Author's response to reviews

Title: The GRADE approach for assessing new technologies as applied to apheresis devices in ulcerative colitis.

Authors:

Nora Ibargoyen-Roteta (n-ibargoyen@ej-gv.es)
Iñaki Gutierrez-Ibarluzea (osteba7-san@ej-gv.es)
Rosa Rico-Iturrioz (rosarico-osteba@ej-gv.es)
Marta Lopez-Argumedo (osteba5-san@ej-gv.es)
Eva Reviriego-Rodrigo (ereviriego@ej-gv.es)
Jose Luis Cabriada-Nuño (jcabriada@gmail.com)
Holger J Schünemann (schuneh@mcmaster.ca)

Version: 2 Date: 29 January 2010

Author's response to reviews: see over
On behalf of myself and my co-authors, I submit the enclosed manuscript for consideration by the Journal, and I warrant that the article is original, has not been formally published in any other peer-reviewed journal, is not under consideration by any other journal and does not infringe any existing copyright or any other third party rights.

We would be grateful if you could consider our article for publication in Implementation Science because we know that it would be freely available online to the entire global biomedical community, what ensures a high visibility for the work made, and because of the high standard of peer review of the BioMed Central’s journals.

We state that we received 100% public funds for the development of the research and for the publication of the article. We also declare that one of the authors (Holger Schunemann) works in the development of the GRADE approach and that his contribution was mainly related to teaching about GRADE, the revision and discussion of the results obtained by this work and helping with the correct application of the GRADE approach by the team.

All authors have read and approved the paper, have met the criteria for authorship as established by the International Committee of Medical Journals Editors, believe that the paper represents honest work and are able to verify the validity of the results reported.

Please, find below the responses to the referee reports.

Signature: Nora Ibargoyen Roteta
RESPONSE TO THE FIRST REFEREE

Comment 1: “Department of gastroenterology instead of digestive department?”
We have replaced “Digestive Department” with “Department of Gastroenterology”.

Comment 2: “Be more explicit: e.g. methods section line 3: At the end the overall quality of the evidence for each question was considered for the formulation of recommendations”
We have added the following sentence to the manuscript: “Finally, the overall quality of each question was taken into account to formulate recommendations following the GRADE approach”.

Comment 3: “It is unclear to what method of formulating HTA recommendations the GRADE method is compared in this study. What is common practice in your centre and/or in your country?”
In our context, the SIGN method is used to make recommendations, mainly in the form of Clinical Practice Guidelines. We have added a comment to this effect (“To evaluate this experience, a SWOT (Strengths, Weaknesses, Opportunities and Threats) analysis was performed to enable a comparison with our previous experience with the SIGN method”).

Comment 4: “The final sentence touches upon the distinction between assessment and appraisal. This could be elaborated upon a bit in the discussion section of the article.”
This is certainly a possibility, therefore we have decided to remove this comment from the manuscript as it has no bearing on the points we want to make.
Comment 5: “For non-clinicians: what is steroid tapering?”

Steroid tapering: in an ulcerative colitis flare, when the steroid dosage (usually 40 or 60 mg per day) has produced significant clinical improvement, the dose of corticosteroids is reduced at a rate of 5-10 mg per week. For a daily dose of 20 mg, tapering generally proceeds at 2.5 mg/week.

Comment 6: “The objective of the study was explorative, a kind of feasibility- or pilot study (the authors use the latter term in the discussion). At first sight, however, one would expect a comparative study, e.g. with SIGN…”

This is indeed the case and the appropriate change has been made. This is more of feasibility or exploratory study to determine whether this methodology can be used to assess new technologies, rather than a pilot study.

Before doing a controlled study we need to know whether using the GRADE approach in this new context is feasible. The SWOT analysis describes the “risks and benefits” of using the GRADE approach in this context (new and emerging technologies), using apheresis treatment for UC patients as an example. In light of the results of this study, a controlled trial should be performed to determine whether the recommendations obtained for the use of different technologies using different approaches differ. We should also explore the participant’s experience when doing this exercise (comment added to the discussion).

Comment 7: “The essential difference between treatment indications seems to be corticosteroid dependent versus non-corticosteroid dependent and corticosteroid refractory versus non-corticosteroid refractory ulcerative colitis. This requires some more information and explanation as part of the background section …”
We have added further information regarding literature precedents. Initially we only considered two questions related to the new treatment, whereas perhaps we should have included further examples. For example, we could have studied whether this treatment can be used in children or in patients with toxicity to corticosteroids and use the results to make a recommendation based on the evidence found in the literature.

