BMJ

RESEARCH

Filtering Medline for a clinical discipline: diagnostic test assessment framework

Amit X Garg, associate professor,[1,2,3] Arthur V Iansavichus, information specialist ,[1] Nancy L Wilczynski, assistant professor ,[3] Monika Kastner, PhD student ,[4] Leslie A Baier, research assistant ,[3] Salimah Z Shariff, PhD student ,[1] Faisal Rehman, assistant professor ,[1] Matthew Weir, research fellow ,[1] K Ann McKibbon, associate professor ,[3] R Brian Haynes, professor [3]

[1]Division of Nephrology, University of Western Ontario, London, ON, Canada N6A 5C1

[2]Department of Epidemiology and Biostatistics, University of Western Ontario

[3]Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, ON, Canada L8N 3Z5

[4]Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada M5T 3M6

Correspondence to: A Garg, London Kidney Clinical Research Unit, Room ELL-101, Westminster, London Health Sciences Centre, 800 Commissioners Road East, London, ON, Canada N6A 4G5 amit.garg@lhsc.on.ca

Cite this as: *BMJ* 2009;339:b3435
doi:10.1136/bmj.b3435

ABSTRACT

Objective To develop and test a Medline filter that allows clinicians to search for articles within a clinical discipline, rather than searching the entire Medline database.

Design Diagnostic test assessment framework with development and validation phases.

Setting Sample of 4657 articles published in 2006 from 40 journals.

Reviews Each article was manually reviewed, and 19.8% contained information relevant to the discipline of nephrology. The performance of 1 155 087 unique renal filters was compared with the manual review.

Main outcome measures Sensitivity, specificity, precision, and accuracy of each filter.

Results The best renal filters combined two to 14 terms or phrases and included the terms "kidney" with multiple endings (that is, truncation), "renal replacement therapy", "renal dialysis", "kidney function tests", "renal", "nephr" truncated, "glomerul" truncated, and "proteinuria". These filters achieved peak sensitivities of 97.8% and specificities of 98.5%. Performance of filters remained excellent in the validation phase.

Conclusions Medline can be filtered for the discipline of nephrology in a reliable manner. Storing these high performance renal filters in PubMed could help clinicians with their everyday searching. Filters can also be developed for other clinical disciplines by using similar methods.

INTRODUCTION

Clinicians search bibliographic databases for information to guide the care of their patients.[1 2] Medline is the most popular of the databases. About 800 million PubMed searches are now done each year; in a survey in 2002, 15% of all searches were done by clinicians (personal communication, National Library of Medicine staff).[3] As of February 2009, this multipurpose electronic database contained information on 18 million articles from 5363 different journals; 12 500 new articles are added each week.[4 5]

However, when clinicians type searches into PubMed, they often do not retrieve all the key articles relevant to the questions they are trying to answer. One way to improve this would be to filter Medline to a discipline of interest when searching. The use of filters is akin to screening for disease in high risk populations. By filtering the database to do the search with a discipline specific set of articles, the likelihood of retrieving relevant information with the remaining search terms is increased.

To search for information on the effectiveness of hepatitis B vaccination in chronic kidney disease, for example, one could type a phrase as shown in figure 1. Alternatively, one could choose to use a renal filter and simply type in the phrase "hepatitis B vaccination" (fig 2). One would then no longer be searching the entire Medline database but, rather, searching within a set of articles relevant to a discipline. Selecting a discipline filter removes the need to type in terms for that discipline. The filter would use a pre-programmed combination of medical subject headings (MeSH), explosions, subheadings, and text words of key concepts, words, and phrases to embody a discipline of interest, in this case nephrology.[6 7]

Members of our team previously developed and tested Medline filters to optimise the retrieval of studies and systematic reviews of treatment, diagnosis, prognosis, aetiology, and clinical prediction guides.[8 9] The filters to retrieve primary studies are part of the PubMed interface in the clinical queries section, where users search a Medline database filtered for articles of high methodological merit.[10] The clinical queries filters are independent of any particular clinical discipline, such as cardiology or nephrology.

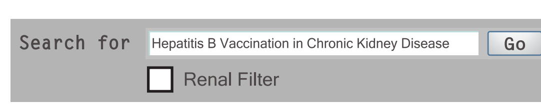

Fig 1 | **Searching without using filter**

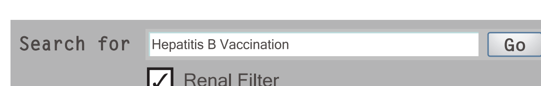

Fig 2 | **Searching with use of filter**

Table 1 | **Formulas for calculating sensitivity, specificity, precision, and accuracy of each filter to identify articles with renal information**

Filter (consisting of single or combined terms)	Manual review of each article	
	Articles with renal information	Articles without renal information
Article identified	a	b
Article not identified	c	d

Sensitivity=a/(a+c): proportion of articles with renal information identified (also called recall in information retrieval studies).
Specificity=d/(b+d): proportion of articles without renal information not identified.
Precision=a/(a+b) (also referred to as positive predictive value in diagnostic test terminology): proportion of articles identified with renal information.
Accuracy=(a+d)/(a+b+c+d): proportion of all articles dealt with correctly by filter.

