ELSEVIER

Journal of Clinical Epidemiology 62 (2009) 944–952

Journal of Clinical Epidemiology

ORIGINAL ARTICLES

An evidence-based practice guideline for the peer review of electronic search strategies

Margaret Sampson[a,b,*], Jessie McGowan[a,c,d,e], Elise Cogo[b,c], Jeremy Grimshaw[c,d,e], David Moher[d,e], Carol Lefebvre[f]

[a]*Department of Information Studies, University of Wales, Aberystwyth, Wales, UK*
[b]*Children's Hospital of Eastern Ontario Research Institute, Ottawa, Canada*
[c]*Institute of Population Health, University of Ottawa, Ottawa, Canada*
[d]*Faculty of Medicine, University of Ottawa, Ottawa, Canada*
[e]*Ottawa Health Research Institute, Ottawa, Canada*
[f]*UK Cochrane Centre, National Institute for Health Research, Oxford, UK*

Accepted 13 October 2008

Abstract

Objective: Complex and highly sensitive electronic literature search strategies are required for systematic reviews; however, no guidelines exist for their peer review. Poor searches may fail to identify existing evidence because of inadequate recall (sensitivity) or increase the resource requirements of reviews as a result of inadequate precision. Our objective was to create an annotated checklist for electronic search strategy peer review.

Study Design: A systematic review of the library and information retrieval literature for important elements in electronic search strategies was conducted, along with a survey of individuals experienced in systematic review searching.

Results: Six elements with a strong consensus as to their importance in peer review were accurate translation of the research question into search concepts, correct choice of Boolean operators and of line numbers, adequate translation of the search strategy for each database, inclusion of relevant subject headings, and absence of spelling errors. Seven additional elements had partial support and are included in this guideline.

Conclusion: This evidence-based guideline facilitates the improvement of search quality through peer review, and thus the improvement in quality of systematic reviews. It is relevant for librarians/information specialists, journal editors, developers of knowledge translation tools, research organizations, and funding bodies.

Keywords: Evidence-based practice; Information retrieval; Peer review; Practice guideline; Search strategies; Systematic reviews

1. Introduction

The quality of systematic reviews and health technology assessments depends on many factors. A key factor is the evidence base, that is the literature and other information on which the analysis is based. A sound evidence base is equally important in related contexts, such as clinical practice guidelines, formal decision analysis, and most health economics studies, and the knowledge translation efforts surrounding this work. Performing a high-quality electronic search of information resources is an essential contribution toward ensuring accuracy and completeness of the evidence base used in these reports. Furthermore, search quality has resource implications for the conduct of a review, because the number of records retrieved and screened is often very large. The aim of this research was to develop an evidence-based guideline in the form of an annotated checklist for peer review of electronic search strategies. Given the inconsistent performance of the peer-review process in the contexts of journal publication [1] and grant funding [2], any new development in peer review should be evidence-based.

Checklists, scales, and instruments (henceforth, instruments) for validating some aspects of the search-reporting methods of the systematic review process have been developed, and some address aspects of the overall search plan. We have reviewed these elsewhere [3]. However, a validated process for evaluating the quality and completeness of the electronic search strategies used to identify the largest portion of the evidence base for systematic reviews, and

* Corresponding author. Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Road, Room 226, Ottawa, Ontario, Canada. Tel.: +613-737-7600 ext. 2207; fax: +613-738-4806.

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

0895-4356/09/$ – see front matter
doi: 10.1016/j.jclinepi.2008.10.012

Key points

- Errors in the electronic search strategy have been demonstrated to reduce the effectiveness of electronic search strategies used in systematic reviews and health technology assessment reports.
- Evidence supported 13 aspects of electronic search strategies that have an impact on search performance.
- An evidence-based peer-review process has been developed to evaluate electronic search strategies for use in systematic reviews and health technology assessment reports.
- The electronic search strategy for systematic reviews and health technology assessment reports should be peer reviewed before the evidence base is created and reviewed.
- The evidence-based peer review of electronic search strategies should be conducted by librarians.

related evidence-based products, such as health technology assessments and clinical practice guidelines, does not exist. The lack of such a process paired with a demonstrable level of error in reported searches [4] leaves this type of research open to debate over the quality of evidence on which they are based.

