Reviewer’s report

**Title:** Developing Optimal Search Strategies for Detecting Clinically Sound Prognostic Studies in MEDLINE

**Version:** 1  **Date:** 3 March 2004

**Reviewer:** Susan Wieland

**Reviewer’s report:**

General

The question posed by the authors (what are the best search strategies for retrieving methodologically sound prognostic studies in MEDLINE) is important, and it is valuable for clinicians to be made aware of the underlying research supporting the development of such search strategies. This paper discusses the strategies that the authors used to develop the sensitive and precise search strategies currently found in PubMed Clinical Queries for prognostic studies. However, while this manuscript is readable and almost completely free of technical jargon, it does not provide enough detail about the development of the search strategies.

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

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

1. In the Methods section (page 6, top of page) and the Results section (page 7, bottom of page) the manuscript is imprecise about the population of studies upon which the search strategies are run. Are the searches tested against all 49,028 studies from the 161 journals, or only the 1,547 studies classified as prognosis studies? This lack of clarity makes it vague whether the research is focused on sorting high from low quality prognostic studies or on sorting high quality prognostic studies from both low quality prognostic studies and all non-prognostic studies. This needs to be clarified in the manuscript.

2. In the Methods section (page 6, second paragraph), the criteria used for considering a report on prognosis to be methodologically sound are quite reasonable. In the Results section (page 7, last paragraph), only 12% of all prognosis articles were considered methodologically sound. Given the small proportion of prognostic studies that meet criteria for methodological soundness, it would be useful to be given an idea of what characterize most prognostic studies. For example, what is the distribution of methodological flaws in prognostic studies -- did most studies fail to be sound due to poor analysis, or lack of followup? Is there any drawback to excluding methodologically unsound prognostic studies from a search strategy for prognosis? Some reassurance about this issue should be discussed in this manuscript.

3. Finally, the authors have written about a purely methodological search filter. They state in the Discussion section (page 9, second paragraph) that limiting searches by clinical content terms might increase precision, but that an increase in precision cannot be assumed. The authors hint in the last sentence of this paragraph that they may focus on using content specific terms in the next phases of their project. The results to be expected from combining disease-specific terms with the methodologic terms are surely of great interest to clinicians and need to be discussed further in this manuscript.

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

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

1. In the Discussion section (page 9, first paragraph) the authors use the term ‘false drops’. This term is jargon and should be replaced with a term that can be understood by the general medical reader.

Discretionary Revisions (which the author can choose to ignore)

1. In the Background section (page 5, second paragraph), the authors discuss how the current study builds upon previous research they have done in this area. The current research relies upon 161 journals (instead of 10 in the authors’ research from the early 1990s, used to develop the current Clinical Queries interface) to develop a gold standard, and upon a broader range of MEDLINE terms to develop the search strategies. In the Methods section (page 6, last paragraph, and page 7, second paragraph), the authors provide information upon the processes used to select the journals and the search terms, and offer to provide a list of the journals and a list of the MeSH terms and textwords upon request. However, this manuscript would be strengthened by presenting more specifics about the journals and the search terms within the text, perhaps in the form of tables highlighting key journals and search terms. This would give readers more confidence in the methods used, and allow replication of the research.

2. In the Methods section (page 7, very top of page) the authors say that inter-rater agreement for application of criteria was very high, and cite a paper from 2001. Do the authors mean to say that they used methods previously described to have high inter-rater agreement, or that the inter-rater agreement for the review of these 161 journal titles was itself 80% beyond chance?

3. In the Methods section (page 7, second paragraph) the authors say that the database was randomly split but they do not state the method used. Do the authors mean that the database was divided in a non-systematic (haphazard) way, or did they use a true random method?

4. In the Discussion section (page 9, last paragraph), the authors mention that the best performing strategy for sensitivity was the same as the strategy they reported in 1991, but the references are to articles in 1993 and 1994. Is it possible that the strategy was developed in 1991 but reported in 1993 and 1994, or is there a reference from 1991?

5. In the Discussion section (page 9, last paragraph), the authors do not discuss whether their 1991 results differ from the current results for specificity, precision, or accuracy. Were there any differences in these areas? If so, what is the implication of these differences?

6. This research was developed using Ovid’s search engine syntax. In the Discussion section (page 10, third paragraph) and the footnotes to Table 1 and Table 2 (pages 14 and 15), the authors mention that the strategy is contained in PubMed’s Clinical Queries. However, the PubMed syntax is not presented anywhere in the manuscript. Given that PubMed is a free resource that is available to anyone with an internet connection, presenting the PubMed syntax would make the research more replicable. The authors should briefly mention why they chose to develop their strategy using Ovid syntax rather than PubMed syntax. Since Ovid and PubMed do not contain exactly the same information, it would also be interesting to know how results from searches run in Ovid might compare with searches run in PubMed.

7. In the Discussion section (page 10, third paragraph) the authors mention that the Clinical Queries interface of MEDLINE will be updated to reflect the new search strategies. Because the authors already discuss (page 9, last paragraph) that their best strategy for sensitivity is the same as that reported earlier, and they do not mention how the best strategy for specificity has changed, it is not clear from the Discussion section how Clinical Queries will be updated. The authors should clarify this point, and also discuss whether Clinical Queries will include a strategy optimizing both sensitivity
and specificity, as presented in Table 2. Finally, it is not clear from looking at Clinical Queries whether the update of Clinical Queries has already happened; if the update has already happened, the authors should mention this, and if the update has not already happened but the authors have any indication of the expected date of the update, the authors should give this date.

8. The first column of Table 1 (page 14) is laid out in a confusing manner so that it appears that there are three terms in the first column. If it is possible, the term should be contained on a single line.

9. The text does not explain why the best single search term presented in Table 1 (page 14) is of interest. When would a searcher want to use the single search term rather than the various combinations of terms? Please discuss this further in the text.

Which journal?: Appropriate or potentially appropriate for BMC Medicine: an article of outstanding merit and interest in its field

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:
None