In this study, we aimed to develop new high performance filters for a clinical discipline in medicine. We chose the area of renal medicine, as clinical information in this field is published across hundreds of multidisciplinary journals and is difficult to track down.[11]

METHODS

Study overview

We used a diagnostic test assessment framework with development and validation phases (fig 3, table 1). We divided a sample of articles from all available articles in Medline into two sets: a development dataset and a validation dataset. We produced a "reference standard" by manually reviewing a sample of articles to determine whether they contained any type of renal information. We then compared the retrieval performance of various filters made up of individual search terms and combinations of terms with the reference standard of manual review. We treated each filter as a "diagnostic test" for the identification (retrieval) of renal articles. For each filter, we constructed a two by two contingency table and quantified agreement (measures outlined in table 1). We then examined in the validation set of articles those filters that performed well in the development set of articles.

Sample of articles

For efficient manual review of full text articles for relevance, we first sampled a set of journals and then sampled a set of articles within those journals. We had previously compiled a list of journals that published at least one article relevant to the care of renal patients in the period from 1961 to 2005. We ranked these 466 journals by the number of articles with renal information.[11] We selected the top 20 ranked journals, divided the remaining 446 journals into five equal groups, and randomly selected four journals from each group. We ordered these 40 journals by rank and randomly divided the list into either the development set or the validation set by using a block size of five journals and a ratio of three to two (table 2). We then manually reviewed all articles published in the first three months of 2006 for each journal and restricted our searches to these articles (fig 1). We reviewed all types of articles indexed in Medline, including original investigations, reviews, letters, and editorials. We initially selected two additional journals through our sampling process,[12 13] but we did not consider them further because they were not available to us in electronic format.

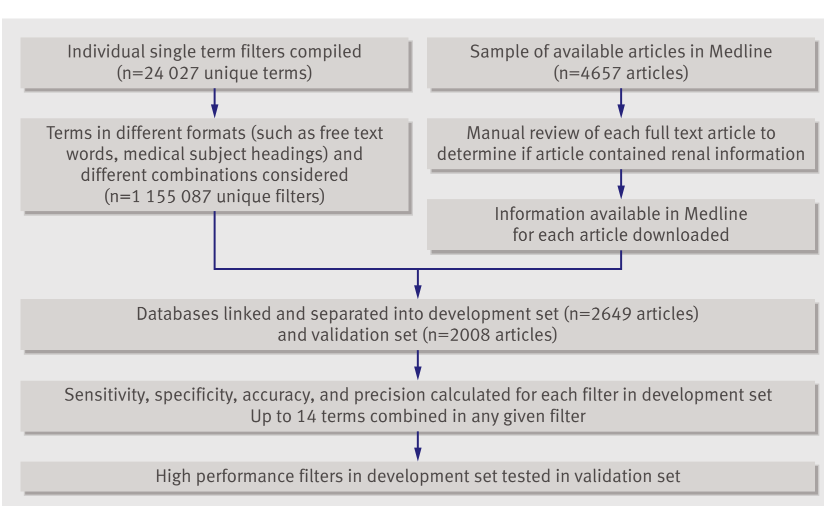

Fig 3 | **Data collection and filter development**

Review of each article

We previously developed a standardised checklist to determine whether an article contained renal information (developed by a team of nephrologists, see web appendix). We derived this checklist by reviewing nephrology textbooks and the MeSH thesaurus. We used this checklist to determine whether the full text of each article was relevant to nephrology (four reviewers: AVI, LAB, MK, and AXG). Using five test sets of 298 articles, all reviewers were calibrated against a nephrologist (AXG) in their application of checklist criteria (agreement beyond chance, κ=0.98).[14]

