Optimal search filters for detecting quality improvement studies in Medline

N L Wilczynski, R B Haynes, for the QI Hedges Team

Health Information Research Unit, McMaster University, Canada

Correspondence to
Professor R Brian Haynes, Health Information Research Unit, McMaster University, 1280 Main Street West, CRL 133, Hamilton, Ontario L8S 4K1, Canada;
bhaynes@mcmaster.ca

Accepted 17 May 2010
Published Online First 29 July 2010

ABSTRACT

Background As the knowledge translation and comparative effectiveness research agendas gain momentum, we can expect more evidence on which to base quality improvement (QI) programmes. Unaided searches for such content in the literature, however, are likely to be daunting, with searches missing key articles while mainly retrieving articles that are irrelevant to the question being asked. The objective of this study was to develop and validate optimal Medline search filters for retrieving original and review articles about clinical QI.

Methods Analytical survey in the McMaster Clinical Hedges database and Health Knowledge Refinery (HKR) of 161 clinical journals to determine the operating characteristics of QI search filters developed by computerised combinations of terms selected to detect original QI studies and systematic reviews meeting basic methodological criteria for scientific merit. Results from a derivation random subset of articles were tested in a validation random subset.

Results The Clinical Hedges QI database contained 49 233 citations of which 471 (0.96%) were original or review QI studies; of those, 282 (60%) were methodologically sound. Combinations of search terms reached peak sensitivities of 100% at a specificity of 89.3% for detecting methodologically sound original and review QI studies, and sensitivities of 97.6% at a specificity of 53.0% for detecting all original and review QI studies independent of rigour. Operating characteristics of the search filters derived in the development database worked similarly in the validation database, without statistical differences.

Conclusion New empirically derived Medline search filters have been validated to optimise retrieval of original and review QI articles.

INTRODUCTION

Clinical quality improvement (QI) is essential if optimal healthcare, balancing patient benefits and safety, is to be achieved. Both health services administrators and health professionals bear responsibility for the quality of care.[1] The Charter on Medical Professionalism[2] outlines 10 professional responsibilities for the physician for the next millennium, one of which is the commitment to improving quality of care. The physician is expected to assist in the creation and implementation of mechanisms designed to encourage continuous QI by taking an active role in reducing medical error, increasing patient safety, minimising overuse of healthcare resources and optimising the outcomes of care.

QI efforts have been implemented with varying degrees of success. Research has shown that physician engagement and leadership are key elements for the success of QI efforts.[3] Engaging physicians can be difficult but one of the facilitators to physician involvement is the use of QI methods that have been shown to work.[4] Determining which QI methods work can be challenging, however, because the evaluation of these methods is usually first widely accessible in the biomedical and health services journal literature available by online searching of electronic databases such as Medline which contains over 19 million articles from over 5000 journals.[5] Finding QI research in these electronic databases can be difficult because it is not well organised in the literature (which is typically better organised within disease categories, while QI research is often cuts across disease categories), and many end users do not find what they need, while investigators doing QI studies often 're-invent the wheel' for lack of easy access to other studies in the field. For QI practice and research to advance, it is important to have dependable and efficient access to this research evidence.

If large electronic bibliographic databases such as Medline are to be helpful to end users, they must be able to retrieve articles that are both scientifically sound and directly relevant to the healthcare problem they are trying to solve, without missing key studies or retrieving excessive numbers of preliminary, irrelevant, outdated or misleading reports. Our approach and the approach of others to these problems has been to develop search filters ('hedges') to improve the retrieval of clinically relevant and scientifically sound study reports from Medline and similar bibliographic databases.[6–19] Search filters can be created by combining disease content terms ('ANDed') with Medical Subject Headings (MeSH), explosions (exp), publication types (pt), subheadings (sh) and textwords (tw) that detect research design features for applied healthcare research. The use of search filters has been shown to increase the relevancy and reduce the volume of information retrieved.[6–12]

To our knowledge, only one published study has tested the efficiency of retrieving QI studies. The study, conducted by Balas and colleagues,[20] set out to measure the efficiency of simple searches in retrieving controlled evidence about seven specific primary healthcare QI interventions (home care, patient education, patient reminder, physician education, physician reminder, provider feedback and telephone follow-up) and seven effect variables (C-section rate, cost of care, follow-up visits, hospitalisation rate, immunisation rate, length of stay and number of prescriptions). All searches were restricted to the MeSH publication type 'randomised controlled trial.' The authors reported

Qual Saf Health Care: first published as 10.1136/qshc.2010.042432 on 29 July 2010. Downloaded from http://qualitysafety.bmj.com/ on April 16, 2024 by guest. Protected by copyright.

that MeSH term searches had an overall recall (sensitivity) of 58% and text word searches performed significantly lower at 11%.