Given the absence of any adequate instruments for peer-reviewing electronic search strategies used for systematic reviews, a context where high recall must be balanced with reasonable precision, we sought evidence regarding various ''elements'' of the electronic search strategy. We define elements as components of the electronic search strategy that could have a positive or negative impact on the performance metrics of recall and precision. Research evidence on the impact of problems with these elements on search performance would be the most compelling evidence. We expected some gaps in the evidence and so we considered reports from the literature on the prevalence of errors of that type, as well as theoretical explanations of their impact. We sought this evidence through a systematic review of the literature. Finally, we sought expert opinion through a survey.

2. Objectives

We had three objectives: the identification of elements associated with the accuracy and completeness of the systematic review evidence base, identified by electronic search strategies; the identification of the impact of problems with any of these elements on the resulting evidence base; and the development of an evidence-based guideline for peer review of the electronic search strategy.

3. Methods

3.1. Systematic review methods

A systematic review of the literature was conducted to address (1) whether instruments exist that evaluate or validate the quality of literature searches in any discipline and (2) which elements of electronic search strategies have demonstrable impact on search performance. These articles need to report performance indicators or measures (such as recall, relevance, and so forth).

The search strategy was developed initially in the Library and Information Science Abstracts (LISA) database by one author (E.C.), and reviewed and revised by two others (M.S., J.M.). The search strategy from LISA was then adapted for the other databases searched. We sought material published from 1980 onward, reflecting the rise in widespread use of electronic searching. Languages were limited to those understood by the review team: English, French, Italian, and Spanish.

The databases searched were LISA (CSA interface) 1969–May 2005; Cochrane Methodology Register and Cochrane Methodology Reviews (completed reviews only; The Cochrane Library 2005, Issue 2, Wiley Interface); MEDLINE (Ovid interface) 1966–June Week 1, 2005; PsycINFO (Ovid interface) 1806–June Week 2, 2005; CINAHL (Ovid interface) 1982–June Week 2, 2005; HealthSTAR (Ovid interface) 1987–May 2005; HAPI (Health and Psychosocial Instruments—Ovid interface) 1985–March 2005. The search strategies used have been previously published [5].

Gray literature was identified through correspondence with librarians and other systematic review experts, by searching the Cochrane Methodology Register for conference abstracts, and by searching our personal databases of information retrieval research accrued over the years, including material such as presentations, dissertations, and prepublication manuscripts.

After screening and data extraction of this material, targeted checking of the reference lists of included studies was conducted to capture literature on those elements for which we had found fewer than five studies. These elements were proximity operators, the Boolean operator NOT, organization, irrelevant terms, truncation errors, wrong line number specified, explosion without the existence of any narrower terms, searching additional fields, redundancy without rationale, and combining index and free-text terms on the same line. Redundancy would be a part of the search that could not yield new material as it was functionally equivalent to or narrower than another part of the search. An example of a rationale would be that both a narrower term and an exploded broader term were included so that readers unfamiliar with subject heading hierarchy would know that both terms factored into the search.

3.1.1. Eligibility criteria

The search context was not restricted to systematic reviews, or health technology assessments. All research

on electronic database searching was eligible, including research outside of health care. At the initial review of the bibliographic record (title, abstract, and indexing), we considered only whether the record appeared relevant to peer review of electronic search strategies. For those that appeared relevant, we retrieved and assessed the full report, retaining those articles that presented an evaluation tool for search strategies, or presented primary evidence on the impact of searching techniques on search results, or presented a theoretical discussion on the impact of searching techniques on search results.