Filters

We compiled renal terms used in the filters from the following sources: US National Library of Medicine (NLM) medical subject heading (MeSH) thesaurus using Medline MeSH browser,[15] Medline permuted index,[16] Emtree thesaurus,[17] SNOMED clinical terms, nephrology textbooks,[18-20] clinical practice guidelines,[21 22] website glossaries,[23-31] 195 renal systematic reviews,[11] 21 clinicians from eight different countries, and seven librarians from three different countries. Any term considered potentially useful by anyone involved in this process was added to the list. Examples of terms used in the filters included kidney, renal, creatinine, nephropathy, uremia, and dialysis. We considered the terms both as MeSH terms and as text words. We considered MeSH terms with and without major focus (major focus refers to records in which an index term has been tagged as the major topic of the article) and as 42 possible subheadings, and with and without explosion capability (for example, exploding the MeSH "renal replacement therapy" means the following MeSH terms are included in the search: renal dialysis, hemodialysis, peritoneal dialysis, hemofiltration, hemodiafiltration, and kidney transplantation). We considered free text words as full and truncated terms (inclusion of multiple endings achieved though use of the $ symbol—for example, nephro$), using both American and British English spelling. Terms could appear anywhere in a citation (title, abstract, subject headings, and so on) but not in the journal name only. We automated the process of combining and testing the filters by using a computer implemented algorithm. We combined single term filters with a sensitivity greater than 10% and a specificity greater than 10% into multiple term filters, as well as two term filters with a sensitivity above 75% and a specificity above 50%. We used Boolean operators "OR," "AND," and "NOT" to combine terms.

Table 2 | Division of 40 journals into development and validation sets

Rank*	Journal	Total contributed articles (n=4657)	Articles with renal information (%)
Development set			
1	Nephrology Dialysis Transplantation	168	94.6
2	Transplantation Proceedings	171	59.7
3	Journal of the American Society of Nephrology	102	87.3
4	American Journal of Kidney Diseases	85	90.6
5	Pediatric Nephrology	83	67.5
6	American Journal of Transplantation	96	44.8
7	New England Journal of Medicine	376	3.7
8	Diabetes Care	184	5.4
9	American Journal of Medicine	107	7.5
10	Journal of Hypertension	89	9
11	Radiology	136	5.1
12	Journal of Vascular Surgery	73	6.8
13	Lancet	386	0.8
14	Archives of Disease in Childhood	87	3.4
15	Journal of Trauma-Injury Infection & Critical Care	120	1.7
16	Diabetic Medicine	60	1.7
17	Surgical Endoscopy	101	1
18	Journal of the Association of Physicians of India	50	2
19	Calcified Tissue International	22	4.5
20	Journal of Human Genetics	41	2.4
21	Journal of Infection	39	0
22	Journal of Viral Hepatitis	29	0
23	American Journal of Clinical Oncology	22	0
24	Netherlands Journal of Medicine	22	0
Total		2649	22.3
Validation set			
1	Kidney International	187	79.7
2	Transplantation	168	37.5
3	Peritoneal Dialysis International	49	100
4	Clinical Transplantation	46	50
5	Journal of Urology	219	10.1
6	Annals of Internal Medicine	130	6.9
7	BMJ	477	1.3
8	Annals of Thoracic Surgery	289	1.7
9	Investigative Radiology	48	6.3
10	Journal of Pediatrics	96	2.1
11	Clinical Pharmacology and Therapeutics	30	3.3
12	American Journal of Clinical Pathology	58	1.7
13	Bone Marrow Transplantation	101	0
14	Family Practice	57	0
15	East African Medical Journal	31	0
16	Diabetes/Metabolism Research Reviews	22	0
Total		2008	16.6

*Ranked by number of articles with renal information.

Statistical analysis

We calculated the sensitivity, specificity, precision, and accuracy of each filter as described in table 1. We developed and tested filters by using Ovid Medline syntax. Compared with Ovid syntax, translations provided for the PubMed interface had an accuracy of more than 99.5%.

Proof of concept searches

To examine the utility of filters, we asked five clinicians independent of the research team to each type in a PubMed search for a focused clinical question. We selected these focused clinical questions because in each case a recent systematic review had used a comprehensive method to compile relevant primary studies.[32-36] We randomly selected the clinicians from a list of nephrologists practising in Canada and asked them to complete an online survey on their medical information gathering practices. The sample included four men and one woman, the average age was 45 (range 37-52) years, the average length of practice was 11 (5-20) years, and the average number of Medline searches done was 5 (1-15) a month. Two clinicians were practising in a centre with a nephrology training programme.

We provided the clinicians with as much time as they needed to complete the survey. We asked each clinician to search for articles on one of the following: the renal effects of statins, the benefits of fenoldopam in acute kidney injury, the benefits of tacrolimus compared with ciclosporin in kidney transplantation, the efficacy of low dose dopamine in acute kidney injury, and the benefits of intradermal compared with intramuscular hepatitis B vaccination in chronic kidney disease. We restricted each search to the search dates provided in the methods of each of the identified systematic reviews and the records indexed in Medline. In each case, we determined how many relevant articles were identified by the clinician's search and how many relevant articles were identified when the physician's search was combined with the best performing filters developed as part of this study.

RESULTS

Sample of articles—We used 4657 articles: 2649 articles from 24 journals in the development set and 2008 articles from another 16 journals in the validation set (fig 3, table 2). We manually reviewed each article, and 19.8% contained renal information (table 2). We compiled a total of 24 027 unique terms, which formed 1 155 087 unique filters (fig 3).

