Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed

Karen A Robinson[a] and Kay Dickersin[b]

Objective	To develop, through revision of the Cochrane Collaboration search strategy for OVID-MEDLINE, a highly sensitive search strategy to retrieve reports of controlled trials using PubMed.
Methods	The original highly sensitive Cochrane strategy was revised to take into account additional Medical Subject Headings (MeSH) and other terminology as well as the current unique features of PubMed. We compared the retrieval of the revised strategy with that of the original Cochrane strategy before and after translation of the strategies into PubMed format. Finally, we used a gold standard database of reports of controlled trials identified by electronic and hand search of selected journals to test the revised strategy in PubMed format.
Results	The revised strategy included a search statement modified for increased precision, and added 'Cross-over Studies' as a MeSH term and the term 'latin square' as a text word. Compared to the original Cochrane strategy, the revised strategy identified 53 additional reports of controlled trials accessing MEDLINE through OVID. When the revised strategy and original Cochrane strategy were translated into PubMed format, the revised strategy retrieved 90 reports of controlled trials not identified by the original strategy. Finally, the revised strategy in PubMed format retrieved all of the reports of controlled trials in the gold standard database. Ninety-eight per cent of the gold standard reports of controlled trials were retrieved by Phase 1 of the optimal PubMed search strategy.
Conclusions	Failure to identify all relevant trials for systematic review could result in bias. We developed a highly sensitive search strategy for the retrieval of reports of controlled trials for use with PubMed that retrieves more relevant citations (greater sensitivity) and fewer non-relevant citations (greater precision) than the original Cochrane search strategy.
Keywords	Search strategies, controlled trials, systematic reviews, MEDLINE, PubMed, Cochrane Collaboration
Accepted	5 October 2001

Failure to identify all possibly relevant reports of controlled trials for systematic review could result in bias. As part of the effort to identify all such reports, the Cochrane Collaboration has been using a 'highly sensitive search strategy' developed in 1993 for the retrieval of reports of controlled trials from MEDLINE.[1]

Many reviewers do not have access to fee-based versions of MEDLINE and will increasingly turn to the free web-based searching provided by PubMed as a means of identifying reports of controlled trials. Therefore, we set out to revise the original Cochrane strategy to develop a highly sensitive search strategy for the retrieval of reports of controlled trials from MEDLINE using PubMed.

[a] Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health.

[b] Department of Community Health, Brown University.

Correspondence: Karen A Robinson, Department of Epidemiology, Johns Hopkins University Bloomberg School Public Health, 2024 E. Monument Street, Suite 2–503, Baltimore, MD 1287, USA. E-mail: krobin@jhmi.edu

Methods

The following search services were used: OVID-MEDLINE (Version 4.0), a non-graphical vendor software, and PubMed.

Unless otherwise stated, searches included publication years 1966–1998.

Three stages were completed: Stages one and two were completed in 1998 and stage three was completed in 2000.

Stage one: Revision of original Cochrane strategy

Because we hypothesized that use of a longer word stem would result in greater precision, the search statement 'clin$ adj25 trial$' from the original Cochrane search strategy was changed to 'clinic$ adj25 trial$' in the revised strategy. Next tested was the addition of MeSH term 'Cross-Over Studies' to the original Cochrane search strategy. The third revision considered was the addition of 'latin[tw] AND square[tw]'.

Revisions were made to the original Cochrane search strategy and the retrieval of the original and revised strategies were compared. The number of uniquely retrieved citations was noted and these were reviewed by a trained searcher to identify reports of controlled trials.

Stage two: Translation of original and revised strategy into PubMed format

The original and revised strategies were translated into PubMed format and, as in stage one, the difference in the total number of citations retrieved was noted and a trained searcher reviewed the unique citations to determine the number of reports of controlled trials retrieved.