3.1.2. Screening and data extraction

Titles and abstracts were screened and the consensus of two reviewers was needed to exclude a study. Articles appearing potentially relevant were retrieved and two reviewers assessed each of the full reports, arriving at consensus regarding eligibility. Reviewers were E.C., M.S., and J.M.

Data extraction was performed by one reviewer (E.C.) and verified by a second reviewer (M.S.). For each included study, we determined which elements of the electronic search were addressed. These initial elements of interest were refined and several additional elements emerged in the course of the review. The final elements considered are listed in Table 1. A summary of the evidence relating to each element is presented elsewhere [5].

3.2. Survey methods

Because of the anticipated scarcity of research evidence in some elements of the electronic search, a Web-based survey was conducted. The aim of the survey was to gather expert opinion regarding both the impact of various search elements on the search results and the importance of each element in peer review of electronic search strategies. The Ottawa Hospital Research Ethics Board approved the survey and participants formally consented to participation.

The survey was conducted after the systematic review was completed, so that elements identified by the review could be included. Added elements were use of additional fields, problems with the use of limits, and conceptualization of the search, which had originally been treated as an aspect of search organization. The survey is reproduced elsewhere [5].

The survey was conducted using SurveyMonkey, a Web-based application (http://www.surveymonkey.com/). The survey was piloted on a convenience sample of 10 searchers, all of whom were professionally qualified as librarians and some were also Cochrane Trials Search Co-ordinators. All were known to the investigators. Recruitment for the main survey was from eight discussion lists in the field of systematic reviews or medical librarianship. Reminders were sent after 1 week to those who had expressed interest but who had neither completed the survey nor withdrawn by declining consent to participate on the first screen of the survey.

For each of the elements considered, the survey respondents were asked five questions: the extent of potential negative impact on recall, potential negative impact on precision, whether problems with the use of the element might indicate unfamiliarity on the part of the searcher with aspects of searching, how important it would be to consider that element in the peer review of a search strategy, and the

Table 1
Type, amount and main impact of research evidence found for each element examined

Element[a]	N	Type of Evidence(*n*)			Main impact[b]
		R	T	F	
Conceptualization	23	10	12	4	Recall
Logical operator errors	40	22	19	10	Recall, precision
Boolean AND or OR	22	9	14	3	Recall, precision
Proximity operators	8	2	4	2	Recall, precision
Boolean NOT	6	3	3	—	Recall, precision
Line number errors	0	—	—	—	—
Search not adapted for each database	27	17	14	1	Recall
Subject heading missing	81	56	31	7	Recall, precision
Spelling errors	13	11	7	9	Recall, precision
Missed free-text terms	71	51	29	10	Recall, precision
Limits missed or used inappropriately	31	14	15	2	Recall, precision
Irrelevant subject headings	3	2	1	1	Precision
Truncation errors	2	1	—	1	Precision
Irrelevant free-text terms	4	1	2	1	Precision
Redundancy without rationale	2	—	2	—	Recall
Failure to use additional fields	13	6	6	1	Recall, precision
Organization of the search	6	1	6	—	Recall, time-cost
Subject headings and free text combined on a line	0	—	—	—	—
Subject heading exploded when no narrower terms	0	—	—	—	—

Abbreviations: R, research; T, theory; F, frequency of error.

[a] Elements are organized according to final ranking of importance (tier 1, 2, or 3—see Table 2), not volume of evidence.

[b] Main impacts only are reported. Other impacts include specificity, relevance and importance in peer review.

respondent's impression of the level of research evidence supporting the element. Response choices varied by question; a four-point scale from *Nil* to *Large* was used for impact questions, a five-point scale from *Strongly Disagree* to *Strongly Agree* was used for the unfamiliarity and importance questions, and a seven-point nominal scale was used for level of evidence.

At the end of the survey, respondents had the opportunity to nominate additional elements and provide an assessment of their impact and importance in peer review.