Single term filters—We tested the filters in the development set of articles. The best single term filters were text word "kidney" and exploded major MeSH "kidney diseases", which achieved sensitivities of 78.7% and 57.5% and specificities of 97.2% and 98.6%

Table 3 | Best single term filters for high sensitivity (keeping specificity ≥50%), high specificity (keeping sensitivity ≥50%), and optimal balance of sensitivity and specificity, and performance of some other single term filters from 24 027 considered. Values are percentages (95% confidence intervals)

Filters						
Ovid†	PubMed translation‡	Journal set	Sensitivity	Specificity	Precision	Accuracy
Best sensitivity						
kidney$.mp	kidney*[tw]	Development	78.7 (75.4 to 82.0)	97.2 (96.5 to 97.9)	88.9 (86.2 to 91.6)	93.0 (92.1 to 94.0)
		Validation	79.6 (75.25 to 83.9)	95.6 (94.7 to 96.6)	78.4 (74.0 to 82.8)	93.0 (91.9 to 94.1)
Best specificity						
exp *kidney diseases/	"kidney diseases"[majr]	Development	57.5 (53.5 to 61.5)	98.6 (98.1 to 99.1)	92.4 (89.7 to 95.1)	89.5 (88.3 to 90.6)
		Validation	41.4 (36.1 to 46.7)	97.6 (96.8 to 98.3)	77.1 (70.9 to 83.3)	88.2 (86.8 to 89.7)
Best optimisation of sensitivity and specificity						
kidney$.mp	kidney*[tw]	Development	78.7 (75.4 to 82.0)	97.2 (96.5 to 97.9)	88.9 (86.2 to 91.6)	93.0 (92.1 to 94.0)
		Validation	79.6 (75.25 to 83.9)	95.6 (94.7 to 96.6)	78.4 (74.0 to 82.8)	93.0 (91.9 to 94.1)
Other single item filters						
*kidney transplantation/	"kidney transplantation" [majr:noexp]	Development	28.6 (24.9 to 32.2)	99.7 (99.5 to 99.9)	96.6 (93.9 to 99.3)	83.8 (82.4 to 85.2)
		Validation	27.0 (22.3 to 31.8)	99.9 (99.7 to 100)	97.8 (94.8 to 100)	87.8 (86.4 to 89.2)
glomerul$.mp	glomerul*[tw]	Development	20.8 (17.5 to 24.1)	99.3 (98.9 to 99.6)	89.1 (83.9 to 94.3)	81.8 (80.3 to 83.2)
		Validation	24.0 (19.4 to 28.6)	99.5 (99.1 to 99.8)	89.9 (83.6 to 96.2)	86.9 (85.5 to 88.4)
glomerular filtration rate/	"glomerular filtration rate"[mh:noexp]	Development	6.3 (4.3 to 8.2)	99.8 (99.7 to 100)	92.5 (84.3 to 100)	78.9 (77.4 to 80.5)
		Validation	7.2 (4.4 to 9.9)	100 (100 to 100)	100 (100 to 100)	84.6 (83.0 to 86.2)
renal.ti.	renal[ti]	Development	34.2 (30.4 to 38.0)	99.5 (99.2 to 99.8)	94.8 (91.9 to 97.8)	84.9 (83.5 to 86.3)
		Validation	35.1 (30.0 to 40.3)	98.7 (98.1 to 99.2)	84.2 (78.1 to 90.2)	88.1 (86.7 to 89.6)
*proteinuria/	"proteinuria"[majr:noexp]	Development	2.9 (1.5 to 4.2)	99.9 (99.8 to 100)	89.5 (75.7 to 100)	78.3 (76.7 to 79.8)
		Validation	1.8 (0.4 to 3.2)	100 (100 to 100)	100 (100 to 100)	83.7 (82.1 to 85.3)
exp renal replacement therapy/	"renal replacement therapy"[mh]	Development	48.1 (44.0 to 52.1)	99.5 (99.2 to 99.8)	96.6 (94.5 to 98.7)	88.0 (86.8 to 89.3)
		Validation	50.5 (45.1 to 55.8)	99.5 (99.1 to 99.8)	94.9 (91.7 to 98.2)	91.3 (90.1 to 92.6)
dialy$.mp.	dialy*[tw]	Development	26.6 (23.0 to 30.1)	99.9 (99.7 to 100)	98.1 (96.0 to 100)	83.5 (82.1 to 84.9)
		Validation	27.0 (22.3 to 31.8)	99.9 (99.7 to 100)	97.8 (94.8 to 100)	87.8 (86.4 to 89.2)

†Ovid fields: exp=exploded search term that automatically includes narrower indexing terms (for example, with exploded term kidney diseases, terms such as glomerulonephritis, renal insufficiency, uremia, and kidney failure are included in search); $=truncation character; mp=multiple posting (term appears in title, abstract, or MeSH); /=MeSH character; *=focused MeSH term; ti=term present in title.
‡PubMed fields: *=truncation character; [majr]=exploded and focused MeSH term; [majr:noexp]=not exploded and focused MeSH term; [tw]=text word present in title, abstract, or MeSH; [ti] =term present in title; [mh]=exploded MeSH term; [mh:noexp]=not exploded MeSH term.

