Cite this article as: BMJ, doi:10.1136/bmj.38336.804167.47 (published 24 December 2004)

Papers

Optimal search strategies for retrieving systematic reviews from Medline: analytical survey

Victor M Montori, Nancy L Wilczynski, Douglas Morgan, R Brian Haynes, for the Hedges Team

Abstract

Objective: To develop optimal search strategies in Medline for retrieving systematic reviews.
Design: Analytical survey.
Data sources: 161 journals published in 2000 indexed in Medline.
Main outcome measures: The sensitivity, specificity, and precision of retrieval of systematic reviews of 4862 unique terms in 782 485 combinations of one to five terms were determined by comparison with a hand search of all articles (the criterion standard) in 161 journals published during 2000 (49 028 articles).
Results: Only 753 (1.5%) of the 49 028 articles were systematic reviews. The most sensitive strategy included five terms and had a sensitivity of 99.9% (95% confidence interval 99.6% to 100%) and a specificity of 52% (51.6% to 52.5%). The strategy that best minimised the difference between sensitivity and specificity had a sensitivity of 98% (97% to 99%) and specificity of 90.8% (90.5% to 91.1%). Highest precision for multiterm strategies, 57% (54% to 60%), was achieved at a sensitivity of 71% (68% to 74%). The term "cochrane database of systematic reviews.jn." was the most precise single term search strategy (sensitivity of 56% (52% to 60%) and precision of 96% (94% to 98%)). These strategies are available through the "limit" screen of Ovid's search interface for Medline.
Conclusions: Systematic reviews can be retrieved from Medline with close to perfect sensitivity or specificity, or with high precision, by using empirical search strategies.

Introduction

Systematic reviews exhaustively search for, identify, and summarise the available evidence that addresses a focused clinical question, with particular attention to methodological quality. When these reviews include meta-analysis, they can provide precise estimates of the association or the treatment effect.[1] Clinicians can then apply these results to the wide array of patients who do not differ importantly from those enrolled in the summarised studies. Systematic reviews can also inform investigators about the frontier of current research. Thus, both clinicians and researchers should be able to reliably and quickly find valid systematic reviews of the literature.

Finding these reviews in Medline poses two challenges. Firstly, only a tiny proportion of citations in Medline are for literature reviews, and only a fraction of these are systematic reviews. Secondly, the National Library of Medicine's Medlars indexing procedures do not include "systematic review" as a "publication type." Rather, the indexing terms and publication types include a number of variants for reviews, including "meta-analysis" (whether or not from a systematic review)[2]; "review, academic"; "review, tutorial"; "review literature"; as well as separate terms for articles that often include reviews, such as "consensus development conference", "guideline", and "practice guideline". The need for special search strategies (hedges) for systematic reviews could be substantially reduced if such reviews were indexed by a separate publication type, but indexers need to be able to dependably distinguish systematic reviews from other reviews. Pending this innovation, there is need for validated search strategies for systematic reviews that optimise retrieval for clinical users and researchers.

Since 1991, our group and others have proposed search strategies to retrieve citations of clinically relevant and scientifically sound studies and reviews from Medline.[3] Our approach relies on developing a database of articles resulting from a painstaking hand search of a set of high impact clinical journals, assessing the methodological quality of the relevant articles, collecting search terms suggested by librarians and clinical users, generating performance metrics from single terms and combination of these terms in a derivation database, and testing the best strategies in a validation database. However, we did not produce a systematic review hedge in 1991 because there were few such studies in the 10 journals we reviewed at that time. Without the benefit of such data, we proposed strategies that have since been reproduced in library websites and tutorials, but for which there are no performance data.[4] Since then, the Cochrane Collaboration has greatly increased the production of systematic reviews, and we have created a new database with 161 journals that are indexed in Medline.

Other groups have published strategies to retrieve systematic reviews from Medline. Researchers at the Centre for Reviews and Dissemination of the University of York developed strategies to identify systematic reviews to populate DARE, the Database of Abstracts of Reviews of Effects, which includes appraised systematic reviews obtained from searching Medline and handsearching selected journals.[5] These strategies resulted from careful statistical analysis of the frequency with which certain words appeared in the abstracts of systematic reviews. Researchers tested these strategies on the Ovid interface and found their sensitivity was ≥98% and precision about 20%.

Shojania and Bero developed the strategy programmed into the searching interfaces for PubMed (as a clinical query) and the Medline database on Ovid (as a limit).[6] The authors nominated terms, assembled them in a logical strategy, and tested this strat-



A table showing PubMed translations of the Ovid search strategies is on bmj.com

BMJ: first published as 10.1136/bmj.38336.804167.47 on 24 December 2004. Downloaded from http://www.bmj.com/ on 29 March 2024 by guest. Protected by copyright.