The objective of the research reported in this paper is to determine if empirical search filters ('hedges') can be created that optimise the yield of original and review articles about QI from Medline, and separate higher-quality studies from others.

METHODS

The basic methodology for our search filter development research has been described in a previous publication.[21] Briefly, search filters are developed and validated using a diagnostic testing model. After constructing the 'hand search' literature which serves as the gold standard, search filters are developed and validated, treating the search terms related to research design features as 'diagnostic tests.' The hand search by research staff is required to determine which articles should be retrieved, and which articles should not be retrieved.

Hand search

To construct the hand search for this project, we used two similarly constructed sources, the McMaster University's Clinical Hedges Database and Health Knowledge Refinery (HKR). Research staff at McMaster University produced the content for both sources using identical critical appraisal criteria for determining the methodological soundness of clinically relevant reviews and primary studies. More specifically, the Clinical Hedges Database was constructed by six research assistants in the Health Information Research Unit (HIRU) hand searching 161 journals titles for the year 2000. The selection of the 161 journal titles reviewed was based on recommendations of clinicians and librarians, Science Citation Index Impact Factors provided by the Institute for Scientific Information, and ongoing assessment of their yield of studies and reviews of scientific merit and clinical relevance for all major medical disciplines. Combined with the HKR content, the coverage includes general medical practice, internal medicine and subdisciplines, mental health, surgery, obstetrics and women's health, nursing care and rehabilitation (a list of journals can be found at http://hiru.mcmaster.ca/hiru/hedges/Medline%20journals%20read.pdf).

The research assistants categorised all original and review studies found in these journals for eight purpose categories (treatment/QI, diagnosis, prognosis, aetiology, clinical prediction guide, economics, cost and qualitative) and then applied methodological criteria to determine if the article was methodologically sound for all categories except cost and qualitative. All purpose category definitions and corresponding methodological rigour were outlined in a previous paper.[22] Research staff were rigorously calibrated prior to reviewing the 2000 literature and inter-rater agreement for application of all criteria exceeded 80% beyond chance.[22]

During the initial hand search of the 161 journals, treatment and QI studies were assessed as one category. As part of this study, we reviewed all citations tagged as original treatment and review treatment (n=8631) for whether the content was about QI defined as 'content pertains directly to interventions intended to improve the quality of healthcare, including studies of continuing education for the purpose of improving the quality of care. The focus is on the providers and processes of care.' The programmers in HIRU produced a web-based interface containing all original and review treatment citations from the Clinical Hedges Database. Two research assistants reviewed the full-text article (if needed) independently and in duplicate to determine if the article should be retagged as QI according to the definition stated above. Disagreements and agreements of 'unsure' were resolved by a third party. Methodological rigour of the QI studies were not retagged because the criteria for a methodologically sound ('pass') QI study (original or review) were the same as for a treatment study. The methodological criteria applied for original studies of QI or treatment were: random allocation of participants to comparison groups; outcome assessment of at least 80% of those entering the investigation accounted for in one major analysis at any given follow-up assessment; and analysis consistent with study design. Randomised controlled trials provide top-quality evidence if conducted rigorously for QI questions where the interventions are changes in the process of care and the quality outcome measures correspond to clinical effects. The methodological criteria applied for review studies of QI or treatment were: a clear statement of the clinical topic of the review, a methods section indicating how the evidence was retrieved and from what sources, an explicit statement of the inclusion and exclusion criteria, and inclusion of at least one study that passed on the methodological criteria for a original study of QI or treatment. The final retagged results were reintegrated into the Clinical Hedges Database while renaming the database to Clinical Hedges QI Database.