(table 3). Table 3 also shows the performance of other terms such as "renal" and the exploded MeSH "renal replacement therapy". The retrieval performance of these filters was similar in the validation set of articles (table 3).

Multiple term filters—We tested 1 131 060 filters using a combination of two to 14 terms in the development set of articles. Top filters achieved peak sensitivities of 97.8% and specificities of 98.5% (table 4). The best filters included the terms "renal replacement therapy", "renal dialysis", "kidney function tests", "renal", "nephr" truncated, "glomerul" truncated, and "proteinuria". The performance of the best filters remained excellent in the validation set of articles (table 4).

Proof of concept searches—The retrieval of relevant studies increased when we combined the best filters with a search by a clinician (table 5). For example, in the case of searching for the renal effects of statins, the clinician's search on its own retrieved six of the 24 relevant articles. This increased to 20/24 when we combined this search with the most sensitive filter and to 16/24 when we combined the search with the most specific filter.

DISCUSSION

Previous attempts to develop Medline filters for a clinical discipline have met with limited success, and many have never been validated.[7 37-39] We succeeded in proving that Medline can be filtered for a clinical discipline in a reliable manner. Our best renal filters had a sensitivity and specificity in excess of 96%. Clinicians retrieved more clinically relevant articles when they used these filters.

Strengths and limitations

We tested more than one million renal filters, using an empirical approach to discover those with the highest performance. However, these filters help only with the renal components of any search. Limitations of the accompanying terms, such as the description of a certain treatment or diagnostic test, will continue to contribute to poor performance of searches. To develop these high performance renal filters, we sampled clinical rather than basic science journals. We also deliberately enriched the sample with primary renal journals. Although the sensitivity and specificity will not change when these filters are applied to all Medline journals, the precision will be reduced from the values shown in table 4. However, this level of precision uses a very strict definition of relevance (referenced in a systematic review), and we expect that other types of articles such as review articles and clinical practice guidelines will also be relevant to the searcher. Finally, although these

Table 4 | Top filters yielding highest sensitivity (keeping specificity >90%) and highest specificity (keeping sensitivity >90%) based on combination of up to 14 terms. Values are percentages (95% confidence intervals)

Filter: Ovid†	Filter: PubMed translation‡	Journal set	Sensitivity	Specificity	Precision	Accuracy
Highest sensitivity						
exp kidney diseases/OR exp renal replacement therapy/OR renal.mp OR kidney$.mp OR nephr$.mp. OR proteinuria.mp	"kidney diseases"[mh] OR "renal replacement therapy"[mh] OR renal[tw] OR kidney*[tw] OR (nephre*[tw] OR nephri*[tw] OR nephroc*[tw] OR nephrog*[tw] OR nephrol*[tw] OR nephron*[tw] OR nephrop*[tw] OR nephros*[tw] OR nephrot*[tw]) OR proteinuria[tw]	Development	97.8 (96.6 to 99.0)	95.0 (94.1 to 95.9)	84.9 (82.2 to 87.6)	95.6 (94.8 to 96.4)
		Validation	96.7 (94.8 to 98.6)	93.9 (92.8 to 95.1)	75.9 (71.9 to 80.0)	94.4 (93.4 to 95.4)
Highest specificity						
(exp *renal replacement therapy/OR exp *kidney diseases/OR kidney$.ti. OR nephr$.ti. OR renal.ti. OR *kidney/OR exp renal dialysis/OR exp *kidney function tests/OR *proteinuria/OR glomerul$.ti.) NOT (exp *kidney neoplasms/OR *pyelonephritis/OR exp *urinary tract infections/OR exp *nephrolithiasis/)	("renal replacement therapy"[majr] OR "kidney diseases"[majr] OR kidney*[ti] OR nephr*[ti] OR renal[ti] OR "kidney"[majr:noexp] OR "renal dialysis"[mh] OR "kidney function tests"[majr] OR "proteinuria"[majr:noexp] OR glomerul*[ti]) NOT ("kidney neoplasms"[majr] OR pyelonephritis[majr:noexp] OR "urinary tract infections"[majr] OR "nephrolithiasis"[majr])	Development	93.4 (91.4 to 95.4)	98.5 (98.0 to 99.0)	94.7 (92.9 to 96.5)	97.4 (96.7 to 98.0)
		Validation	91.3 (88.3 to 94.3)	98.9 (98.4 to 99.4)	94.4 (91.9 to 96.9)	97.7 (97.0 to 98.3)