3.2.1. Survey analysis

Elements were ranked according to the median score for each question. For instance, an element with a median score indicating a larger negative impact on recall would be ranked higher (more important) than an element with a lower median on that question. For elements with equal medians, those with the smaller interquartile range (indicating greater consensus) were ranked higher. Where there was still a tie, the elements were considered equal. Elements were also ranked by the percentage of respondents who felt that there was some support from research evidence for the element.

Elements were classified into three tiers of importance. Tier 1 elements were those that received only first- or second-place rankings on importance in peer review, recall, or precision. Tier 2 elements had some, but not consistent first- or second-place rankings. Tier 3 elements received no first- or second-place rankings on these dimensions.

3.3. Synthesis of evidence and formulation of recommendations

Findings from all subject areas of information retrieval research were considered for each element. The salient features that would provide guidance to peer review were extracted and summarized. For instance, for subject headings, considered a Tier 1 (i.e., very important) element, those aspects of usage that would improve recall were summarized.

The evidence summaries of the elements have undergone peer review and have been published in a technical report [5]. These summaries form the basis for the peer-review guideline presented here. Survey ratings were used to determine the importance of the elements in the peer-review process and Tier 1 and 2 elements are included in this guideline.

4. Results

4.1. Systematic review

Our electronic database search strategies identified 9,155 records. After broad screening of the titles, abstracts, and keywords, 256 appeared potentially eligible and were obtained for data abstraction. Of these, 113 articles were

Table 2
Survey respondents' ranking of importance or degree of impact on critical dimensions of search performance for all search elements examined in a survey of search practitioners in the area of health evidence summaries

Element	Tier[a]	Important in peer review	Rank on each dimension			Research evidence[b]
			Recall	Precision	Unfamiliarity	
Conceptualization	1	1	1	1	1	6
Logical operator errors	1	1	1	1	1	12
Line number errors	1	1	1	1	5	18
Search not adapted for each database	1	1	2	2	5	1
Subject heading missing	1	2	1	2	1	2
Spelling errors	1	2	1	2	7	15
Free-text missing terms	2	2	2	3	3	3
Irrelevant limits	2	2	2	4	3	8
Spelling variants missing	2	2	2	5	2	4
Irrelevant subject headings	2	2	4	1	2	9
Truncation errors	2	3	2	2	3	14
Relevant limits missed	2	3	3	2	3	5
Irrelevant free-text terms	2	3	6	2	3	13
Redundancies	3	3	8	7	5	17
Additional fields	3	4	5	5	5	7
Organization of search	3	4	3	6	7	10
Subject headings and free-text in a single search statement	3	4	6	7	6	11
Irrelevant explosions	3	5	7	8	4	16

[a] Tier 1 elements received only first- or second-place rankings on importance in peer review, recall or precision. Tier 2 had some, but not consistent first- or second-place rankings. Tier 3 elements received no first- or second-place rankings on these dimensions.

[b] The 18 elements were ranked from high (1) to low (18) based on the percentage of respondents who felt that there was at least some research evidence supporting the importance of the element.

assessed as eligible for some aspect of the systematic review process (Fig. 1). The remaining articles were excluded either because they did not meet the inclusion criteria ($n = 139$) or because they could not be obtained ($n = 4$).

4.2. Existing scales and instruments

Twenty-six articles discussed aspects that could be used as search assessment tools. A number of the articles addressed the conduct or reporting of the broader search plan used to form the evidence base for systematic reviews, or clinical practice guidelines, but did not address the quality of the electronic search strategies themselves. These are described elsewhere [3]. A number of articles described additional instruments that assessed student learning [5].

4.3. Research evidence pertaining to the search elements

The number of articles and the type of evidence identified from the literature for each element is presented in Table 1. The amount of research evidence varied considerably, with more than 70 articles addressing the consequences of missed search terms, and little or no identifiable research evidence on several elements, such as redundancy without rationale, subject headings exploded when no narrower term exists, index and free-text terms combined on a line, line number errors, or truncation errors.