Stage three: Testing of revised strategy in PubMed format

A gold standard database of reports of controlled trials was created using CENTRAL, the Cochrane Collaboration database that contains electronic and hand searching results. The reports of controlled trials in CENTRAL (Issue 4, 2000) from 11 general US medical journals for 1998 were downloaded into a database (n = 308). PubMed was then searched using the revised strategy and results were downloaded for comparison to the gold standard database.

Phase	Set	Search statement
Phase 1	001	randomized controlled trial.pt.
	002	controlled clinical trial.pt.
	003	Randomized Controlled Trials/
	004	Random Allocation/
	005	Double-Blind Method/
	006	Single-Blind Method/
	007	or/1–6
	008	Animal/ not Human/
	009	7 not 8
Phase 2	010	clinical trial.pt.
	011	explode Clinical Trials/
	012	(clinic$ adj25 trial$).tw.
	013	((singl$ or doubl$ or trebl$ or tripl$) adj (mask$ or blind$)).tw.
	014	Placebos/
	015	placebo$.tw.
	016	random$.tw.
	017	Research Design/
	018	(latin adj square).tw.
	019	or/10–18
	020	19 not 8
	021	20 not 9
Phase 3	022	Comparative Study/
	023	explode Evaluation Studies/
	024	Follow-Up Studies/
	025	Prospective Studies/
	026	(control$ or prospectiv$ or volunteer$).tw.
	027	Cross-Over Studies/
	028	or/22–27
	029	28 not 8
	030	29 not (9 or 21)
All Phases	031	9 or 21 or 30

Figure 1 Revised search strategy for retrieval of reports of controlled trials from OVID-MEDLINE. Revised or added search studies are underlined

Results

Stage one: Revision of original Cochrane strategy

When we truncated 'clinical' with a longer word stem, 29 fewer citations were retrieved, none of which was classified as a report of a controlled trial. This more precise longer word stem was used in Phase 2 of the revised strategy. Separately adding the term 'latin[tw] AND square[tw]' resulted in 153 additional citations, of which 119 (78%) were reports of controlled trials. This term was added to Phase 2 of the revised search strategy. Separately adding the term 'Cross-Over Studies[mh]' to the strategy resulted in 227 additional citations but this term was low in precision (4% identified as reports of controlled trials) and so was added to the least precise Phase 3 of the search strategy. Figure 1 shows the revised strategy for OVID-MEDLINE.

The retrieval of the revised strategy was then compared to the retrieval of the original Cochrane search strategy. There were no changes made to Phase 1 and thus no unique citations retrieved by either strategy. For Phase 2, the revised strategy retrieved 119 additional reports of controlled trials and did not retrieve

29 non-relevant citations retrieved by the original Cochrane strategy. In Phase 3, the original Cochrane strategy retrieved 74 reports of controlled trials not retrieved by Phase 3 of the revised strategy, however all of these reports were already retrieved by Phase 2 of the revised strategy. Also in Phase 3, the revised strategy identified 10 (14% of unique citations) additional reports of controlled trials. Using all phases, the revised strategy for OVID-MEDLINE retrieved 53 (44% of unique citations) additional reports and no reports of controlled trials were missed.