It was necessary to increase the sample size of methodologically sound QI studies that were in the hand search database (n=77 after retagging the content of the Clinical Hedges Database) because previous research shows that at least 99 articles in the category of interest are needed when developing and validating search filters.[23] To increase the sample size, a second source was used; McMaster University's Health Knowledge Refinery (HKR) and more specifically McMaster PLUS (http://hiru.mcmaster.ca/hiru/HIRU_McMaster_PLUS_projects.aspx). McMaster PLUS is produced by research staff in HIRU. They critically appraise the context of more than 120 clinically relevant journals (list of journals found at http://hiru.mcmaster.ca/hiru/journalslist.asp) on an ongoing basis. Original and review articles in each issue of these journals are categorised and assessed for methodological rigour according to the criteria found at http://hiru.mcmaster.ca/hiru/InclusionCriteria.html. Among other categories, QI studies are identified and assessed for methods. Those that are methodologically sound are rated for relevancy and newsworthiness by practising physicians from around the world and subsequently enter the McMaster PLUS database. The McMaster PLUS database contains content from 2003 on, and methodologically sound original and review QI studies added between 2003 and 23 September 2008 were integrated into the Clinical Hedges QI Database.

Search terms

During the Clinical Hedges Study conducted in 2000, an initial list of MeSH terms and text words was compiled. Input was then sought from clinicians and librarians in the USA and Canada through interviews of known searchers, requests at meetings and conferences, and requests to the National Library of Medicine. Individuals were asked to identify which terms or phrases they used when searching for studies of prognosis, causation, diagnosis, treatment, economics, clinical prediction guides, reviews and costs, and of a qualitative nature. Terms could be from MeSH, including publication types and subheadings, or could be text words denoting methodology in titles and abstracts of articles. We compiled a list of 5395 terms of which 4862 were unique (list of terms tested provided by the authors upon request). All search terms originally tested in the

Qual Saf Health Care: first published as 10.1136/qshc.2010.042432 on 29 July 2010. Downloaded from http://qualitysafety.bmj.com/ on April 16, 2024 by guest. Protected by copyright.

Clinical Hedges Study as well as all the terms identified by other researchers that have been tested since we developed our search filters which are available on the Clinical Queries page of PubMed (http://www.ncbi.nlm.nih.gov/sites/pubmedutils/clinical) were tested in this study. Additionally, we reviewed this list of terms and added text words and MeSH that are relevant to identifying QI studies that did not appear. The term list was expanded by reviewing the Cochrane Effective Practice and Organisation of Care (EPOC) site (http://www.mrw.interscience.wiley.com/cochrane/clsysrev/articles/CD003319/frame.html) and the search filters used in a Technical Review of QI strategies for the Agency for Healthcare Research and Quality.[24] We also reviewed the indexing of QI studies included in McMaster PLUS and contacted searching experts at McMaster University, the University of Ottawa, and the US National Library of Medicine.

Clinical hedges QI database and search filter development

The Clinical Hedges QI database was randomly split using Microsoft Windows' random number generator into components of 60% and 40%. Search strategies were initially tested and developed in 60% of the database (development) and then validated in 40% of the database (validation). The sensitivity, specificity, precision and accuracy of Medline searches were determined (table 1). Sensitivity for a given topic is defined as the proportion of QI articles or high-quality QI articles that are retrieved; specificity is the proportion of non-QI articles or low-quality QI articles not retrieved; precision is the proportion of retrieved articles that are QI or of high-quality QI; and accuracy is the proportion of all articles that are correctly classified.

Four types of search filters were developed: (1) maximise the sensitivity for retrieving original and review QI studies; (2) optimise the balance of sensitivity and specificity for retrieving original and review QI studies; (3) maximise the sensitivity for retrieving original and review QI studies that are methodologically sound (ie, 'pass' on the methods criteria outlined earlier); and (4) optimise the balance of sensitivity and specificity for retrieving original and review QI studies that are methodologically sound. We developed single-term search filters and ORed combinations of 4 terms. All terms were tested to determine the best single term, and the following cut-offs were used for the testing of ORed combinations: elimination of single terms with ≤ 25% sensitivity and ≤75% specificity when developing two-term combinations; elimination of two-term combinations with ≤75% sensitivity and ≤50% specificity when developing three-term combinations; and elimination of two-term combinations with ≤75% sensitivity and ≤50% specificity when developing 4-term combinations.