The impacts considered were recall (or sensitivity), precision (or positive predictive value), specificity, relevance, cost or time, and any discussion of importance in peer reviewing. Most available evidence related to recall.

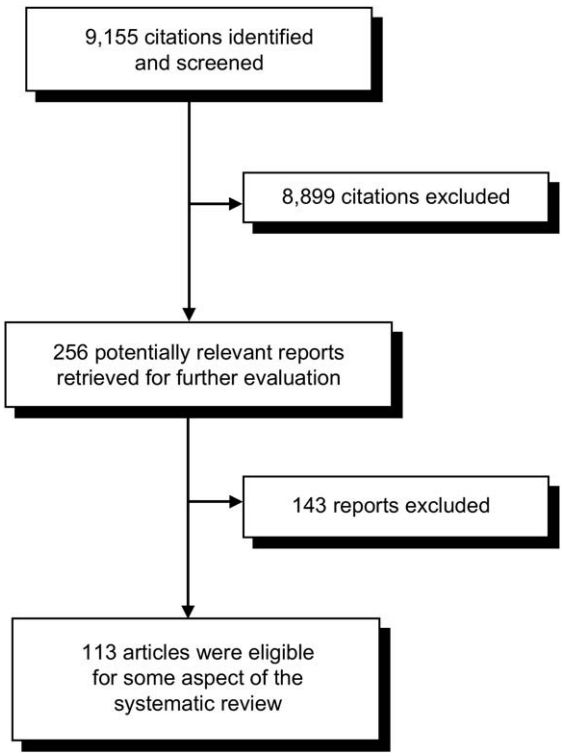

Fig. 1. Study flow chart.

4.4. Survey results

Surveys where the respondents consented to participation and provided at least one response beyond the demographics data were deemed to be valid and analyzed ($n = 58$).

The average number of years of searching experience of respondents was 12.6 years (standard deviation: 7.4 years). The average experience with systematic reviews or health technology assessment reports was 5.3 years (standard deviation: 3.8 years). Searchers reported prior involvement in an average of 18 systematic review or health technology assessment reports (standard deviation: 23.6). Most (79%) had a library science degree or similar. Nearly 85% of the respondents were from Canada, the United Kingdom, or the United States. Other countries represented were China, Denmark, Malaysia, New Zealand, Norway, South Africa, Spain, and Switzerland.

4.5. Survey rankings of elements

The survey ratings of elements clustered into three discernable tiers based on median levels of importance after response choices were converted to ordinal ranks (Table 2). The first tier had only first- and second-place rankings when ranked on the three variables *impact on recall*, *impact on precision*, and *importance in peer review*. These elements were perceived to have research support or be self-evident. The second tier had strong support. Examination of the actual scores showed these to be high, even though slightly lower or less consistently higher than for the first-tier elements. The survey respondents supported the inclusion of these elements in peer review. The elements in the third tier had no first- or second-place ranking on recall, precision or importance in peer review, and were not perceived to be supported by research evidence, nor a marker of inexperience in searching. The survey respondents did not support the inclusion of these elements in peer review.

5. Peer-review guideline

5.1. Tier 1 elements

5.1.1. Conceptualization

Recommendation

➢

Assess whether the research question has been correctly translated into search concepts.

Guidance

Review the conceptual elements of the search before considering any of the mechanical elements.

Consider that the combination of too many search concepts reduces recall.

The search strategy will ideally be submitted for peer review accompanied by the research question structured according to a format, such as *P*opulation, *I*ntervention, *C*omparison, *O*utcome (PICO) [6] and an explanation of how the reference interview informed the search.

5.1.2. Logical operators

Recommendation

➢
Assess whether the elements addressing the search question have been correctly combined with Boolean and/or proximity operators.

Guidance

Note any instances where an AND may have been used to connect words or phrases (e.g., as a conjunction) rather than as a Boolean operator, or where AND may have been inadvertently substituted for OR (or vice versa).