†Ovid fields: exp=exploded MeSH term; $=truncation character; mp=multiple posting (term appears in title, abstract, or MeSH); /=MeSH character; *=focused MeSH; ti=term present in title.
‡PubMed fields: *=truncation character; [majr]=exploded and focused MeSH term; [majr:noexp]=not exploded and focused MeSH term; [tw]=text word present in title, abstract, or MeSH; [ti] =term present in title; [mh]=exploded MeSH term.

filters should improve the retrieval of relevant articles compared with unaided searches, they may return a greater number of non-relevant articles (table 5).

Of course, some articles are never indexed in Medline and can only be found through other bibliographic databases such as Embase. However, even when present in Medline, some articles may never be retrieved with the filters or otherwise because of poor indexing.[40-42] For example, the subject heading for a recent citation on diabetic nephropathy was entered as diabetic neuropathy.[43] Other articles lack accurate subject headings, key words, or a proper descriptive abstract,[44-47] and some medical concepts lack existing MeSH terms.[40] These filters may also need future updates if important changes in vocabulary occur, as happened when the concept of "chronic renal insufficiency" began to be referred to as "chronic kidney disease."[21 48]

Using these renal search filters

These best performing filters are complex, with multiple terms. Coding these renal filters into the PubMed and Ovid search engine interfaces will permit their easy use by anyone doing a search (as done with our "clinical queries," which as of March 2009 were located on the left hand menu of the PubMed screen). In the meantime, we provide these filters at http://hiru.mcmaster.ca/hiru/HIRU_Hedges_Nephrology_Filters.aspx. By selecting a simple filter option, one can query only those articles

Table 5 | Number of relevant articles retrieved with and without search filters

Clinical question*	No of relevant articles retrieved: Clinician's search alone	No of relevant articles retrieved: Clinician's search with most sensitive renal content filter	No of relevant articles retrieved: Clinician's search with most specific renal content filter	No of non-relevant articles retrieved†: Clinician's search alone	No of non-relevant articles retrieved†: Clinician's search with most sensitive renal content filter	No of non-relevant articles retrieved†: Clinician's search with most specific renal content filter
What are the effects of statins on change in kidney function and urinary protein excretion? (24 relevant articles)	6	20	16	86	48	39
What is the impact of fenoldopam on acute kidney injury, patients' mortality, and length of hospital stay in critically ill patients? (12 relevant articles)	1	11	11	7	32	21
When tacrolimus is compared directly with ciclosporin in the treatment of kidney transplant recipients, what is the evidence on transplant outcomes, toxicity, and adverse effects? (63 relevant articles)	10	60	60	18	20	15
What is the efficacy of low dose dopamine (<5 μg/kg of body weight per minute) compared with no therapy in patients with or at risk for acute renal failure? (52 relevant articles)	6	15	12	13	16	15
How does intradermal *v* intramuscular hepatitis B vaccine compare regarding response rate among chronic kidney disease patients? (11 relevant articles)	6	10	10	22	34	31

*Five clinicians were asked to type in a PubMed search to answer a focused clinical question for which relevant articles were summarised in a recent systematic review.[33-37] Each search was restricted to search dates provided in methods of each review. Search phrases as determined and typed in by clinicians were: "statins and kidney function", "fenoldopam and acute kidney injury", "'kidney transplant outcome tacrolimus cyclosporin", "low-dose dopamine and acute renal failure", and "hepatitis b vaccination in chronic kidney disease". Renal terms were removed when the clinician's search was performed with renal content filters.
†Number of non-relevant articles retrieved is expressed per relevant article, rounded up to nearest whole number.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Previous attempts to filter Medline for a clinical discipline have met with limited success

WHAT THIS STUDY ADDS

Medline can be filtered for a clinical discipline in a reliable manner

The best renal filters had sensitivity and specificity in excess of 97%

These filters can be programmed into the PubMed interface, so they are available for everyone to use

filtered for renal information. As of March 2009, our most sensitive filter reduced the Medline database from 18 million citations to about 780 000 citations, and the most specific filter reduced it to about 435 000 citations.

Future research

Ongoing development of filters will help to prevent relevant articles from being missed. The best filters should also minimise the number of non-relevant articles retrieved. Future research should quantify the impact of filters on real searches by clinicians, clinicians' knowledge, medical decision making, and even patients' outcomes.[49] Such research can also consider whether searchers' characteristics, such as expertise in searching, influence filters' utility. The impact of different types of filters in combination should be considered, including filters made for clinical disciplines, methodological characteristics, and subsets of journals. Developing filters for specific areas within a discipline may also have additional benefits, such as filters for transplantation or acute kidney injury within the discipline of nephrology. Finally, the methods described in this study can be used to develop filters for other disciplines. Whether high performance filters can be developed for other clinical disciplines, as we have done for the renal vocabulary, remains to be seen.