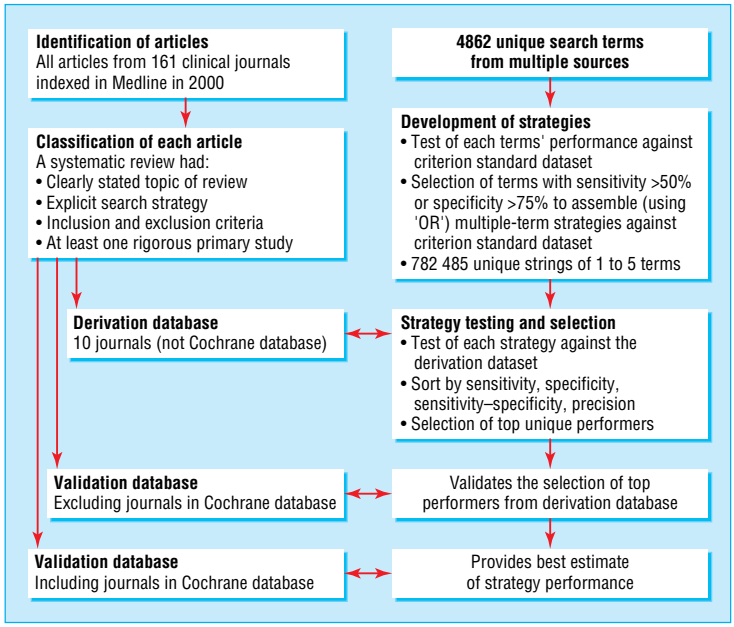


Fig 1 Manual review to build criterion standard dataset

egy in PubMed against a criterion standard. This standard comprised 100 reviews found on DARE and 100 systematic reviews highlighted in *ACP Journal Club* because of their methodological quality and clinical relevance. This strategy had a reported sensitivity ≥ 90% and a precision (for a given clinical topic) ≥ 50%.

In this paper we report on the generation, validation, and performance characteristics of new search strategies to identify systematic reviews in Medline, and compare them with previously published strategies.

Methods

We developed search strategies by using methodological search terms and phrases in a subset of Medline records matched with a handsearch of the contents of 161 journal titles for 2000; this manual review of the literature represented the criterion standard dataset. We have described elsewhere and in detail the process of building the gold standard dataset and classifying the records by purpose (for example, diagnosis, therapy), and methodological quality (for example, valid systematic review, unsystematic review).[7] Here, we describe this process only briefly (see fig 1).

Guided by recommendations from clinicians and librarians, and taking into account the Science Citation Index impact factors, as well as an iterative process of evaluation of over 400 journals for yield of studies and reviews of scientific merit and clinical relevance, we defined a set of 161 clinical journals for the fields of general internal medicine, family practice, nursing, and mental health that were indexed in Medline.[8] Data about all items found in each journal for the year 2000 were entered into a data collection form, verified, and entered onto a Microsoft Access database. Then, each article was searched in Medline for 2000 (using the Ovid interface) and the full Medline record captured and linked with the handsearch data.

Study classification

We defined a review as any full text article that was shown as a review, overview, or meta-analysis in the title or in a section heading, or that indicated in the text that the intention of the authors was to review or summarise the literature on a particular topic.[9] For an article to be considered a systematic review, the authors had to clearly state the clinical topic of the review and how the evidence was retrieved and from what sources, and they had to provide explicit inclusion and exclusion criteria and include at least one study that passed methodological criteria for the purpose category. For example, reviews of interventions had to have at least one study with random allocation of participants to comparison groups and assessment of at least one clinical outcome.

Six research associates, rigorously calibrated and periodically checked, applied methodological criteria to each item in each issue of the 161 journals.[9] Inter-rater agreement adjusted for chance was 81% (95% confidence interval 79% to 84%) for identifying the purpose of an article and 89% (78% to 99%) for identifying articles that met all methodological criteria.[9]

Search terms

To construct a comprehensive set of possible search terms, we listed indexing terms (for example, subject headings and subheadings, publication types) and text words used to describe systematic reviews (single words or phrases that may appear in titles or abstracts, both in full and in various truncations). We sought further terms from clinicians and librarians, and from published strategies from other groups. We compiled a list of 4862 unique terms and tested them using the Ovid Technologies searching system.