Stage two: Translation of revised strategy into PubMed format

The original and revised search strategies in PubMed format are shown in Figure 2. In PubMed, the revised strategy retrieved all citations retrieved by the original Cochrane strategy. The revised strategy identified 161 additional citations, 90 (55%) of which were classified as reports of controlled trials (73 additional reports of controlled trials in Phase 2 and 17 in Phase 3).
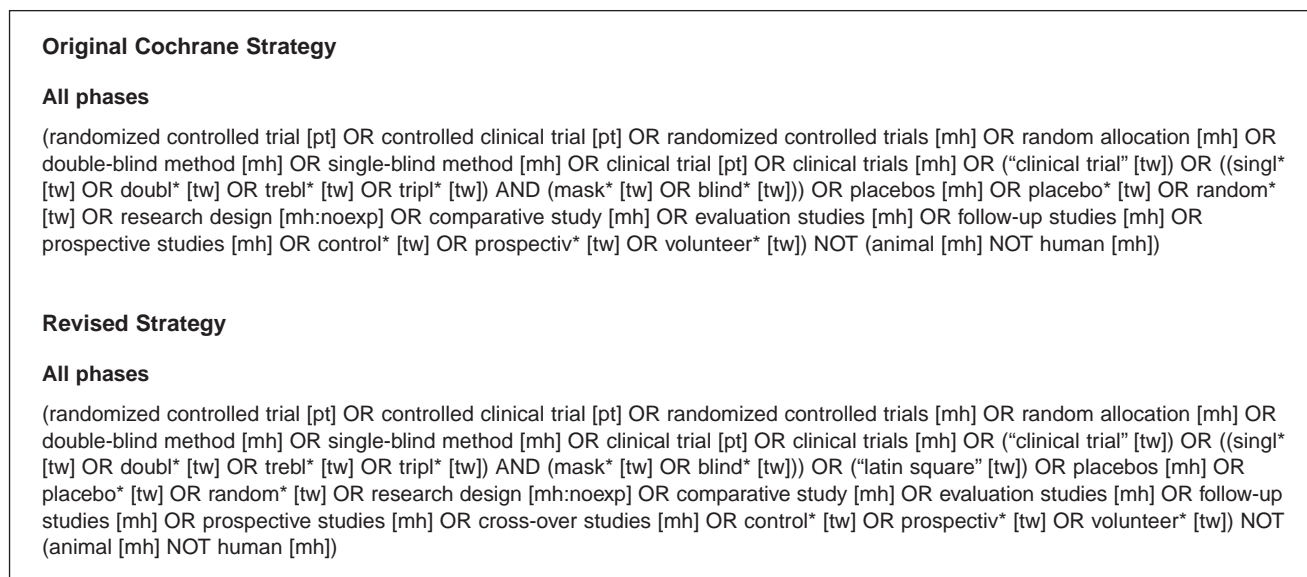

Original Cochrane Strategy

All phases

(randomized controlled trial [pt] OR controlled clinical trial [pt] OR randomized controlled trials [mh] OR random allocation [mh] OR double-blind method [mh] OR single-blind method [mh] OR clinical trial [pt] OR clinical trials [mh] OR ("clinical trial" [tw]) OR ((singl* [tw] OR doubl* [tw] OR trebl* [tw] OR tripl* [tw]) AND (mask* [tw] OR blind* [tw])) OR placebos [mh] OR placebo* [tw] OR random* [tw] OR research design [mh:noexp] OR comparative study [mh] OR evaluation studies [mh] OR follow-up studies [mh] OR prospective studies [mh] OR control* [tw] OR prospectiv* [tw] OR volunteer* [tw]) NOT (animal [mh] NOT human [mh])

Revised Strategy

All phases

(randomized controlled trial [pt] OR controlled clinical trial [pt] OR randomized controlled trials [mh] OR random allocation [mh] OR double-blind method [mh] OR single-blind method [mh] OR clinical trial [pt] OR clinical trials [mh] OR ("clinical trial" [tw]) OR ((singl* [tw] OR doubl* [tw] OR trebl* [tw] OR tripl* [tw]) AND (mask* [tw] OR blind* [tw])) OR ("latin square" [tw]) OR placebos [mh] OR placebo* [tw] OR random* [tw] OR research design [mh:noexp] OR comparative study [mh] OR evaluation studies [mh] OR follow-up studies [mh] OR prospective studies [mh] OR cross-over studies [mh] OR control* [tw] OR prospectiv* [tw] OR volunteer* [tw]) NOT (animal [mh] NOT human [mh])

Figure 2 Strategies in PubMed format

Stage three: Testing of revised strategy in PubMed format

The revised strategy in PubMed format retrieved all of the citations in the gold standard database. Phase 1 alone retrieved 301 (98%) and adding Phase 2 retrieved 307 of the 308 reports of controlled trials in the gold standard database.