We thank other members of our research team: Nicholas Hobson and Chris Cotoi, who did the computer programming, and Robert Yang who helped to develop the criteria used to assess renal information.

Contributors: AXG, AVI, NLW, KAM, and RBH conceived the study. AVI compiled articles and managed data. AXG, AVI, MK, and LAB rated the articles for renal relevance. NLW and RBH supervised the computer programming. All authors had full access to data and aided the interpretation. SZS organised the clinicians' searches. AXG drafted the manuscript, and all authors revised it. AXG is the guarantor.

Funding: This study was funded by the Kidney Foundation of Canada. AXG was supported by a clinician scientist award from the Canadian Institutes of Health Research. The researchers were independent of the funders. The funders had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Competing interests: None declared.

Ethics approval: The study was approved by the regional ethics board of the University of Western Ontario. The five clinician searchers provided informed consent for study participation.

1 Coumou HC, Meijman FJ. How do primary care physicians seek answers to clinical questions? A literature review. *J Med Libr Assoc* 2006;94:55-60.
2 Cullen RJ. In search of evidence: family practitioners' use of the internet for clinical information. *J Med Libr Assoc* 2002;90:370-9.
3 National Library of Medicine. Key Medline indicators. 2008. www.nlm.nih.gov/bsd/bsd_key.html.
4 NLM Systems Data News and Update Information. 2009. www.nlm.nih.gov/bsd/revup/revup_pub.html.
5 National Library of Medicine. Number of titles currently indexed for Index Medicus and Medline on PubMed. 2009. www.nlm.nih.gov/bsd/num_titles.html.
6 Haynes RB, McKibbon KA, Wilczynski NL, Walter SD, Werre SR, Hedges Team. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005;330:1179.
7 Haafkens J, Moerman C, Schuring M, van Dijk F. Searching bibliographic databases for literature on chronic disease and work participation. *Occup Med (Lond)* 2006;56:39-45.
8 Haynes RB, McKibbon KA, Wilczynski NL, Walter SD, Werre SR. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005;330:1179.
9 Wilczynski NL, Morgan D, Haynes RB. An overview of the design and methods for retrieving high-quality studies for clinical care. *BMC Med Inform Decis Mak* 2005;5:20.
10 PubMed. Clinical queries. 2008. www.ncbi.nlm.nih.gov/entrez/query/static/clinical.shtml.
11 Garg AX, Iansavichus AV, Kastner M, Walters LA, Wilczynski N, McKibbon KA, et al. Lost in publication: half of all renal practice evidence is published in non-renal journals. *Kidney Int* 2006;70:1995-2005.
12 International Society of Nephrology. *Nephron*. Basel, New York: S Karger, 2006.
13 Feistle K. *Clinical nephrology*. München-Deisenhofen: Dustri-Verlag, 2006.
14 Holman CD. Analysis of interobserver variation on a programmable calculator. *Am J Epidemiol* 1984;120:154-60.
15 National Library of Medicine. MeSH Browser (2008 MeSH). 2008. www.nlm.nih.gov/mesh/MBrowser.html.
16 National Library of Medicine. *Permuted medical subject headings*. Bethesda, MD: US Dept of Commerce, National Technical Information Center, 2003.
17 Excerpta Medica. *EMTREE thesaurus*. Amsterdam: Excerpta Medica, 2001.
18 Daugirdas JT, Ing TS. *Handbook of dialysis*. 2nd ed. Boston: Little, Brown, 1994.
19 Danovitch GM. *Handbook of kidney transplantation*. 3rd ed. Philadelphia: Lippincott Williams & Wilkins, 2001.
20 Brenner BM, Rector FC. *Brenner and Rector's the kidney*. 5th ed. Philadelphia: Saunders, 1996.
21 National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis* 2002;39(2 suppl 1):S1-266.
22 National Kidney Foundation. NKF-KDOQI guidelines. 2008. www.kidney.org/professionals/kdoqi/guidelines.cfm.
23 University of Southern California Department of Surgery Kidney Transplant Program. Kidney transplant information and glossary of kidney terms. 2007. www.kidneytransplant.org/kidneyglossary.html.
24 Loma Linda University Medical Center. Kidney transplant glossary—LLUMC Transplantation Institute. 2007. www.llu.edu/llumc/transplant/kidneyglossary.html.
25 Kidney Foundation of Canada Ontario. Glossary of terms. 2007 www.kidney.on.ca/kidney_glossary/;jsessionid=5AF75A1565ABB3A845D4EFACF8A8D605.
26 Department of Health. National service framework for renal services: glossary of terms for part one: dialysis and transplantation and part two: chronic kidney disease, acute renal failure and end of life care. 2007. www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4077621.
27 American Kidney Fund. The kidney disease dictionary. 2007. www.kidneyfund.org/kf_dictionary.asp.
28 National Kidney and Urologic Diseases Information Clearinghouse (NKUDIC). Kidney failure glossary. 2007. http://kidney.niddk.nih.gov/kudiseases/pubs/glossary/index.htm.
29 National Kidney Federation. Glossary. 2007. www.kidney.org.uk/Medical-Info/glossary/glossary.html.
30 Kidney Foundation of Canada. KFoC SK Branch—glossary of terms. 2007. www.kidney.ca.
31 Family Practice Notebook. Nephrology. 2007. www.fpnotebook.com/REN.htm.
32 Sandhu S, Wiebe N, Fried LF, Tonelli M. Statins for improving renal outcomes: a meta-analysis. *J Am Soc Nephrol* 2006;17:2006-16.
33 Landoni G, Biondi-Zoccai GG, Tumlin JA, Bove T, De LM, Calabro MG, et al. Beneficial impact of fenoldopam in critically ill patients with or at risk for acute renal failure: a meta-analysis of randomized clinical trials. *Am J Kidney Dis* 2007;49:56-68.
34 Webster A, Woodroffe RC, Taylor RS, Chapman JR, Craig JC. Tacrolimus versus cyclosporin as primary immunosuppression for kidney transplant recipients. *Cochrane Database Syst Rev* 2005;(4):CD003961.