Building search strategies

We determined the sensitivity, specificity, and precision of single term and multiple term search strategies against the criterion standard dataset. Sensitivity for a given topic is defined as the proportion of systematic reviews for that topic that are retrieved; specificity is the proportion of non-systematic reviews not retrieved; and precision is the proportion of retrieved articles that are systematic reviews. The inverse of precision offers a sense of how many records need to be reviewed before one finds a relevant record or hit ("number needed to read") and depends on the proportion of systematic reviews in the database (this proportion does not affect the estimates of sensitivity and specificity).

BMJ: first published as 10.1136/bmj.38336.804167.47 on 24 December 2004. Downloaded from http://www.bmj.com/ on 29 March 2024 by guest. Protected by copyright.

Individual search terms with a sensitivity of more than 50% (to develop strategies that optimised sensitivity) and a specificity of more than 75% (to develop strategies that optimised specificity) for identifying systematic reviews were incorporated into the development of two term strategies; we continued this process to build five term strategies that optimised either sensitivity or specificity. All combinations of terms used the boolean OR—for example, "search OR review". This procedure yielded 782 485 unique strings of one to five terms.

Performing search strategies

We iteratively tested these strategies against a validation dataset—a set of articles from 10 journals in which the highest proportion of systematic reviews are published[7] without including the Cochrane Database of Systematic Reviews. We sorted these strings by sensitivity, by specificity, by the absolute difference of sensitivity and specificity, and by precision. Then, we selected representative strings of terms among those with similar top performance (within 3% of each other). To create generic strategies that were independent of the topic of a review, we excluded terms that would apply exclusively to a given purpose. For example, we excluded "random" or "controlled trials" as these terms would likely apply exclusively to studies of therapy and prevention, and we excluded "sensitivity", which would apply exclusively to studies of diagnosis. This allows a user to add purpose specific terms to a search strategy when constructing a topic specific search. Given that the majority of published systematic reviews relate to therapy and prevention, this decision meant basically excluding terms related to randomised trials from our generic strings. As a secondary approach, we assessed the extent to which our best strings improved when we considered terms specific to the "therapy and prevention" purpose category.

To validate the strategies, we tested candidate strategies in the two validation datasets: database comprising records from 161 journals, one with and one without the Cochrane Database of Systematic Reviews. The ability to find all Cochrane reviews by using the specific term ("journal name") for the Cochrane Database of Systematic Reviews that publishes all of them, and the high prevalence of Cochrane reviews (Cochrane reviews represented 421 of 753 (60%) systematic reviews in the database) justified excluding the Cochrane database from the derivation procedure. The validation database without the Cochrane database tested the validity of the derivation database, whereas the full validation database, including the Cochrane database, offered the best estimates of strategy performance.

Results

The derivation database had 10 446 records, of which 133 (1.3%) were systematic reviews. The full validation database (including the Cochrane Database of Systematic Reviews) included 49 028 records, of which 753 (1.5%) were systematic reviews. Table 1 shows the single term strategies that performed best. This table excludes the term "random:.tw.", a top performer that pertains mostly to systematic reviews of effectiveness (for which the sensitivity in the full validation database was 76% (73% to 79%) and specificity 92.2% (91.9% to 92.5%)). It also excludes the term "cochrane database of systematic reviews.jn." the journal name for the Cochrane Database of Systematic Reviews, by far the most specific single term search strategy. This term pertains solely to Cochrane reviews, with sensitivity in the full validation database of 56% (52% to 60%), specificity of 99.9% (99.9% to 100%), and precision of 96% (94% to 98%). pgwide = "D"

Table 2 shows the top strategies that maximise sensitivity and minimise the absolute difference between sensitivity and specificity (while keeping both ≥90%), a strategy that optimises the balance of sensitivity and specificity. Table 3 shows a strategy with top precision and a set of strategies that identify systematic reviews with greater sensitivity and precision resulting from combining the term "cochrane database of systematic reviews.jn.", a top precision performer, with each of the terms that performed best described in table 1. The combination of any of the strategies using the boolean NOT with publication type terms (such as editorial, comment, or letter) produced negligible improvements in precision and decrements in sensitivity (data not shown). pgwide = "D" pgwide = "D"

Table 4 describes the performance of the most popular strategies available to search for systematic reviews when tested against our full validation database. Compared to the 16-term "high sensitivity" strategy from Centre for Reviews and Dissemination of the University of York, our five term sensitive query has 2.3% (1.2% to 3.4%) greater sensitivity, and our three term balanced query has similar sensitivity and 21.2% (21.16% to 21.24%) greater specificity. The latter strategy performs similarly to the centre's 12-term "high sensitivity and precision" strategy. Although the five term balanced query performs similarly (simi-