Check that nesting with brackets is logical and has been used where necessary. Consider whether precision could be improved by using a proximity operator (adjacent, near, within) instead of AND.

For proximity operators, consider if the width chosen is too restrictive to retrieve all likely circumstances of the use of the search terms, or too broad with resulting negative impact on precision, taking into account whether or not the database being searched uses stop words.

In searches translated from one search interface and/or database to another, correct specification of the proximity operators should be assessed, as there are many variations in implementation between databases and search interfaces.

If NOT is used, consider if there may be any unintended exclusions and whether another mechanism (e.g., use of a subject heading, limit, or check tag) would be a more suitable alternative.

For restrictions (e.g., human or child populations), verify that the construction used is correct. For example, it is better to exclude records by using phrases such as ‘‘not (animals/ not humans/)’’ instead of excluding all records indexed as ‘‘animals’’ or limiting searches to ‘‘humans’’ only.

5.1.3. Spelling errors

Recommendation

➢
Assess all spelling and system syntax.

Guidance

Breaking long strings of terms into discrete search statements (with interfaces that permit such an approach) will make null or smaller than expected retrieval because of misspelled terms more obvious and therefore easier to detect.

Search strategies should be reviewed for misspelled words, where possible by cutting and pasting them into a spell-checker but also by reviewing for errors in system syntax not typically amenable to spell-checking.

Note: for spelling variants see 5.2.3 below.

5.1.4. Wrong line numbers

Recommendation

➢
Assess for correct search implementation by checking each line number and combination of line numbers.

Guidance

To facilitate peer review and subsequent replication and updating of searches, lines related to the same search concept should be arranged contiguously, and experimental terms that do not belong in the final search should be deleted.

5.1.5. Translation of the search strategy to different databases and/or interfaces

Recommendation

➢
Assess whether the search strategy was correctly adapted for each database used.

Guidance

Check that subject headings, special features (including truncation and proximity operators), special indexes (including limits), and system syntax are correctly adapted for each database and interface searched.

It may not be practical to display each version of the electronic search strategy in a published systematic review, however, the methods section should contain a clear statement that the search was or was not tailored for various databases, and all versions of the search need to be peer reviewed. When the publishing format permits, all versions of the search should be published in full so that readers may assess the adequacy of the various adaptations. Where the adaptations are not all published in full, it should be stated that they are available on request from the corresponding author. Additionally, they may be made available on a Web site and the URL given.

5.1.6. Missed subject headings

Recommendation

➢
Assess whether there is enough scope in the selection of subject headings to optimize recall.

Guidance

Consider the following aspects of usage of subject headings:

the relevance of the terms
if terms are missing
if incorrect terms have been used
if a term should have been exploded to include relevant narrower terms
or if items might be missed because relevant documents could be indexed with a broader or more specific term than those selected by the searcher.

Floating subheadings are preferred to subheadings attached to specific subject headings (e.g., in MEDLINE [Spinal Cord Injuries/ and rh.fs.] instead of Spinal Cord Injuries/rh, where ''rh'' is the abbreviation for the subheading ''rehabilitation''). They should also normally be used in preference to the related subject heading (e.g., in MEDLINE, rh.fs. should normally be used rather than the subject heading Rehabilitation/). Finally, it is important to remember that subject headings and subheadings are database specific.

5.2. Tier 2 elements

5.2.1. Free-text terms missing

Recommendation

➢

Assess whether the search terms without adequate subject heading coverage are well represented by free-text terms, and whether additional synonyms and related terms are needed.

Guidance

As subject headings vary by database, the need for free-text terms may also vary. Consider the following aspects of usage of free-text terms:

the relevance of the terms
whether they are too broad or too specific
and whether synonyms are included.

5.2.2. Subject headings and free-text terms both used

Although this was not surveyed, considerable evidence supported the use of both subject headings and free-text terms for optimal search performance

Recommendation

➢

Assess the adequate use of both subject headings and free-text terms used in combination.