Discussion

To avoid bias, it is essential that a systematic review identify and consider all possibly relevant trials. An integral component of a comprehensive search plan is an effective MEDLINE search. Compared to the original Cochrane strategy, the revised strategy retrieved more reports of controlled trials (increased sensitivity), thereby minimizing the potential for bias caused by the omission of possibly relevant trials. Because the revised strategy was also more precise, its use limits the amount of non-relevant citations retrieved and thus lowers the amount of work in reviewing citations for eligibility. The revised strategy in PubMed format is an essential tool for reviewers.

Since its initial description,[2] this highly sensitive strategy for the retrieval of reports of controlled trials using PubMed has been used as part of comprehensive search plans for a number of systematic reviews in a variety of fields.[3–7] In each of these reviews, the revised strategy in PubMed format was combined with a topic-specific strategy designed to balance sensitivity and precision.

Additional testing of the revised strategy in PubMed format may be warranted. For instance, retrieval may differ for citations published in other years, such as those published in 1975 or earlier when MEDLINE did not include abstracts. Also, the gold standard database was relatively small and limited to US general medical journals for practical reasons. The revised strategy could have different sensitivity and precision for journals published outside of the US or in specific health fields. Citations from such journals may include different terms or be indexed with MeSH differently. Finally, CENTRAL was not analysed in terms of validity and completeness before using it as the gold standard database. Further testing with a verified gold standard database covering additional journals and years, as well as an examination of any additional citations retrieved by the revised strategy, may be useful.

To continue to be an effective and efficient strategy, the revised strategy should be examined periodically to take into account new features available on PubMed, as well as developments in indexing by the National Library of Medicine. For instance, additional guidance on the use of the MeSH term 'Cross-Over Studies' has been recently published.[8] In addition, the consequences of current features of PubMed, such as the inability to use modifiers, need further examination.

Acknowledgement

KEY MESSAGES

• Systematic reviews, meta-analyses and decision analyses are evidence-based methods increasingly used and reported in the medical literature; failure to identify all possibly relevant trials for such research may lead to bias.

• The highly sensitive search strategy for controlled trials used by the Cochrane Collaboration was revised and then translated for use in PubMed.

• The new strategy provides researchers and others completing evidence syntheses with an effective and efficient method of searching PubMed for controlled trials.

References

[1] Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *Br Med J* 1994;**309(6964):**1286–91.

[2] Robinson KA, Hinegardner PG, Lansing P. Development of an optimal search strategy for the retrieval of controlled trials using PubMed [abstract]. Proceedings of the 6th International Cochrane Colloquium; 1998 Oct 22–26. *Abstr Workshops Sci Sess Int Cochrane Coll* 1998;**6:**85.

[3] Lehmann HP, Andrews JA, Robinson KA, Holloway VL, Goodman SG. *Evidence Report on Management of Acne*. AHRQ Publication No. 01-E019. Rockville, MD: Agency for Healthcare Research and Quality, 2001.

[4] Schein OD, Friedman DS, Fleisher LA *et al. Evidence Report on Anesthesia Management During Cataract Surgery*. AHRQ Publication No. 00-E015. Rockville, MD: Agency for Healthcare Research and Quality, 2000.

[5] Jampel HD, Lubomski LH, Friedman DS *et al. Evidence Report on Treatment of Coexisting Cataract and Glaucoma*. AHRQ Publication No. 01-E050. Rockville, MD: Agency for Healthcare Research and Quality, 2001.

[6] Friedman DS, Bass EB, Lubomski LH *et al.* Synthesis of the literature on the effectiveness of regional anesthesia for cataract surgery. *Ophthalmology* 2001;**108:**519–29.

[7] Friedman DS, Bass EB, Lubomski LH *et al.* The methodologic quality of clinical trials on regional anesthesia for cataract surgery. *Ophthalmology* 2001;**108:**530–41.

[8] Lefebvre C, Clarke MJ. Identifying randomized trials. In: Egger M, Davey Smith G, Altman DG (eds). *Systematic Reviews in Health Care: Meta-Analysis in Context, 2nd Edn*. London: BMJ Books, 2001.