RESULTS

The Clinical Hedges QI database contained 49 233 citations, of which 471 (0.96%) were original or review QI studies; of those 282 (60%) were methodologically sound (ie, 'pass'). Of the 282 original or review QI studies that were methodologically sound, 77 (27%) were published in the year 2000, and the remaining 205 (73%) were published in the years 2003 to 2008. All other citations (n=49 028) in the database were published in the year 2000. The breakdown of the number of citations included in the development and validation databases is shown in table 2.

Table 3 shows the single terms with the best sensitivity (keeping specificity ≥50%) and best optimisation of sensitivity and specificity (based on the lowest possible absolute difference between sensitivity and specificity) for detecting original and review articles about QI, and original and review articles about QI that pass for methods in MEDLINE. Comparison of the operating characteristics in the development and validation databases is also shown. The single term 'exp health services administration' had the best sensitivity (77.8%) and the best optimisation of sensitivity (77.8%) and specificity (77.8%) for detecting original and review articles about QI. The single term 'control:.mp.' had the best sensitivity (90.8%) for detecting original and review articles about QI that pass, whereas 'random:.mp.' provided the best optimisation of sensitivity (90.2%) and specificity (90.5%). The operating characteristics for the single terms shown in table 3 did not differ when comparing the results from the development and validation databases.

Table 4 shows the results for four-term ORed combinations (three-term combinations are shown if they outperform four-term combinations). The search filter 'exp health services administration OR random: .mp. OR review.pt. OR compare:.tw.' had 97.6% sensitivity with a specificity of 53.0% for detecting original and review articles about QI, whereas the search filter 'random:.ti,ab. OR educat:.tw. OR exp patient care management' had the best balance between sensitivity (83.3%) and specificity (83.9%). The search filter 'effectiveness.tw. OR journal.mp. OR MEDLINE.tw. OR random:.tw.' had 100% sensitivity with a specificity of 89.3% for detecting original and review articles about QI that pass for methods, and the search filter 'control: trial:.mp. OR journal.mp. OR MEDLINE .tw. OR random: trial:.tw.' had the best balance between sensitivity (94.8%) and specificity (95.7%). The operating characteristics for the search filters shown in table 4 did not differ when comparing the results from the development and validation databases.

DISCUSSION

We developed search filters that can assist clinicians and researchers in retrieving original and review articles about QI and in retrieving methodologically sound original and review studies about QI from Medline. Search filters that maximise sensitivity were developed as opposed to those that maximise specificity because there are so few QI studies in the medical literature. Additionally, we developed search filters that optimised the balance between sensitivity and specificity as an option for searching when the retrieval using the sensitive search becomes unmanageable due to time constraints. When using the latter search filter, the user risks missing a few QI studies. Additionally, search filters that retrieve both original (primary) studies and reviews about QI were developed because it is

Table 1 Formulae for calculating the sensitivity, specificity, precision and accuracy of medline searches

		Hand search of the literature	
		Meets criteria	Does not meet criteria
Search terms	Detected	a	b
	Not detected	c	d
		a+c	b+d

Sensitivity=a/(a+c). Specificity=d/(b+d). Precision=a/(a+b). Accuracy=(a+d)/(a+b+c+d).

Table 2 Number of citations included in the development and validation Clinical Hedges QI Databases

Article type	Development database	Validation database
All articles	29540	19693
Original or review QI	288	183
Original or review QI pass	173	109

Qual Saf Health Care: first published as 10.1136/qshc.2010.042432 on 29 July 2010. Downloaded from http://qualitysafety.bmj.com/ on April 16, 2024 by guest. Protected by copyright.

Table 3 Single term with the best sensitivity (keeping specificity ≥50%) and best optimisation of sensitivity and specificity (based on the lowest possible absolute difference between sensitivity and specificity) for detecting original and review articles about quality improvement (QI) and for detecting original and review articles about QI that pass for methods in Medline