Guidance

If a search does not use both subject headings and free-text terms within each concept, the searcher should provide a sound reason for not using both.

Search terms should be grouped by concept—not according to whether they are free-text or index terms. If the searcher has essentially constructed two parallel searches, one with all concepts expressed in free text and the other with all concepts expressed using subject headings, this must be considered a deficiency.

5.2.3. Spelling variants and truncation

Recommendation

➢

Assess whether all relevant spelling variants are covered by the search terms.

Guidance

Ensure that British and American spellings are both used for free-text terms.
Consider whether the (truncated) word stem used will capture all necessary variants in spelling, without being unnecessarily short and likely to reduce precision.
Variants not amenable to be captured by truncation should be included as search terms.
Both the acronyms and the full terms should be used to obtain maximum retrieval.

5.2.4. Irrelevant subject headings and irrelevant free-text terms

Recommendation

➢

Assess the relevance of the terms used to the search question.

Guidance

The inclusion of irrelevant terms lowers precision and adds to the burden of scanning retrieved references and can result in relevant records being missed at the scanning stage.

5.3. Limits

Recommendation

➢

Assess whether the limits (including filters) used were appropriate and have been applied correctly.

Guidance

When limits are applied that are not relevant to the clinical question or eligible study designs, the potential introduction of epidemiological bias must be considered.
Ensure that methodological search filters are used appropriately, for instance, that searches for studies for inclusion in systematic reviews of diagnostic tests are not limited to randomized controlled trials.

Subject headings should not be restricted to focus or major emphasis (also known as 'starring').

See 5.1.2 when NOT is used to impose a limit.

5.4. Tier 3 elements

Recommendation

➢

It is not necessary to assess the following elements

organization of the search
use of additional database-specific fields
redundancies in the search
subject headings and free text combined in a single search statement.

6. Discussion

Evidence-based peer review of electronic search strategies needs to focus on those elements that will negatively impact search performance. A preliminary list of elements was explored through a systematic review of the literature. Several additional elements, with either some supporting research evidence or support from the survey, were identified, and some initial elements were refined. After seeking opinion from the community of interest, the list was reduced to those elements with at least some research evidence or strong support from the survey. The Tier 1 and 2 elements qualify in this regard and are recommended as the basis for peer review of electronic search strategies.

The evidence-based peer review of electronic search strategies requires the same body of expert knowledge needed to create search strategies. Peer review should be undertaken by librarians or other suitably qualified and experienced information specialists. Because a systematic review is based on material identified through the search, peer review of the search will most usefully occur at the beginning of the review, not as part of the peer review of the end product, be that a journal manuscript, technical report, or Cochrane review.

The findings of this research, as well as the peer review elements derived from them, are largely consistent with current best practice as described in authoritative sources, such as The Cochrane Handbook for Systematic reviews of Interventions [7]. Input was also sought from the health science library community, the health technology assessment information retrieval community, and The Cochrane Collaboration. Preliminary results of this work have been presented to these communities for feedback (cf. [8–10]).

Research evidence was drawn from all aspects of library and information science, not just health sciences librarianship. Thus, the peer-review elements identified could be applied to any type of literature search. In saying that, we caution that in assessing the evidence and formulating recommendations, we gave the greatest weight to maximizing recall because recall is the most important parameter of search performance for systematic reviews and health technology assessments [7]. In other situations, parameters such as precision or cost might be of greater importance than recall, and so the focus of the peer review should be adjusted accordingly.

The peer review process may slightly delay the development of the electronic search strategies. Time to complete a PRESS review is less than 1 hour; however, there may be some delay introduced by finding a willing reviewer whose availability matches the timeframe of the review. The peer review could result in improved search precision, which translates to time-saving in the screening phase of the review. Other foreseeable time-savings are avoidance of late corrections to the search with additional screening needed and late identification of important evidence and consequential revisions to the review. Even when no time-savings are achieved, there is greater assurance of the overall quality of the final systematic review, health technology assessment report, or guideline.