**Comment 8:** “Mild adverse events instead of effects?”

Change made as requested.

**Comment 9:** “Isn’t the IBDQ relevant as an outcome measure in the context of the first question as well?”

One would expect this to be considered for the first question also, although the working group did not define this outcome as important or critical in this context (see the table with the scores), which is why it was not considered here.

**Comment 10:** “How was consensus reached on the importance of the defined outcomes??…”

Consensus was obtained as follows: each researcher scored the relevance of each defined outcome individually. The results were then shared. In the event of doubts concerning the relevance of any of the outcomes, it was discussed amongst the five researchers until consensus was obtained. Throughout this process, we attempted to keep in mind which could be the possible values and preferences of target patient groups. A final score of 7-9 points was considered critical, and a final score of 4-6 was considered important, as defined in the GRADE approach (see table 1 for the points awarded by each researcher).
Comment 11: “Assessment of the outcomes. The criteria may need a bit more explanation (the reader now has to continue two more pages to learn more).”

A more detailed description of each criterion has been provided.

Comment 12: “How did you agree on recommendations? Could you give a few examples of your considerations in specific cases (e.g. in text boxes)?…”

We have added this point to the discussion as a limitation of our method. Our recommendations are based on our results, the risk-benefit ratio, possible patient values and preferences and the costs. We decided to include an expert to check how well we were doing in terms of the outcomes definition and PICO questions etc. We note in the discussion that this was a limitation of our study as this expert could introduce a bias (a larger number of experts would be more interesting in terms of being able to include different points of view). For further studies it would be interesting to include clinicians when elaborating recommendations regarding new technologies. The GRADE working-group member (HJS) was included to ensure that we were applying the GRADE approach correctly.

Comment 13: “On page 6 you write that you will use the Endoscopic Mayo Subindex to assess endoscopic remission. How does this compare to the information on page 10 on the Rachmilewich Endoscopic Index…”

The problem with this disease is that is very heterogeneous in terms of both the definition of patients and the heterogeneity of the activity indexes used to assess its activity. We defined endoscopic remission using the Mayo Sub-endoscopic index. The main problem encountered for both questions is that all the studies selected used another index to measure endoscopic activity, namely the Rachmilewich Index. Moreover, these studies
did not define endoscopic remission as an outcome of interest—they only measured the endoscopic index before and after treatment. We therefore decided to use these data to assess the endoscopic remission indirectly in order to avoid losing this information (and that is why we downgraded the quality of these findings).

Comment 14: “Balance between risks and benefits. It seems that a decision on the choice of treatment for each (sub) indication goes together with/is based on many different considerations…”

This is indeed the case. We made a “sub-recommendation”, whereas we should have defined another PICO question for that purpose and searched for studies related to that question. We would then be able to consider the quality of the evidence found for this purpose and subsequently elaborate a recommendation. We have therefore eliminated the comment related to children, and the “sub-recommendation”, from the corresponding paragraph.

Comment 15: “Recommendation 1: Shouldn’t you write: for most corticosteroid-dependent etc.?…”

The second question was specifically defined for this type of patient as, according to the consensus Spanish document and published studies this is likely to be the most important target group (we have added more, and hopefully sufficient, background information).

Comment 16: “Why was the study with the prednisolone-resistant or dependent patients excluded?”

The study of Jo et al was a retrospective study where apheresis treatment was more likely to have been applied to those patients who had been resistant to or dependent on
prednisolone. That would mean heterogeneity between patient groups: in fact, the authors state that previous prednisolone dose was significantly higher in the leukocytapheresis group than in the prednisolone group and that a lower remission rate and a higher proctocolectomy rate were found in the leukocytapheresis group compared with the prednisolone group. These were the reasons why this study was not included.

**Comment 17:** “Balance between risks and benefits. There seems to be more benefits than harms of apheresis systems treatment?”
Ok, change made as requested.