Table 1 Best single terms for high sensitivity searches, high specificity searches, and high precision searches for retrieving systematic reviews. Values are percentages (95% confidence intervals)

Term in Ovid format	Sensitivity*	Specificity**	Precision†
search:.tw.‡ §			
Development	52.6 (43.8 to 61.4)	99.2 (99.0 to 99.4)	46.4 (38.2 to 54.7)
Validation without CDSR	52.4 (47.0 to 57.8)	98.0 (97.9 to 98.1)	15.2 (13.1 to 17.3)
Validation	78.9 (75.9 to 81.8)	98.9 (98.8 to 99.0)	51.8 (48.9 to 54.7)
meta-analysis.mp,pt §			
Development	51.1 (42.3 to 59.9)	99.1 (98.9 to 99.3)	42.5 (34.7 to 50.6)
Validation without CDSR	51.5 (46.1 to 56.9)	99.3 (99.2 to 99.3)	32.6 (28.6 to 36.6)
Validation	32.8 (29.5 to 36.2)	99.4 (99.3 to 99.5)	47.1 (42.8 to 51.3)
meta-analysis.pt.¶			
Development	42.1 (33.6 to 51.0)	99.7 (99.6 to 99.8)	67.5 (56.3 to 77.4)
Validation without CDSR	43.0 (37.8 to 48.4)	99.7 (99.7 to 99.8)	50.2 (44.4 to 56.0)
Validation	19.0 (16.2 to 21.8)	99.7 (99.7 to 99.8)	50.2 (44.4 to 56.0)
systematic review.tw.¶			
Development	31.6 (23.8 to 40.2)	99.8 (99.7 to 99.9)	65.6 (52.7 to 77.1)
Validation without CDSR	29.8 (24.9 to 34.7)	99.8 (99.7 to 99.8)	47.4 (40.6 to 54.1)
Validation	19.5 (16.7 to 22.4)	99.9 (99.8 to 99.9)	70.3 (64.1 to 76.5)
Medline.tw.††			
Development	43.6 (35.0 to 52.5)	99.7 (99.6 to 99.8)	64.4 (53.6 to 74.3)
Validation without CDSR	41.9 (36.6 to 47.2)	98.9 (98.8 to 99.0)	21.1 (17.9 to 24.2)
Validation	55.1 (51.6 to 58.7)	99.5 (99.4 to 99.6)	62.9 (59.2 to 66.6)

CDSR=Cochrane Database of Systematic Reviews.
*Development dataset (n=133); validation dataset without CDSR (n=332);validation dataset (n=753).
**Development dataset (n=10 313); validation dataset without CDSR (n=48 258);validation dataset (n=48 275).
†Numbers vary by row.
‡Top term for maximising sensitivity (keeping specificity ≥50%).
§Top terms for maximising specificity (keeping sensitivity ≥50%).
¶Top terms for maximising precision (keeping sensitivity ≥25% and specificity ≥50%).
††Term with superior performance in all categories.

BMJ: first published as 10.1136/bmj.38336.804167.47 on 24 December 2004. Downloaded from http://www.bmj.com/ on 29 March 2024 by guest. Protected by copyright.

Table 2 Best multiple-term strategies maximising sensitivity and minimising the difference between sensitivity and specificity. Values are percentages (95% confidence intervals)

Search strategy in Ovid format	Sensitivity*	Specificity**	Precision†
Top sensitivity strategies‡			
search:.tw. or meta-analysis.mp,pt. or review.pt. or di.xs. or associated.tw.			
Development	100 (97.3 to 100)	63.5 (62.5 to 64.4)	3.41 (2.86 to 4.03)
Validation without CDSR	99.7 (99.1 to 100)	51.1 (50.7 to 51.6)	1.4 (1.2 to 1.5)
Validation	99.9 (99.6 to 100)	52.0 (51.6 to 52.5)	3.14 (2.92 to 3.37)
Top strategy minimising the difference between sensitivity and specificity§			
meta-analysis.mp,pt. or review.pt or search:.tw.			
Development	92.5 (86.6 to 96.3)	93.0 (92.5 to 93.5)	14.6 (12.3 to 17.2)
Validation without CDSR	95.5 (93.3 to 97.7)	89.9 (89.7 to 90.2)	6.1 (5.5 to 6.8)
Validation	98.0 (97.0 to 99.0)	90.8 (90.5 to 91.1)	14.2 (13.3 to 15.2)