Search term OVID search*	Sensitivity (%) development validation diff (95% CI)	Specificity (%) development validation diff (95% CI)	Precision (%) development validation diff (95% CI)	Accuracy (%) development validation diff (95% CI)
Best sensitivity				
All QI				
exp health services administration	77.8 83.1 5.3 (−2.2 to 12.4)	77.8 78.5 0.7 (−0.1 to 1.4)	3.3 3.5 0.2 (−0.5 to 0.9)	77.8 78.5 0.7 (−0.0 to 1.5)
QI pass for methods				
control:.mp.	90.8 87.2 −3.6 (−12.0 to 3.7)	83.5 84.0 0.5 (−0.2 to 1.1)	3.1 2.9 −0.2 (−1.0 to 0.5)	83.5 84.0 0.5 (−0.2 to 1.1)
Best optimisation of sensitivity and specificity				
All QI				
exp health services administration	77.8 83.1 5.3 (−2.2 to 12.4)	77.8 78.5 0.7 (−0.1 to 1.4)	3.3 3.5 0.2 (−0.5 to 0.9)	77.8 78.5 0.7 (−0.0 to 1.5)
QI pass for methods				
random:.mp.	90.2 84.4 −5.8 (−14.6 to 2.0)	90.5 90.2 −0.3 (−0.8 to 0.2)	5.3 4.6 −0.7 (−1.9 to 0.5)	90.5 90.1 −0.4 (−0.9 to 0.2)

*The search strategy is reported using Ovid's search engine syntax for Medline.
diff, difference, comparing the development and validation data sets using the iterative method of Miettinen and Nurminen for two independent binomial proportions. None of the differences were statistically significant; exp, explode, a search term that automatically includes closely related indexing terms; :, truncation; mp, multiple posting (term appears in the title, abstract, or MeSH heading).

advantageous to retrieve the systematic reviews while retrieving the original studies.

Our multiple term search results shown in table 4 show that 100% sensitivity can be achieved when searching for methodologically sound original and review articles about QI and that 97.6% sensitivity can be achieved when searching for original and review articles about QI independent of methodological rigour. Search filters are a helpful but imperfect solution to accurate literature retrieval in large online bibliographic databases. Even our top-performing strategies had generally low precision, 5.0% and 2.0%, respectively, because Medline is such a large and broad-ranging database. Because precision in large multipurpose databases is inevitably low due to the small concentration of relevant articles, it is especially important that terminology (eg, methodological terminology) used in QI studies be as accurate, explicit and consistent as possible to facilitate the indexing process and improve the search success. Typically, however, when searching for QI evidence, content terms are added, and our previous research shows that the addition of content terms increases precision.[25] For example, when searching Ovid Medline for the previous year, the sensitive search filter for detecting sound original and review QI studies

Table 4 Combination of terms with the best sensitivity (keeping specificity ≥50%) and best optimisation of sensitivity and specificity (based on the lowest possible absolute difference between sensitivity and specificity) for detecting original and review articles about quality improvement (QI) and for detecting original and review articles about QI that pass for methods in Medline

Search strategy OVID search*	Sensitivity (%) development validation diff (95% CI)	Specificity (%) development validation diff (95% CI)	Precision (%) development validation diff (95% CI)	Accuracy (%) development validation diff (95% CI)
Best sensitivity				
All QI				
exp health services administration OR random:.mp. OR review.pt. OR compare:.tw	97.6 98.4 0.8 (−2.5 to 3.6)	53.0 53.3 0.3 (−0.6 to 1.1)	2.0 1.9 −0.1 (−0.4 to 0.3)	53.5 53.7 0.2 (−0.7 to 1.1)
QI pass for methods				
effectiveness.tw. OR journal.mp. OR MEDLINE.tw. OR random:.tw	100.0 96.3 −3.7 (−9.1 to 1.4)	89.3 89.3 0	5.2 4.8 −0.4 (−1.6 to 0.7)	89.3 0
Best optimisation of sensitivity and specificity All QI				
random:.ti,ab. OR educat:.tw. OR exp patient care management	83.3 84.7 1.4 (−5.7 to 8.0)	83.9 84.1 0.2 (−0.5 to 0.9)	4.9 4.8 −0.1 (−1.0 to 0.9)	83.9 84.1 0.2 (−0.5 to 0.9)
QI pass for methods				
control: trial:.mp. OR journal.mp. OR MEDLINE.tw. OR random: trial:.tw	94.8 89.9 −4.9 (−12.5 to 1.3)	95.7 95.8 0.1 (−0.3 to 0.5)	11.4 10.6 −0.8 (−3.4 to 1.8)	95.7 95.7 0

*Search strategies are reported using Ovid's search engine syntax for MEDLINE.
diff, difference, comparing the development and validation data sets using the iterative method of Miettinen and Nurminen for two independent binomial proportions. None of the differences were statistically significant; exp, explode, a search term that automatically includes closely related indexing terms; :, truncation; mp, multiple posting (term appears in title, abstract, or MeSH heading); pt, publication type; tw, textword (word or phrase appears in a text word field, ie, appropriate for a subject search—includes the title and abstract fields); ti,ab, term appears in the title or abstract.