**Comment 18:** “Please explain how the SWOT analysis process considers patients values and why the outcomes reported in the literature need to be avoided…?”
The SWOT analysis doesn’t consider the patients’ values and preferences (this is the GRADE approach). We have already changed this in the manuscript (the wording was confusing…).
When we say that the outcomes discussed in the literature need to be avoided, we mean that it is more important to determine the most relevant outcomes from the patients’ point of view before searching the literature. This prevents the outcomes of previous clinical trials or other studies having any influence in the definition of the outcomes of interest when using the GRADE approach.

**Comment 19:** “halfway the page: strength of[0]recommendations?”
Could this error only be present in the PDF produced by the program?
Comment 20: “Conclusions. How about this formulation? Our study suggests that the GRADE approach could be a proper approach to make the process of the formulation of recommendations a more transparent part of the overall process of producing HTA reports…”

Thank you! That sounds much better! We have added about a brief comment concerning the relationship between our work and the value of an information approach to new research.

RESPONSE TO THE SECOND REFEREE

**Major compulsory comments**

**Comment 1:** “Page 11 and second main recommendation on page 12: the authors state: “Therefore, this treatment could be used as the first choice in patients presenting toxicity to corticosteroids or especially…”

We should not have referred to children and patients with corticosteroid toxicity as we have not searched for evidence for this. We only studied two PICO questions although further questions concerning the assessment of this technology could be proposed (one of these would be related to this population, as noted in the discussion). We have eliminated this “sub-recommendation” as we did not search for any evidence to support it.

**Comment 2:** “Table 2: Footnote 1 states that one included trial was not randomized but rather observational in nature. The authors should clarify if mixing experimental with observational data within a meta-analysis is desirable”
The study of Nishioka et al. was not randomized and this issue has been pointed out in the evidence profile. Although it wouldn’t be recommended to mix experimental with “observational” data, we analysed the heterogeneity and used this information to downgrade the quality when necessary, as the GRADE approach gives the opportunity to be explicit also at this point.

**Comment 3:** “Table 2 (and 3): Footnote 6 (and 1 in table 3): GRADE does not recommend to automatically downgrading for inconsistency just for the reason of having only one study. Consider revising.”

We have revised this issue and finally we did not downgrade for inconsistency when having only one study.

**Comment 4:** “Table 3: Footnote 2: After describing the limitations, the authors note that the limitations were not serious enough to warrant downgrading. However, the profile shows that the quality was downgraded one level. Please clarify”

Sorry. This was considered a serious enough reason to downgrade. We have modified the wording.

**Minor Essential Revisions:**

**Comment 5:** “Page 13: “Moreover, in case series, it is reported that a high percentage of corticodependent patients are free for corticosteroids after adjunct apheresis treatment…””

We have eliminated this and changed the order of the references.
**Comment 6:** “Table 4: S4: The authors state that one strength of the GRADE system is its “quantitative assessment of the quality of studies”. However...”

Changed (see manuscript)

**Comment 7:** “Table 4: W2: Not sure what the authors mean with not a direct method establishing recommendations”.

With this sentence we wanted to express that the GRADE approach explicitly provides the possibility to “modulate” the strength of the recommendation on the basis of other issues besides the level of evidence found. Although the SIGN approach already began to include this process through the "Considered Judgment", we believe that issue has been improved in the GRADE approach.

**Comment 8:** “Table 4: W4: “In continued development: there is not [the word not is assumed– it is not correctly displayed in the pdf version of the manuscript] a final version yet”...”.

We have changed the wording to: “Some elements continue to be developed”.

**Comment 9:** “Table 4: T1: Threats: “Difficulties with new technologies: low number of studies, heterogeneity, unsuitable outcomes...” Not sure...”.

We have added the definition of Strengths, Weaknesses, Opportunities and Threats in a SWOT analysis to the “Material and Methods” section. In this case, Threats are external elements that might restrict the use of GRADE in the context of new technologies; this is external to GRADE, which is why it is not considered a Weakness of the GRADE approach itself.
Discretionary revisions

Comment 10: “Page 11: Instead of stating: “using apheresis systems instead of corticosteroids seems to have the same efficacy…” would suggest to state there were “no differences in efficacy”.

Changed

Comment 11: “Table 1: Some overall judgments of the importance of outcomes are not entirely intuitive. For example, if researchers R1 to R5 judge the clinical remission at 12 months…”

Footnote added.

Comment 12: “Table 2: Would recommend changing the heading from GRADEpro table to the suggested GRADE terminology: “GRADE evidence profile”.

Wording changed.