Author's response to reviews

Title: An overview of the design and methods for retrieving high-quality studies for clinical care

Authors:

Nancy L Wilczynski (wilczyn@mcmaster.ca)
Douglas Morgan (dmorgan127@cogeco.ca)
R. BRIAN Haynes (bhaynes@mcmaster.ca)

Version: 2 Date: 31 March 2005

Author's response to reviews: see over
Reviewer's report

Title: An Overview of the Design and Methods of an Information Retrieval Study
Version: 1 Date: 27 December 2004
Reviewer: William Hersh

Reviewer's report:
General
This paper provides details of the methodology used in the important and well-known work from this group in identifying optimal strategies to retrieve studies likely to contain the best evidence applicable in clinical care. Their work has had substantial impact, being incorporated into the NLM PubMed system. This paper provides much greater detail into the methods used for this work than has been allowed by the clinical journals, such as BMJ and JAMA that have published the results. As a result, BMC is an excellent place to publish this paper. The paper adheres to the high standards one expects of this group. I do, however, have a few suggestions/questions.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1. I think the title could be better. This is not just “an information retrieval study”; but rather methods to retrieve studies with the best evidence for clinical care. The title should reflect this more specific task. How about something like, “An Overview of the Design and Methods for Determining Studies Containing the Best Evidence for Clinical Care”; or something like that?

We agree with the reviewer and thank him for his suggestion. The title of the paper has been changed to: “An Overview of the Design and Methods for Retrieving High-Quality Studies for Clinical Care.”

2. One concern I have always had with this work is how these hedges fare in the “real world” searching environment. That is, how well do they aid the user when he or she is adding clinical search terms to these strategies? As such, this work does not represent the last work in clinical information retrieval. The authors should address this issue and perhaps summarize other work in the field that looks at using retrieval systems to apply evidence in the clinical setting.

We acknowledge the reviewer’s concern. We recently submitted a manuscript for consideration for publication that looks at “real world” use of these methodologic strategies when ANDed with content terms in the discipline area of mental health. The following sentence has been included in this manuscript just prior to the conclusion:

“Additional research is underway in search strategy development including testing the strategies developed through this research, when combined with disease content terms, and when combined with terms using the Boolean “AND” and/or “NOT”.”

3. Table 6 states that data collection for Etiology is “prospective.” My previous
understanding of EBM is that retrospective case-control studies are an acceptably good form of evidence for studies of harm. Will their search strategies not identify case-control studies? If so, is that inconsistent with commonly accepted practice of EBM? Related to this, none of the other cells in the table mention prospective vs. retrospective data collection.

Our criteria include case-control studies, but only if the data collection is prospective, as in a nested case-control study. Case-control studies that are based on retrospective data collection are simply too prone to bias to be classified as “high quality.” To clarify the source of our criteria, we have added the following sentence:

“The methodologic criteria outlined in Table 6 are the same as those used for critically appraising articles for inclusion in 4 evidence-based medicine journals that our research group produced in 2000 (i.e., ACP Journal Club, Evidence-Based Medicine, Evidence-Based Nursing, and Evidence-Based Mental Health).”

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Discretionary Revisions (which the author can choose to ignore)

What next?: Accept after minor essential revisions
Level of interest: An article of importance in its field
Quality of written English: Acceptable
Statistical review: No
Reviewer’s report
Title: An Overview of the Design and Methods of an Information Retrieval Study
Version: 1 Date: 21 March 2005
Reviewer: Kaveh Shojania

Reviewer’s report:
This group has produced a number of very useful articles and related work products for supporting effective literature searching. The work is commendable and serves an important purpose. The present work provides more methodologic details than allowed in the brief methods sections of the articles published to date.

---------------------------------------------------------------------

Major Compulsory Revisions
None

---------------------------------------------------------------------

Minor Essential Revisions

p.5 “Each item… were classified” should be “…was classified”

Manuscript changed as indicated by the reviewer.