Qual Saf Health Care: first published as 10.1136/qshc.2010.042432 on 29 July 2010. Downloaded from http://qualitysafety.bmj.com/ on April 16, 2024 by guest. Protected by copyright.

('effectiveness.tw. OR journal.mp. OR MEDLINE.tw. OR random:.tw.') yields over 500 000 citations. When this filter is ANDed with the content term 'diabetes care.tw.' the yield decreases to 180 citations with 13 of the first 20 articles being directly related to QI in diabetes care. For example, the fourth citation retrieved investigated, in an RCT, the impact on A1c and other diabetes outcomes of providing literacy- and numeracy-sensitive diabetes care within an enhanced diabetes care programme.[26]

Another mechanism that could potentially facilitate the retrieval of QI or quality of care articles is to have a designated index term that is assigned exclusively to these types of articles. In this study, we tested several index terms that were related to quality improvement, including 'quality assurance, healthcare' and 'quality of healthcare,' and found low levels of sensitivity (<10%).

Although one set of search filters have been developed to retrieve QI studies meeting certain methodological rigour criteria, it does not necessarily imply that the articles retrieved will meet the highest methodological standards. Ultimately, the end user has the responsibility for appraising the retrieved literature for quality and relevance before applying it to clinical practice.

A possible limitation of our study is the use of a literature database constructed in the year 2000. This would only affect the results if there were substantive changes in the indexing and reporting of studies in Medline since 2000. We do not know of any such changes and in any event, augmented the database with QI studies and reviews meeting similar criteria during the period from 2003 to 2008.

Comparing the results of this study with those of Balas and colleagues,[20] we found that the retrieval of original and review articles about QI and the retrieval of methodologically sound original and review articles about QI in Medline can be enhanced by the use of search filters. Although beneficial filters exist, such as those reported here, further work is needed to improve dissemination and publication of filters so that health service providers, clinicians, researchers and librarians are aware of them, but also have greater knowledge and proficiency in using them effectively.

CONCLUSIONS

Several Medline search filters can achieve high performance in retrieving original and review articles about QI and in retrieving original and review articles that are methodologically sound. Using a sensitive search filter will ensure that few QI articles are missed. Alternatively, using a search filter that optimises the balance between sensitivity and specificity will result in a more manageable retrieval given time constraints.

Acknowledgements Members of the QI Hedges Team are B Haynes, N Wilczynski, A McKibbon, S Marks, A Eady, C Walker-Dilks, S Wong, L Baier, J Moffatt and N Hobson.