CDSR=Cochrane Database of Systematic Reviews.
*Development dataset (n=133); validation dataset without CDSR (n=332); validation dataset (n=753).
**Development dataset (n=10 313); validation dataset without CDSR (n=48 258); validation dataset (n=48 275).
†Numbers vary by row.
‡Keeping specificity ≥50%; adding the Cochrane Database of Systematic Reviews (using Boolean OR) did not improve performance.
§Keeping sensitivity ≥90%.

lar sensitivity with 1.2% (1.16% to 1.24%) greater specificity) to the 71-term PubMed query, we offer three simpler strategies: two with greater sensitivity (a five term sensitive query, difference 9.9% (8.7% to 11%), and a three term balanced query, difference 8% (6.8% to 9.1%)) and one strategy (a three term specific query) with better specificity (difference 2%, 1.96% to 2.04%). With the exception of Hunt and McKibbon strategies, the five term balanced query and the three term specific query offer higher specificity and precision than the other strategies. pgwide = "D"

Table 3 Best multiple-term strategies maximising precision. Values are percentages (95% confidence intervals)

Search strategy in Ovid format	Sensitivity*	Specificity**	Precision†
Top precision performer‡			
Medline.tw. or systematic review.tw. or meta-analysis.pt.			
Development	75.2 (67.0 to 82.3)	99.4 (99.2 to 99.5)	60.2 (52.4 to 67.7)
Validation without CDSR	74.4 (69.7 to 79.1)	98.6 (98.5 to 98.7)	26.3 (23.5 to 29.1)
Validation	71.2 (68.0 to 74.4)	99.2 (99.1 to 99.3)	57.1 (53.9 to 60.3)
Combining most precise term with most sensitive terms (see table 1)§			
Cochrane database of systematic reviews.jn. or search.tw. or meta-analysis.pt. or Medline.tw. or systematic review.tw.	90.2 (88.1 to 92.3)	98.4 (98.3 to 98.5)	46.5 (43.9 to 49.0)
Cochrane database of systematic reviews.jn. or search:.tw.	79.0 (76.1 to 81.9)	98.8 (98.8 to 98.9)	51.8 (48.9 to 54.7)
Cochrane database of systematic reviews.jn. or Medline.tw.	74. 4 (71.3 to 77.5)	99.5 (99.4 to 99.5)	68.8 (65.6 to 72.0)

CDSR=Cochrane Database of Systematic Reviews.
*Development dataset (n=133); validation dataset without CDSR (n=332);validation dataset (n=753).
**Development dataset (n=10 313); validation dataset without CDSR (n=48 258);validation dataset (n=48 275).
†Numbers vary by row.
‡Keeping sensitivity ≥75%.
§Reported for the complete validation dataset. The term for the Cochrane Database of Systematic Reviews (Cochrane database of systematic reviews.jn.) does not apply to the development or to the validation dataset without CDSR since it does not include records from that source.

Table 4 Performance from published strategies to identify systematic reviews in Medline tested in our full validation database. Values are percentages (95% confidence intervals)

Search strategy	Sensitivity*	Specificity**	Precision†
Centre for Reviews and Dissemination			
High sensitivity (16 terms)	97.6 (96.5 to 98.7)	69.6 (69.2 to 70.0)	4.77 (4.43 to 5.11)
Intermediate sensitivity and precision (29 terms)	96.7 (95.4 to 98.0)	79.7 (79.3 to 80.0)	6.91 (6.42 to 7.39)
High sensitivity and precision (12 terms)	95.8 (94.3 to 97.2)	89.7 (89.4 to 90.0)	12.7 (11.8 to 13.5)
Hunt and McKibbon			
Simple query (4 terms)	68.8 (65.5 to 72.1)	99.2 (99.1 to 99.3)	56.7 (53.5 to 59.9)
Sensitive query (8 terms)	73.4 (70.3 to 76.6)	99.1 (99.0 to 99.2)	55.1 (52.0 to 58.2)
Shojania and Bero			
PubMed based query (71 terms)	90.0 (87.9 to 92.2)	97.2 (97.0 to 97.4)	33.2 (31.2 to 35.2)
Hedges (this report)			
Sensitive query (5 terms)‡	99.9 (99.6 to 100)	52.0 (51.6 to 52.5)	3.14 (2.92 to 3.37)
Balanced query, sensitivity>specificity (3 terms)§	98.0 (97.0 to 99.0)	90.8 (90.5 to 91.1)	14.2 (13.3 to 15.2)
Balanced query, specificity>sensitivity (5 terms)¶	90.2 (88.1 to 92.3)	98.4 (98.3 to 98.5)	46.5 (43.9 to 49.0)
Specific query (3 terms)††	71.2 (68.0 to 74.4)	99.2 (99.1 to 99.3)	57.1 (53.9 to 60.3)

