Reviewer's report

Title: A pilot study of a new system for grading the quality of evidence and the strength of recommendations

Version: 1 Date: 25 March 2004

Reviewer: Joseph WATINE

Reviewer's report:

Minor Essential Revisions which the author can be trusted to correct:

Second manuscript: ‘A pilot study of a new system for grading the quality of evidence and the strength of recommendations’

Please read my comments to the first manuscript before reading the comments below.

1) The question posed by the authors is new but they only answer to a part of it. In fact, what the authors have done is “a pilot study of a new system for grading the quality of evidence and the strength of recommendations about the effectiveness of therapeutic and prophylactic interventions”. The system which has been pilot-tested is their preliminary and “oldest” system. Their modified and “newest” system, which they propose at the end of their manuscript (and which is also described in a paper currently “in press” in the BMJ) is not the system which is pilot tested in this study.

Also, in the Background section, mention should be made of the paper “in press” in the BMJ.

2) Are the methods appropriate and well described, and are sufficient details provided to replicate the work?

The Methods section could perhaps be divided into under-sections.

At the very beginning of the Methods section, it is written that “17 people independently judged the quality of evidence”. There are 18 authors for this manuscript. Who is the author who did not participate in the judgment of the quality of evidence? Why does this 18th author deserve to be a co-author?

How were the evidence profiles made? Whom composed these evidence profiles? What is the utility of the whole grading system if the systematic reviews, on which the evidence profiles are based, and therefore on which all the grading processes are based, are not properly done? Which system, if any, did the authors use to judge the methodological rigor of the 12 systematic reviews that they used to compose the 12 evidence profiles? As the authors probably know, many such systems have been proposed and published, e.g. that of McAlister FA, Clark HD, van Walraven C, Straus SE, Lawson FM, Moher D, Mulrow CD [The medical review article revisited: has the science improved? Ann Intern Med 1999 Dec 21;131(12):947-51]. If such a system has been used, this should be said in the Methods section, and probably discussed in the Discussion section. The reference of the system should also be quoted.

3) Are the data sound and well controlled?

Yes except for the way the evidence profiles were made, as already written above.
4) Does the manuscript adhere to the relevant standards for reporting and data deposition?

No keywords are provided.

In table 7, summary of findings: “absolute” (not “absloute”).

In table 8, the word “moderate” is used in the table, whereas in the legend, it is the word “intermediate“ which is used.

5) Are the discussion and conclusions well balanced and adequately supported by the data?

Couldn’t the author discuss the fact that the methodological quality of a systematic reviews that is used to compose an evidence-profile might be another item to be included in the evidence profile?

The issue of patients’ choice is perhaps not discussed as one might have hoped.

In the table of the appendix, page 30, we do not know what the differences are between “serious flaws” and “very serious flaws” or between “some uncertainty” and “major uncertainty”. Is the quality of the evidence of a randomized controlled trial with serious flaws the same as that of an observational study with a strong, consistent and direct association and no plausible confounders?

When reference is made of the first manuscript which is being submitted to BioMed Central, it would be worth mentioning that it included not only questions about effectiveness and harm, but also about diagnosis and prognosis, whereas these latter questions are not really the subject of this second manuscript. It should be made clear that the system developed in the second manuscript only applies to judgments about the effectiveness of therapeutic or prophylactic interventions, and not to diagnostic interventions, neither to economic, etiological, or to prognostic studies.

6) Do the title and abstract accurately convey what has been found?

As already suggested above, a more accurate title would be: 'A pilot study of a new system for grading the quality of evidence and the strength of recommendations about the effectiveness of therapeutic and prophylactic interventions'. This should be made clearer in the abstract too.

Also in the Methods section of the abstract, the first sentence could perhaps be rewritten something like this: “Twelve evidence profiles were prepared based on the results of 12 systematic reviews” (the word “example” is already used in the following sentence).

7) Is the writing acceptable?

Being French, I am not sure that my English is as good as that of the GRADE Working Group. However, it seems to me that their style might be improved in order to make the text easier to be read and understood.

Page 11, first sentence of the “Sensibility and understandability” section: there is a spelling mistake (“rates” instead of “raters”).

8) If these two manuscripts are accepted for publication, I confirm that you may feel free to post my signed reports, together with the manuscript, as part of the pre-publication history.

Dr Joseph Watine, PharmD, AIHP, AHU, EurClinChem
Laboratoire de Biologie Polyvalente
Centre Hospitalier Général
Rodez
France

**What next?:** Accept after minor essential revisions

**Level of interest:** An article of importance in its field

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

**Statistical review:** No

**Declaration of competing interests:**

See my conflicts of interest in my comments to the first manuscript.