68.5%	66.6%
	Specificity	69.5%	94.8%	99.9%	98.0%	88.4%	90.8%	96.4%	70.9%
Questionnaire	**Sensitivity**	97.4%	96.0%	75.8%	92.4%	92.8%	93.0%	92.0%	86.6%
	Specificity	89.5%	99.3%	99.7%	99.4%	97.5%	97.8%	99.3%	95.6%
Overall	**Sensitivity**	96.8%	90.3%	67.1%	84.1%	84.0%	84.3%	83.1%	74.8%
	Specificity	74.9%	97.4%	99.7%	98.8%	93.9%	93.9%	98.1%	87.6%

of relevant articles could have been limited to 1,478. In this example, no relevant articles were missed. Comparing our sensitive filter to the 'common practice' of search strategies used when conducting, for example, Cochrane systematic reviews, all tested literature searches showed a lack of good sensitivity (see online supplementary materials). Thus, it can be assumed that relevant references were missed in these reviews which might have been found when applying our sensitive search filter.

The present filters do not use the Boolean operators AND or NOT, but combined single search terms and phrases in an OR relationship. However, in our methodology, 'phrases' were already separately identified as combination of words in an AND matter that could potentially discriminate between FM/GP-relevant or not. For example 'primary health care' was identified in this way. These three words are combined in an AND matter, but the quotes also demand it to be in this specific order. Using an objective method, the developed filters did not always end up with phrases that one would expect, like 'family physician'. However, our objective method suggested the single words 'family' and 'physician' to be more distinctive. Finally, using NOT would imply a substantial risk of excluding relevant articles and was therefore rejected.

There are two important arguments for manually improving the sensitive filter. In order for this methodology to create a completely objective filter without manual improvement, it was estimated that about 30,000 articles had to be scored. Manually evaluating the false negatives overcomes the use of a relatively small 'reference standard'. Furthermore, in order for AntConc to find phrases, the 126 candidate filter terms were used. Words like 'general' and 'practice' did not meet the requirements for inclusion in the list of candidate filter terms, because the words themselves are not specific enough.

Translating the search strategies developed for PubMed to the syntax of the other databases (Ovid, Embase and Cochrane), carries a small risk of losing some sensitivity and specificity. Ideally, one would use the candidate filter terms and start constructing the search filter using the different interfaces. Unfortunately, the other databases did not have an 'application programming interface' (a set of routines, protocols, and tools for building software applications) that allowed communication with the software programs that were used for the development of these search filters. Instead the filters were directly translated into the syntax of the other databases, without optimization for that specific database.

We noticed that, in many cases, the title and abstract did not disclose sufficient information to determine whether (or not) an article was relevant for family medicine. In many cases the setting and/or relevance to family medicine could only be determined by scrutinizing the full text; this omission will influence both the

sensitivity and specificity of search filters. We emphasize that mentioning the setting in the title or abstract will help to find all relevant literature available for family medicine.

Conclusions

Two useful filters were created for a search on articles related to family medicine. The sensitive filter has a sensitivity of 96.8% with an adequate specificity of 74.9%. The specific filter has a specificity of 97.4% with an adequate sensitivity of 90.3%.

Acknowledgments

The authors thank Heleen Moed, Dahri Al Rashad, Elvira van Alphen and Youssou Gueye for their assistance in the preparation of this study.

References

1. Chiu YW, Weng YH, Lo HL, Ting HW, Hsu CC, Shih YH, et al. Physicians' characteristics in the usage of online database: A representative nationwide survey of regional hospitals in Taiwan. Inform Health Soc Care. 2009;34(3):127-35.
2. Shariff SZ, Bejaimal SAD, Sontrop JM, Iansavichus AV, Weir MA, Haynes RB, et al. Searching for medical information online: A survey of Canadian nephrologists. J Nephrol. 2011;24(6):723-32.
3. Weng YH, Kuo KN, Yang CY, Lo HL, Shih YH, Chiu YW. Information-searching behaviors of main and allied health professionals: A nationwide survey in Taiwan. J Eval Clin Pract. 2013; 19(5):902-8.
4. Younger P. Internet-based information-seeking behaviour amongst doctors and nurses: a short review of the literature. Health Info Libr J. 2010;27(1):2-10.
5. Davies K, Harrison J. The information-seeking behaviour of doctors: a review of the evidence. Health Info Libr J. 2007;24(2):78-94.
6. Brown L, Carne A, Bywood P, McIntyre E, Damarell R, Lawrence M, et al., editors. Facilitating access to evidence: Primary Health Care Search Filter2014.
7. Jelercic S, Lingard H, Spiegel W, Pichlhofer O, Maier M. Assessment of publication output in the field of general practice and family medicine and by general practitioners and general practice institutions. Fam Pract. 2010;27(5):582-9.
8. Glanville J, Kendrick T, McNally R, Campbell J, Hobbs FD. Research output on primary care in Australia, Canada, Germany, the Netherlands, the United Kingdom, and the United States: bibliometric analysis. BMJ. 2011;342:d1028.
9. Gill PJ, Roberts NW, Wang KY, Heneghan C. Development of a search filter for identifying studies completed in primary care. Fam Pract. 2014;31(6):739-45.
10. The European definition of general practice / family medicine. 2011 [cited 2014 3-2-2014]; Available from: http://www.woncaeurope.org/sites/default/files/documents/Definition%20 3rd%20ed%202011%20with%20revised%20wonca%20tree.pdf.
11. Koster J. PubMed PubReminer. Amsterdam: AMC; 2008.
12. Anthony L. AntConc. 3.2.4 ed. Tokyo, Japan: Waseda University; 2011.
13. Pols DH, Wartna JB, Moed H, van Alphen EI, Bohnen AM, Bindels PJ. Atopic dermatitis, asthma and allergic rhinitis in general practice and the open population: a systematic review. Scand J Prim Health Care. 2016;34(2):143-50. Epub 2016/03/25.
14. Rosendal M, Blankenstein AH, Morriss R, Fink P, Sharpe M, Burton C. Enhanced care by generalists for functional somatic symptoms and disorders in primary care. Cochrane Database Syst Rev. 2013;10:CD008142. Epub 2013/10/22.
15. Smith SM, Soubhi H, Fortin M, Hudon C, O'Dowd T. Interventions for improving outcomes in patients with multimorbidity in primary care and community settings. Cochrane Database Syst Rev. 2012;4:CD006560. Epub 2012/04/20.
16. Scott A, Sivey P, Ait Ouakrim D, Willenberg L, Naccarella L, Furler J, et al. The effect of financial incentives on the quality of health care provided by primary care physicians. Cochrane Database Syst Rev. 2011(9):CD008451. Epub 2011/09/09.
17. Hoedeman R, Blankenstein AH, van der Feltz-Cornelis CM, Krol B, Stewart R, Groothoff JW. Consultation letters for medically unexplained physical symptoms in primary care. Cochrane Database Syst Rev. 2010(12):CD006524. Epub 2010/12/15.

18. Kaner EF, Beyer F, Dickinson HO, Pienaar E, Campbell F, Schlesinger C, et al. Effectiveness of brief alcohol interventions in primary care populations. Cochrane Database Syst Rev. 2007(2):CD004148. Epub 2007/04/20.