*Validation (n=753).
** Validation (n=48 275).
†Numbers vary by row.
‡search:.tw. or meta-analysis.mp,pt. or review.pt. or di.xs. or associated.tw.
§meta-analysis.mp,pt. or review.pt. or search:.tw.
¶Cochrane database of systematic reviews.jn. or search.tw. or meta-analysis.pt. or Medline.tw. or systematic review.tw.
††Medline.tw. or systematic review.tw. or meta-analysis.pt.

Discussion

Our study documents search terms with best sensitivity, best specificity and precision, and smallest difference between sensitivity and specificity for retrieving systematic reviews from Medline. These strategies were derived from a small subset of the database (of 10 high yield journals) that excluded the Cochrane Database of Systematic Reviews and were validated in two databases of 161 high yield journals, one with and one without the Cochrane database. This approach allowed us to identify strategies that adequately detect both Cochrane and non-Cochrane systematic reviews. As a result, we offer Medline hedges with fewer terms and better performance than the best available strategies. These strategies can be implemented using the "limit" top button on the Ovid's search interface for Medline (www.ovid.com) or users can "copy-and-paste" the strategies in table A (see bmj.com) into a PubMed query (www.pubmed.gov), with the chosen strategy combined (by using boolean AND) with appropriate content

BMJ: first published as 10.1136/bmj.38336.804167.47 on 24 December 2004. Downloaded from http://www.bmj.com/ on 29 March 2024 by guest. Protected by copyright.

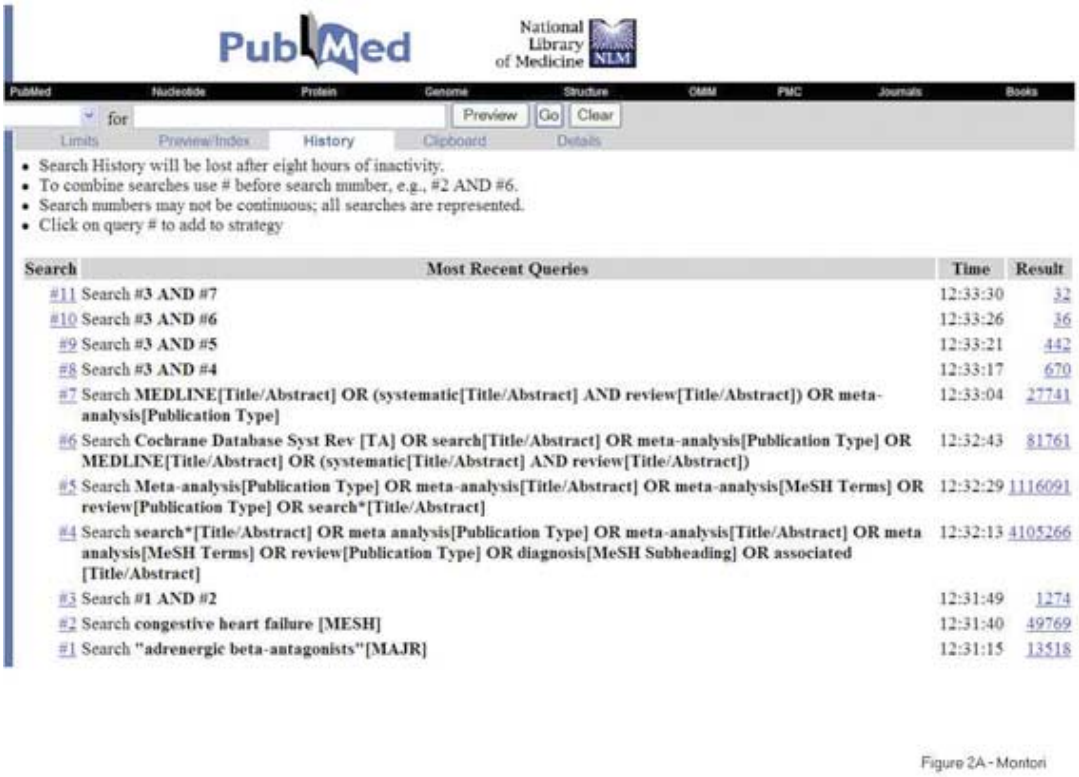


Fig 2 Retrieval on PubMed by combining each of the strategies (see table on bmj.com, using boolean AND, with topic terms to identify systematic reviews about β blockers for congestive heart failure

terms (see figs 2 and 3 for examples). Also, the top strategy minimising the difference between sensitivity and specificity has been incorporated into Skolar (www.skolar.com).