35 Friedrich JO, Adhikari N, Herridge MS, Beyene J. Meta-analysis: low-dose dopamine increases urine output but does not prevent renal dysfunction or death. *Ann Intern Med* 2005;142:510-24.

36 Fabrizi F, Dixit V, Magnini M, Elli A, Martin P. Meta-analysis: intradermal vs. intramuscular vaccination against hepatitis B virus in patients with chronic kidney disease. *Aliment Pharmacol Ther* 2006;24:497-506.

37 Sladek RM, Tieman J, Currow DC. Improving search filter development: a study of palliative care literature. *BMC Med Inform Decis Mak* 2007;7:18.

38 Wilczynski NL, Haynes RB. Optimal search strategies for identifying mental health content in MEDLINE: an analytic survey. *Ann Gen Psychiatry* 2006;5:4.

39 Ramos-Remus C, Suarez-Almazor M, Dorgan M, Gomez-Vargas A, Russell AS. Performance of online biomedical databases in rheumatology. *J Rheumatol* 1994;21:1912-21.

40 Portaluppi F. Consistency and accuracy of the Medical Subject Headings thesaurus for electronic indexing and retrieval of chronobiologic references. *Chronobiol Int* 2007;24:1213-29.

41 Funk ME, Reid CA. Indexing consistency in MEDLINE. *Bull Med Libr Assoc* 1983;71:176-83.

42 Boynton J, Glanville J, McDaid D, Lefebvre C. Identifying systematic reviews in MEDLINE: developing an objective approach to search strategy design. *J Inform Sci* 1998;24:137-54.

43 Boersma C, Atthobari J, Gansevoort RT, de Jong-Van den Berg LT, de Jong PE, de Zeeuw D, et al. Pharmacoeconomics of angiotensin II antagonists in type 2 diabetic patients with nephropathy: implications for decision making. *Pharmacoeconomics* 2006;24:523-35.

44 Mazzaferro S, Barresi G, Baldinelli M, Malagnino E, Matera G, Otranto G, et al. Utilization of guidelines for the management of hypertension in cardiovascular risk scoring of renal patients. *Clin Ter* 2006;157:327-32.

45 Rodriguez GM, Naves DM, Cannata Andia JB. Bone metabolism, vascular calcifications and mortality: associations beyond mere coincidence. *J Nephrol* 2005;18:458-63.

46 Campbell SB, Hothersall E, Preston J, Brown AM, Hawley CM, Wall D, et al. Frequency and severity of acute rejection in live- versus cadaveric-donor renal transplants. *Transplantation* 2003;76:1452-7.

47 Salmela K. Quality and safety of organ transplantation. *Ann Transplant* 2004;9:51-2.

48 Hsu CY, Chertow GM. Chronic renal confusion: insufficiency, failure, dysfunction, or disease. *Am J Kidney Dis* 2000;36:415-8.

49 Hersh WR, Crabtree MK, Hickam DH, Sacherek L, Friedman CP, Tidmarsh P, et al. Factors associated with success in searching MEDLINE and applying evidence to answer clinical questions. *J Am Med Inform Assoc* 2002;9:283-93.

Accepted: 30 March 2009