Journal of Clinical Epidemiology 78 (2016) 108–115

Journal of Clinical Epidemiology

Complementary approaches to searching MEDLINE may be sufficient for updating systematic reviews

Margaret Sampson[a,*], Berry de Bruijn[b], Christine Urquhart[c], Kaveh Shojania[d]

[a]*Library and Media Services, Children's Hospital of Eastern Ontario, 401 Smyth Road, Ottawa, Ontario K1H 8L1, Canada*

[b]*National Research Council—Information and Communications Technologies Portfolio (NRC-ICT), 1200 Montreal Rd, bldg M-50, Ottawa, Ontario K1A 0R6, Canada*

[c]*Department of Information Studies, Sunnybrook Health Sciences Centre, Aberystwyth University, Llanbadarn Campus, Aberystwyth, Wales SY23 3AL, United Kingdom*

[d]*Centre for Quality Improvement and Patient Safety, Room H468, 2075 Bayview Avenue Toronto, Ontario M4N 3M5, Canada*

Accepted 7 March 2016; Published online 11 March 2016

Abstract

Objectives: To maximize the proportion of relevant studies identified for inclusion in systematic reviews (recall), complex time-consuming Boolean searches across multiple databases are common. Although MEDLINE provides excellent coverage of health science evidence, it has proved challenging to achieve high levels of recall through Boolean searches alone.

Study Design and Setting: Recall of one Boolean search method, the clinical query (CQ), combined with a ranking method, support vector machine (SVM), or PubMed-related articles, was tested against a gold standard of studies added to 6 updated Cochrane reviews and 10 Agency for Healthcare Research and Quality (AHRQ) evidence reviews. For the AHRQ sample, precision and temporal stability were examined for each method.

Results: Recall of new studies was 0.69 for the CQ, 0.66 for related articles, 0.50 for SVM, 0.91 for the combination of CQ and related articles, and 0.89 for the combination of CQ and SVM. Precision was 0.11 for CQ and related articles combined, and 0.11 for CQ and SVM combined. Related articles showed least stability over time.

Conclusions: The complementary combination of a Boolean search strategy and a ranking strategy appears to provide a robust method for identifying relevant studies in MEDLINE.

Keywords: Information retrieval; Systematic reviews; Support vector machine; Clinical query; PubMed similar articles; Searches; Updating; MEDLINE

1. Introduction

Systematic review searches need to have high recall. Mechanisms to achieve this usually include expansive Boolean searches of multiple databases. This approach leads to long development times for the searches [1], necessitates accessing multiple sources that may not be accessible in some institutions [2], and entails time-consuming removal of duplicate records for articles indexed in more than one of the databases searched [3].

MEDLINE gives excellent coverage of most biomedical topics, in particular, intervention studies. However, in 1994, a landmark article by Dickersin et al. [4] established that only about half the studies included in systematic reviews were identified through MEDLINE. Recent research has demonstrated that a much higher percentage is present in MEDLINE, but sometimes their retrieval is problematic [5–8].

Successful retrieval through a Boolean search is operator dependent, with search performance being influenced by skill of the indexer and the searcher. The search of multiple databases can therefore be helpful. The target records may be indexed differently in the second, third, or subsequent source searched, increasing the probability of a match between the search terms entered and the indexing of the additional records. A text search, querying terms appearing in the title or abstract, may help improve retrieval, as may the use of indexing terms that are broader (less specific) than, or related to, the single best indexing term. These tactics on the part of the searcher lead

Conflict of interest: None.

Funding: Some of this work was funded by the US Agency for Healthcare Research and Quality, Department of Health and Human Services (contract no. 290-02-0021).

* Corresponding author. Tel.: 613-737-7600.

E-mail address: msampson@cheo.on.ca (M. Sampson).

http://dx.doi.org/10.1016/j.jclinepi.2016.03.004

0895-4356/

What is new?

Key findings

- Searches using known relevant studies and the similar articles feature of PubMed will identify and rank additional articles of potential relevance.
- For a given question, if the Boolean search has low recall, the ranking search tends to have higher recall, and vice versa. The two approaches complement each other.
- The precision of this complementary paired method appears better than the precision of exhaustive Boolean searches.

What this adds to what was known?

- The paired approach performed well regardless of which of two tested similarity searches were used. It is the use of independent retrieval methods that is important.

What is the implication and what should change now?

- Using the simple and universally available PubMed similar feature makes this paired approach practical for most systematic review teams.
- If the paired complementary approach is used, the recall may be sufficient to consider using only MEDLINE.

to large retrievals with low precision [9]. These resulting problems may be particularly challenging for complex or newly emerging interventions with highly variant terminology, where alternatives to traditional Boolean searches have been sought [10].

Most recent systematic review information retrieval research has focused on text mining approaches [11–14]. These approaches often harvest an intentionally overinclusive set of records and then use machine learning, similarity ranking and other techniques to refine the set to identify the material most likely to be relevant, thereby reducing human screening effort. These methods show promise, but are not yet widely available to reviewers.

We examined one method, support vector machine (SVM), and compared it with a simple and readily available method based on the PubMed similar articles feature. We call this method *related articles* to distinguish the method from the similar articles feature itself. We paired both with a focused Boolean search within MEDLINE. We tested this approach in an updating context where studies included in the original review comprise the reference standard for SVM and seed articles for the PubMed-related articles search. We therefore sought to determine if a focused Boolean search paired with one of the search methods that does not depend on operator skill could provide consistently complete retrieval of relevant new studies.

Comparison of a number of searches, including two tested here, has been previously reported [15]. In this current article, two of the most successful methods in a larger sample of 72 journal-published systematic reviews, clinical query (CQ), and PubMed-related articles are tested in a cohort of six updated Cochrane reviews, as well as in a previously unreported sample of 10 Agency for Healthcare Research and Quality (AHRQ) evidence reports. The Cochrane reviews provide a true gold standard as updates were made by the review team based on evidence identified through comprehensive searches; however, the new relevant studies proved fairly easy to find. The replication in the AHRQ cohort, of more complex interventions, provided a means to validate the generalizability of the approach [16]. All records were assessed for eligibility by two reviewers, and this complete screening allowed the precision of the methods to be calculated for the first time.

Other research [17,18] suggests that searches using the PubMed similar articles feature are effective in increasing recall of relevant items for reviews or more general clinical searching when combined with a Boolean-type search of MEDLINE. We tested an additional search method, SVM, in the Cochrane and AHRQ samples to permit comparison with our PubMed-related articles search and assess whether the complementary effect generalized beyond the PubMed similar articles method.

The aim of this article was to test whether the combined approach of a focused Boolean search paired with a second search using the similar articles feature of PubMed or SVM can yield high recall with reasonable precision.

2. Methods

2.1. Formation of the study cohorts

This analysis uses a data set created for an updating study sponsored by AHRQ [16]. Methods for the selection of the cohorts, search approaches tested, and rigorous mechanisms to screen search results for relevance have been previously reported, along with the criteria to determine if a review was in need of update [16].

Briefly, Cochrane reviews were identified through a search of the ACP Journal Club database (Ovid) using the strategy:

1. review$.ti. 2. meta-analy$.mp. 3. data sources.ab. 4. (search$ or MEDLINE®).ab. 5. or/1-4 6. limit 5 to articles with commentary.

AHRQ reports were identified through the PubMed query "Evid Rep Technol Assess (Summ)"[Journal:__ jrid21544]. Screening was undertaken in two phases with two reviewers reaching consensus on eligibility. Screening continued until the predetermined sample size was reached.