Limitations and strengths

We used a small subset of the database to derive our strategies, and this approach may have overestimated the performance metrics of the identified strategies. We limited our strategy generation to adding terms (boolean OR); strategies with better performance may result from restricting the strategy adding terms using the boolean AND and NOT operators, but the former typically diminished sensitivity when tested in our other work in this database (unpublished data) and the latter resulted in trivial improvements in the current study. Users might observe improvements in the precision of highly sensitive strategies (that is, reduce the number of records reviewed to find one systematic review on the topic) when they combine these (using boolean AND) with terms related to the clinical topic or research topic of interest, but we have not tested this.

Strengths of our approach, unique compared to previous efforts, include the rigorous process followed to generate a criterion standard dataset of systematic reviews, to generate a large number of terms and search strategies, to identify best strategies in a derivation subset, and to validate these in a full validation database of 161 journals. We think that the exclusion of Cochrane reviews (easily retrievable using the Cochrane database term) allowed for the identification of terms that adequately retrieve non-Cochrane reviews. Our approach yielded search strategies with fewer terms and optimal sensitivity, optimal precision, and balance of both sensitivity and precision that exceed the performance of the best published strategies. However, there is room for improvement, and we invite other developers to send us candidate strategies for evaluation against the criterion standard dataset.

Implications for future research and users of medical literature

Ideally, systematic reviews should be indexed using an exact publication type term. Unfortunately, accurate application of this term requires judgment based on assessment of the methods reported in the original articles being indexed. Our process for doing this is highly reproducible, but it has not been shown whether indexers can be trained to do this, and to do so in the time that they have for applying index terms. Pending this, users of the medical literature will need to use hedges, such as those offered here, to identify systematic reviews in Medline.

Currently, clinicians can search systematic reviews within the Cochrane Library, where they can find the Cochrane Database of Systematic Reviews and DARE. If they cannot find a pertinent review, or their interest is other than prevention and treatment, or if they want to conduct a comprehensive search, they can use the strategies presented here to identify systematic reviews in Medline. Quick searches or searching for systematic reviews in topic areas where many are available may be optimal with a high precision strategy or with the strategy that balanced sensitivity and precision. On the other hand, guideline developers and researchers may want to use a highly sensitive strategy. For all, our strategies are most useful when they are pre-programmed into search interfaces, such as the Clinical Queries in PubMed, ready to be combined with topic specific terms.

The Hedges Team includes Angela Eady, Brian Haynes, Susan Marks, Ann McKibbon, Doug Morgan, Cindy Walker-Dilks, Stephen Walter, Stephen Were, Nancy Wilczynski, and Sharon Wong, all at McMaster University Faculty of Health Sciences.

BMJ: first published as 10.1136/bmj.38336.804167.47 on 24 December 2004. Downloaded from http://www.bmj.com/ on 29 March 2024 by guest. Protected by copyright.