Several of the elements examined lacked research evidence of their impact on search performance. In most cases, the survey respondents assessed the impact as self-evident; however, a number of gaps remain. Pertinent to the fields of systematic review and health technology assessment reports, the importance of tailoring searches to individual databases has not been extensively researched, despite consensus that it is important. Although many studies demonstrate that both text words and subject headings are needed for complete retrieval, refinements that help maintain precision would greatly increase the efficiency of systematic reviews. For example, an improved understanding of when indexing practices may be inconsistent could inform when related subject headings may be needed to capture all possible indexing practices or when adding text words or abbreviations would improve retrieval vs. when such additions could be expected to simply lower precision.

Finally, additional work remains to be done to implement this evidence-based guideline and checklist in a peer-review process to occur early in the review, and validate that such a process does improve search quality. Training materials and an interactive Web-based forum for peer review of searches are in development.

7. Conclusions

The development of this peer review guideline fills a gap in the quality assurance of search methodology in systematic reviews and health technology assessment reports. Errors in electronic search strategies have been demonstrated to reduce the effectiveness of electronic search strategies used in systematic reviews and health technology assessment reports. This evidence-based guideline is based on a rigorous methodology, and it focuses only on those

elements of the electronic search where there was consensus on the importance of the element or research evidence support.

Those who sponsor, conduct, publish, or use systematic reviews and health technology assessment reports should demand peer-reviewed electronic search strategies, as the greatest portion of their evidence base will be identified by these searches.

Acknowledgments

We wish to acknowledge and thank Janet Joyce, who served as an advisor to this project, Raymond Daniel for technical assistance, Tamara Rader for writing assistance, Nick Barrowman and Isabelle Gaboury for statistical assistance, and all of those who participated anonymously in the survey. This research was supported by a capacity building grant from the Canadian Agency for Drugs and Technologies in Health. David Moher is supported by a University of Ottawa Research Chair.

References

[1] Jefferson T, Rudin M, Brodney Folse S, Davidoff F. Editorial peer review for improving the quality of reports of biomedical studies MR 000016. Cochrane Database Syst Rev 2007;2.

[2] Demicheli V, Di Pietrantonj C. Peer review for improving the quality of grant applications. MR000003. Cochrane Database Syst Rev 2007;2.

[3] Sampson M, McGowan J, Tetzlaff J, Cogo E, Moher D. No consensus exists on search reporting methods for systematic reviews. J Clin Epidemiol 2008;61:748−54.

[4] Sampson M, McGowan J. Errors in search strategies were identified by type and frequency. J Clin Epidemiol 2006;59:1057−63.

[5] Sampson M, McGowan J, Lefebvre C, Moher D, Grimshaw JM. PRESS: peer review of electronic search strategies. CADTH Technical Report 2008; Available at: http://cadth ca/index php/en/publication/781 http://www.cadth ca/media/pdf/477_PRESS-Peer-Review-Electronic-Search-Strategies_tr_Appendices pdf.

[6] Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. ACP J Club 1995;123:A12−3.

[7] Lefebvre C, Manheimer E, Glanville J. Searching for studies. In: Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions. Chichester, UK: Wiley; 2008. Available at: www.cochrane -handbook.org.

[8] McGowan J, Sampson M, Lefebvre C. Can the quality of literature searches be measured and improved? The EHTAS project. Brisbane, Australia: 3rd International Evidence Based Librarianship Conference; 2005.

[9] Sampson M, McGowan J. Evidence review—evaluating health technology assessment searches. Presented at: XIII Cochrane Colloquium, Melbourne, Australia; 2005.

[10] McGowan J, Sampson M, Joyce J. Peer review electronic search strategy [oral presentation]. In Fifth Annual Meeting of Health Technology Assessment International (HTAi), Information Retrieval Group (IRG). Montréal July 7, 2008.