Funding Canadian Institutes of Health Research, Ottawa, Ontario, Canada.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Wharam JF,** Sulmasy D. Improving the quality of care: who is responsible for what? *JAMA* 2009;**301**:215—17.
2. **ABIM Foundation.** American Board of Internal Medicine; ACP—ASIM Foundation. American College of Physicians—American Society of Internal Medicine; European Federation of Internal Medicine. Medical professionalism in the new millennium: a physician charter. *Ann Intern Med* 2002;**136**:243—6.
3. **Bradley EH,** Holmboe ES, Mattera J, *et al*. A qualitative study of increasing beta-blockers use after myocardial infarction: Why do some hospitals succeed? *JAMA* 2001;**285**:2604—11.
4. **Holmboe ES,** Cassel CK. The role of physicians and certification boards to improve quality. *Am J Med Qual* 2007;**22**:18—25.
5. NLM Bibliographic Services Division, viewed February 22, 2010. http://www.nlm.nih.gov/pubs/factsheets/medline.html.
6. **Haynes RB,** McKibbon KA, Wilczynski NL, *et al*; Hedges Team. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005;**330**:1179—82.
7. **Montori VM,** Wilczynski NL, Morgan D, *et al*; Hedges Team. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. *BMJ* 2005;**330**:68—73.
8. **Wilczynski NL,** Haynes RB, Lavis JN, *et al*; HSR Hedges team. Optimal search strategies for detecting health services research studies in MEDLINE. *CMAJ* 2004;**171**:1179—85.
9. **Wong SS,** Wilczynski NL, Haynes RB; Hedges Team. Developing optimal search strategies for detecting clinically relevant qualitative studies in MEDLINE. *Stud Health Technol Inform* 2004;**107**(Pt 1):311—16.
10. **Wilczynski NL,** Haynes RB; Hedges Team. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey. *BMC Med* 2004;**2**:23.
11. **Haynes RB,** Wilczynski NL. Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: analytical survey. *BMJ* 2004;**328**:1040—42.
12. **Wong SS,** Wilczynski NL, Haynes RB, *et al*; Hedges Team. Developing optimal search strategies for detecting sound clinical prediction studies in MEDLINE. *AMIA Annu Symp Proc* 2003:728—32.
13. **Wilczynski NL,** Haynes RB; Hedges Team. Developing optimal search strategies for detecting clinically sound causation studies in MEDLINE. *AMIA Annu Symp Proc* 2003:719—23.
14. **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;**31**:150—3.
15. **Dumbrigue HB,** Esquivel JF, Jones JS. Assessment of MEDLINE search strategies for randomized controlled trials in prosthodontics. *J Prosthodont* 2000;**9**:8—13.
16. **Sjogren P,** Halling A. Medline search validity for randomised controlled trials in different areas of dental research. *Br Dent J* 2002;**192**:97—9.
17. **Bachmann LM,** Coray R, Estermann P, *et al*. Identifying diagnostic studies in MEDLINE: reducing the number needed to read. *J Am Med Inform Assoc* 2002;**9**:653—8.
18. **Deville WL,** Bezemer PD, Bouter LM. Publications on diagnostic test evaluation in family medicine journals: an optimal search strategy. *J Clin Epidemiol* 2000;**5**:65—9.
19. **Vincent S,** Greenley S, Beaven O. Clinical Evidence diagnosis: Developing a sensitive search strategy to retrieve diagnostic studies on deep vein thrombosis: a pragmatic approach. *Health Info Libr J* 2003;**20**:150—9.
20. **Balas EA,** Stockham MG, Mitchell JA, *et al*. In search of controlled evidence for health care quality improvement. *J Med Syst* 1997;**21**:21—32.
21. **Wilczynski NL,** Morgan D, Haynes RB; Hedges Team. An overview of the design and methods for retrieving high-quality studies for clinical care. *BMC Med Inform Decis Mak* 2005;**5**:20.
22. **Wilczynski NL,** McKibbon KA, Haynes RB. Enhancing retrieval of best evidence for health care from bibliographic databases: calibration of the hand search of the literature. *Stud Health Technol Inform* 2001;**84**(Pt 1):390—3.
23. **Yao X,** Wilczynski NL, Walter SD, *et al*. Sample size determination for bibliographic retrieval studies. *BMC Med Inform Decis Mak* 2008;**8**:43.
24. **Ranji SR,** Steinman MA, Shojania KG, *et al*. Antibiotic prescribing behavior. InVol. 4 of: Shojania KG, McDonald KM, Wachter RM, Owens DK, eds. *Closing The Quality Gap: A Critical Analysis of Quality Improvement Strategies. Technical Review 9 (Contract No 290-02-0017 to the Stanford University-UCSF Evidence-based Practice Center). AHRQ Publication No 04(06)-0051-2*. Rockville, MD: Agency for Healthcare Research and Quality, 2006.
25. **Wilczynski NL,** Haynes RB; Team hedges. Optimal search strategies for identifying mental health content in MEDLINE: an analytic survey. *Ann Gen Psychiatry* 2006;**5**:4.
26. **Cavanaugh K,** Wallston KA, Gebretsadik T, *et al*. Addressing literacy and numeracy to improve diabetes care: two randomized controlled trials. *Diabetes Care* 2009;**32**:2149—55.

Qual Saf Health Care: first published as 10.1136/qshc.2010.042432 on 29 July 2010. Downloaded from http://qualitysafety.bmj.com/ on April 16, 2024 by guest. Protected by copyright.