“Data” is plural, so p. 6 “Hand search data was recorded…” should be “data were recorded” and p.9 “These data was organized by…” should be “These data were…”

Manuscript changed as indicated by the reviewer.

---------------------------------------------------------------------

Discretionary Revisions

1. The extensive iterative automated testing amounts to a derivation set. Was performance for the search strategies re-evaluated on a different set of article (i.e. a validation set?) if not, the methodology runs the risk of over fitting, as in any automated analysis of a database. The authors might comment on this issue (or clarify why it is not an issue, if I have misunderstood).

As indicated on page 13 of the manuscript:

“We also found that when strategies were developed in 60% of the database and validated in the remaining 40% (a random allocation method was used to assign individual items to the development or validation datasets) there were no statistical differences in performance.”

Thus, for several purpose categories strategies were developed in one database and re-evaluated (validated) in a different database. Since there was no statistical difference noted this practice was not continued.

No change to the manuscript.

2. The automated, iterative testing of the search terms is impressive. However, it remains unclear to me if there may be problems related to the requirement that all terms be combined in strings of OR operators. Surely some terms could perform quite well if they were ANDed with something or if they could have a NOT phrase attached. For
instance, suppose one has a string of the form $x \ OR \ y$, but it retrieves too many basic science articles might be improving by changing to $x \ OR \ y \ BUTNOT$ (Physiological Processes [mh] OR Biological Phenomena, Cell Phenomena, and Immunity [mh]). I made up this example for illustrative purposes. The point is that many terms might perform better of they could be combined in different ways with respect to OR, AND, BUTNOT, but the methodology used by the authors seems to allow only the possibility of adding or subtracting to strings of ORed terms.

Excellent point brought up by the reviewer. Initial search strategy development concentrated on “OR” combinations of terms. Preliminary testing of AND strategies invariably resulted in reduced sensitivity and AND NOT strategies led to minimal changes in performance. However, we intend to extend this aspect of the research. The following sentence now appears in the manuscript prior to the conclusion:

“Additional research is underway in search strategy development including testing the search strategies developed through this research in journal subsets, when combined with content terms, and when combined with terms using the Boolean “AND” and/or “NOT”.”

3. p.8 “Since the primary goal of the Clinical hedges Study was to deliver search strategies which could be used by clinicians…” Although OVID is very common among academic institutions, it is not among the total population of clinicians. It is not clear then why they used OVID instead of PUBMED, as the latter would be accessible by a far greater number of clinicians. The authors’ BMJ article on review articles addresses this problem by providing translations for PubMed. Were translation performed for all of the searches and were they evaluated in the same way as described for OVID?

All strategies developed for MEDLINE have been translated for use in PubMed by staff of the National Library of Medicine, and compared for performance by the senior author (RBH). The following sentence has been included in the manuscript on page 13:

“Search strategies developed for use in MEDLINE have been translated for use in PubMed by staff of the National Library of Medicine, and compared for performance by the senior author (RBH).”

4. Table 2 focuses on whether or not articles were “concerned with understanding human health” but this seems quite vague. This judgment must have been operationalized more specifically. For instance, were articles that did not involve human subjects automatically excluded? Gender and equality issues are mentioned as excluded, but what about the rest of health services research? Quality measurement? Some of these articles would likely be difficult to judge as not “of interest to understanding human health”.

In table 2 we have now included the exclusion of animal studies as an example. An extensive coding manual was used by the research assistants when reviewing articles. The development of the manual and the calibration exercise to prepare for data collection for this study occurred over a 14-month period (details are outlined in the first referenced paper in this manuscript). Inter-rater reliability was very high for the readers for all classifications (all kappas > 0.80).
What next?: Accept after discretionary revisions
Level of interest: An article whose findings are important to those with closely related research interests
Quality of written English: Acceptable
Statistical review: No
Declaration of competing interests:
I declare that I have no competing interests