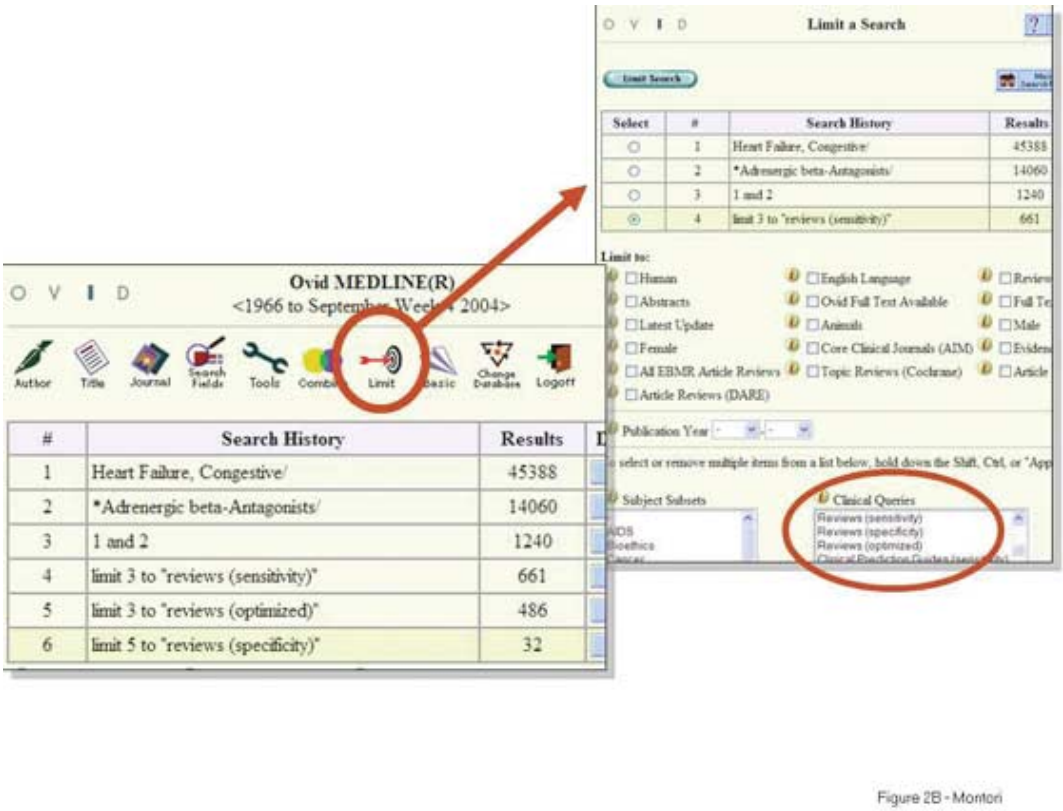


Fig 3 Retrieval on Ovid by entering topic terms (see table on bmj.com) about β blockers for congestive heart failure, then choosing Ovid strategies by clicking on the Limit button on the top of the screen and selecting a strategy for retrieving systematic reviews from the list under Clinical Queries

Contributors: VMM designed the protocol, analysed and interpreted the data, and drafted this report. NLW supervised the research staff and data collection. DM programmed the dataset and analysed the data. RBH planned the study, designed the protocol, and interpreted the data; he will act as guarantor. Members of the Hedges Team collected the data.

Funding: This research was funded by the National Library of Medicine, USA (grant No 1 RO1 LM06866). VMM is a Mayo Foundation Scholar. These funding sources had no additional role.

Competing interests: None declared.

1 Montori VM, Swiontkowski MF, Cook DJ. Methodologic issues in systematic reviews and meta-analyses. *Clin Orthop* 2003;(413):43-54.
2 Dickersin K, Higgins K, Meinert CL. Identification of meta-analyses. The need for standard terminology. *Control Clin Trials* 1990;11:52-66.
3 Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in Medline. *J Am Med Inform Assoc* 1994;1:447-58.
4 Hunt DL, McKibbon KA. Locating and appraising systematic reviews. *Ann Intern Med* 1997;126:532-8.
5 White V, Glanville J, Lefebvre C, Sheldon T. A statistical approach to designing search filters to find systematic reviews: objectivity enhances accuracy. *J Information Sci* 2001;27:357-70.
6 Shojania KG, Bero LA. Taking advantage of the explosion of systematic reviews: an efficient Medline search strategy. *Effective Clin Pract* 2001;4:157-62.
7 Montori VM, Wilczynski NL, Morgan D, Haynes RB. Systematic reviews: a cross-sectional study of location and citation counts. *BMC Med* 2003;1:2.
8 McKibbon KA, Wilczynski NL, Haynes RB. What do evidence-based secondary journals tell us about the publication of clinically important articles in primary health-care journals? *BMC Med* 2004;2:33.
9 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. *Medinfo* 2001;10:390-3.

(Accepted 4 October 2004)

doi 10.1136/bmj.38336.804167.47

What is already known on this topic

Systematic reviews are important for advancing science and evidence based clinical practice, but they may be difficult to retrieve from Medline

What this study adds

Special search strategies retrieved up to 99.9% of systematic reviews, or were able to maximise the proportion of citations retrieved that are systematic reviews

Department of Medicine, Mayo Clinic College of Medicine, Rochester, MN 55905, USA
Victor M Montori *assistant professor*

Health Information Research Unit, Department of Clinical Epidemiology and Biostatistics, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada L8N 3J5
Nancy L Wilczynski *doctoral student*
Douglas Morgan *data analyst*
R Brian Haynes *professor*

Correspondence to: R B Haynes, Department of Clinical Epidemiology and Biostatistics, Room 2C10B Health Sciences Center, McMaster University Faculty of Health Sciences, 1200 Main Street West, Hamilton, ON, Canada L8N 3J5
bhaynes@mcmaster.ca

BMJ: first published as 10.1136/bmj.38336.804167.47 on 24 December 2004. Downloaded from http://www.bmj.com/ on 29 March 2024 by guest. Protected by copyright.