Reviewer's report

Title: Systematic Reviews: A Cross-Sectional Study of Location and Citation Counts

Version: 1 Date: 23 August 2003

Reviewer: Michael Callaham

Reviewer's report:

General

The authors have conducted a fairly carefully crafted and objective assessment of the state of systematic reviews in the clinical medical literature. The topic is important, since systematic reviews are touted as being an important advance in the quality of information available to practitioners, but little is known about their prevalence and effectiveness. This paper does not address the latter, but does provide a unique profile of which journals are publishing this type of study. Less clearly described, but even more interesting, are some associations found between review quality, citations, and journal impact factor. The paper is for the most part clearly written and well organized, although certain sections need more detail and occasionally the phrasing is cryptic.

There are a number of areas that are unclear in the text and which I have listed below as minor compulsory revisions (although none are so minor as the examples given in this review form template). There are also several larger topics which are not fully addressed or clarified, and which I think would greatly strengthen the paper if addressed. I have listed these under major compulsory revisions, although I suppose another approach would be to ignore them or delete the topics altogether, but then the paper would have considerably less interest.

Overall, the paper adds some important information to our knowledge of the “epidemiology” of these reviews, although it does not (and is not meant to) address issues of their usefulness and effectiveness. Although not of great interest to clinicians, the information expands our knowledge of peer reviewed publication. However, it needs some substantial revision of the methods section to provide sufficient detail. Furthermore, the current discussion is very brief and does not much put the findings in perspective for the reader.

Discretionary Revisions (which the author can choose to ignore)

p. 3 para 2, “such as those who do not appear to different importantly....” This really contradicts your statement about a “wide array of patients”, doesn’t it? Most of the RCTs on which meta-analyses are based are limited by the very particular and often selective nature of the patients enrolled, which makes generalization difficult.

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

There are no page numbers

P. 4 para 1 – there is no citation for the statement that others have authored “3 to 5 times that number” which are “dispersed throughout the medical literature”. Where does this information come from?
Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

P. 5 Methods – Your methods are only superficially described and leave too many questions unanswered. Some of the statements are quite confusing. For example:

What recommendations by what clinicians and librarians? Something systematic, or just folks you knew?
How did you take into account impact factors? Did you choose the highest? Or some other method?
What does the “iterative process of evaluation of journal yield of studies…” etc. mean?
You can see from the above that no one could know exactly what you did, or could duplicate it, from this description

Why in particular did you choose nursing and mental health? Why family medicine (a specialty not much given to RCTs and meta-analyses) as well as internal medicine? What was your reasoning?

What methodologic criteria did the research associates apply, where did they come from, have they been validated? Was each article rated by only one researcher, or did you require a certain level of agreement among all 6? You refer later to kappa (belonging in Results) but don’t describe how agreement was reached on categorization.

Are you stating they actually reviewed 60,330 articles? If so, this was a massive undertaking; is your study a sub-analysis of some much larger endeavor? If so, the other papers and the methods should be cited. (By the way, this information belongs in the results section).

Why were you indexing original articles by purpose (and again, how did you choose these categories)? You have lumped together 29,887 articles of 2 types; separate reporting would be clearer.

“staff were rigorously calibrated”… do you mean trained? Tested? Their methods validated? You describe them as “highly trained” in the discussion section, but have not given us any of the details. Typically with this methodology the measurement tools you used, and their background and validation, would be described in detail, as would be the training of the staff and your examination of their reproducibility and agreement. Most other studies of the “quality” of scientific reports have demonstrated that this kind of methodology is in fact not very reproducible, and not well proven to actually represent “quality”.

P. 6 “queried the database for … reviews….” Please describe this in more detail. What search terms were used? I suspect the search was too simple for this description and you might want to say you simply examined the XX number of reviews in the database. Was this database assembled for another purpose (see above)?

Your definitions of “review” and “systematic reviews” (especially the latter) – are there standardized definitions in the literature? Did you use them? Or are these definitions you made up de novo for your study?

p. 6 para 2 “A second level of classification….” I really do not understand what this sentence means. What definition did you use for “understanding of healthcare in humans?”

Classification in purpose categories (etiology, etc.) – do these come from some previously defined source or publication? Or did you create them for this study? How did you select these categories, and what was the purpose in having them?
Fourth level of classification – again, not clearly defined here and seems incomplete; is it based on prior study? The literature to date suggests such definitions of quality, even when carried out meticulously, are not very reliable.

Citation query – how did you choose the date of February 2003 for all 2002 publications? First of all, most clinical medical publications are not cited much in the first year after publication. Second, the intervals after publication of the articles you studied would vary widely and were not controlled for in your analysis. Third, this methodology is simultaneously measuring both impact factor (in the form of citations) and immediacy factor (in other words, the speed of citation varies between journals and specialties).

Subset of 5 journals – what exactly was your rationale for picking these 5 journals? Had you previously determined you would pick the top 5 (instead of the top 6, or 7) or was it based on some proportion of total publications? In Table 1 you refer instead to the top 20 journals.

Results and Discussion

I agree that your findings demonstrate relative priority of systematic reviews, which is an important finding and potentially reassuring. On the other hand, although the association between impact factor and number of “rigorous” reviews was statistically significant, I would conclude it is not “clinically” significant because the association is so weak (in the full dataset) that it accounts for almost none of the variance. An equally valid conclusion would be that number of rigorous reviews is essentially not related to impact factor, and that even in the general medical journals it accounts for a minority of the variance.

Do you take this association to suggest that higher quality (impact factor) journals can attract and publish better reviews, or that better reviews in journals tends to increase their impact factor?

Your finding that 11% of all journals published 80% of all systematic reviews is one of the more interesting you report. What do you think may account for this maldistribution? Can you account for it by specialty or other factors? What can (or should) smaller specialty journals want to do about this maldistribution?

Your finding of the citation number for narrative vs. systematic review published in the same journal is one of your more important results, and deserves more discussion.

Overall, your discussion is a bit lean and it would be helpful if you discussed some of the implications of your findings and how they relate to what we already know about systematic reviews. What do you think of the absence from your list of some of the most influential and widely read journals, such as NEJM?

P. 9 limitations

Another limitation (beyond your control) is that your study could only measure citations, and in fact the chief intent of reviews is to influence the behavior of clinicians. You have no data on how much these reviews were either read or followed by clinicians, as compared to the relatively few academics who cited them.

Tables:

Table 3

It’s discouraging (and a major limitation) that this analysis covers really only 78 systematic reviews in four medical journals. Nursing is really a separate field (with numbers much smaller) and the
Cochrane Library is also in its own category. An interesting question is, what is the “readership” of the Cochrane Library? My own experience suggests to me that even many academicians do not regularly consult it, and practicing clinicians very seldom.

**Advice on publication:** Accept after minor compulsory revisions

**Level of interest:** A paper whose findings are important to those with closely related research interests

**Quality of written English:** Needs some language corrections before being published

**Declaration of competing interests:**
None