

**Draft minutes GRADE Working Group meeting Singapore
15 October 2009**

Participants: Alessandro Liberati, Alessandro Montedori, Benjamin Djulbegovic, Brice Kitio Dschassi, Elie Akl, Gordon Guyatt, Gunn Elisabeth Vist, Hilda Bastian, Holger Schünemann, Ian Shemilt, Ilkka Kunnamo, Iosief Abraha, Jan Brozek, Jenny Doust, Jeremy Howick, Jörg Meerpohl, Luke Vale, Mark Helfand, Mona Nasser, Nancy Santesso, Paul Glasziou, Philipp Dahm, Piet Post, Regina Kunz, Shahnaz Sultan, Shannon Sullivan, Susan Norris, Taixiang Wu, Ton Kuijpers, Vijay Shukla, Tahany

Regrets: Davina Gherji, Joao-Paulo Souza

1. **The minutes from Edinburgh** were approved without comment.

2. **GRADE membership of industry employees**

The issue had been raised by a member who is moving from an academic position to work for industry, is it ok to remain a member of the GRADE Working Group?

The group discussed the issue and agreed that there is no problem with membership and participation, but if a conflict is perceived they should let the group know.

We were reminded that we agreed in Rome that we would have no sponsorship of meetings, and workshops to the industry by members are left to the individual.

3. **JCE GRADE series – new and revisions**

Gordon wants to submit the first four papers for publication within a month – all suggestions for changes should be sent to him before then.

JCE article 1, Introduction – GRADE evidence profiles and Summary of Findings Tables - revised

Philip suggested that although the article is well written, it could benefit from being more enthusiastic. He also suggested moving the limitation and modifications further back in the article. It was also noted that the SoF table (and article 9) has different layout, it was suggested to try to harmonize and make consistent.

Paul suggested specifying in the quality rating of evidence that this is for interventions (p11), “For interventions RCTs start at high quality evidence”. It was agreed to replace with the Table from BMJ2 article to give more information about the up and downgrading. The terminology should reflect agreement with RoB- study limitations as risk of bias. Jorg volunteered to help Gordon with consistent use of terminology in the articles.

Action: Jorg

The group discussed the use of different GRADEpro versions of presentation. We do not know the optimal way to present information, (this optimum probably do not exist) – therefore Gordon would like to see different approaches presented. We agreed to say in paper one about the SoF table that we do not know that this is the ultimate format and that further articles will present different versions. In article 1, the current format produced by GRADEpro will be the one used in Table 4.

Gordon strongly thinks that the current presentation of AR in the SoF table is clearly suboptimal. However, the evidence we currently have points to the current way of presentation, and this is what has been agreed with Cochrane.

Holger reminded us that issues we discuss should not be focused around what the GRADE profiler currently do, but what we want it to do!

JCE article 2, Framing the question – revised

Ton pointed out that the article is about the outcomes and the title is about framing the question. He suggested changing the title to: framing the question and deciding on the important outcomes.

Mona suggested using the HMLIC terminology for different resource settings. And she suggested separating what is happening in the literature from what we are suggesting to do. Mona will write a suggested sentence to reflect the differences and not between the systematic review authors and the guideline producers and suggest where to place it.

Paul will provide a sentence/ paragraph about prognosis questions that it may just be a PO question

The group discussed including the issue of multiple comparators, and that the importance of the different framing of the multiple comparisons reviews may have different implications for the results and interpretations, i.e. specify whether comparators are in competition or a broad group of comparators. Vijay will send a sentence to reflect this. The group discussed and agreed to remove the breast cancer example.

JCE article 3, Quality of evidence introduction – new

This article still needs an abstract.

Ilkka would like an example on p14 for indirectness of patient important outcomes next to the definition.

The group discussed whether or not to include an example of an organization that includes expert opinion in the hierarchy. We agreed that it was important to make our position absolutely clear, and to do so without naming anyone.

Holger suggested to clarify the up and downgrading rules on p1 and that it might be useful to have a general comment somewhere that this is not only about guideline production. This is a long paper and if needed, the section on p3 could be deleted as it is covered later under risk of bias. He also suggested including a citation for use of quality on p7. The criteria for RCT are cited repeatedly, he think it would be helpful to also include risk of bias issues for observational studies. The GRADE web pages could be cited on p19.

Table 2 needs circulation and further discussion.

Hilda suggested that example 2 on p10 about chronic pain should be specified, for example for which part of the body, for topical remedies for pain relief you expect a short term effect only (placebo works 30% of the time).

Brice suggested making a clear statement regarding the overall quality of the evidence that systematic review authors are not supposed to make overall quality judgements – only for guideline developers.

JCE article 4, Rating the quality of evidence – risk of bias – revised

Piet suggested clarifying in the abstract that all studies are initially considered to be good quality studies that will be downgraded for limitations if it turns out that they are not.

Elie suggested adding reporting on p4 first sentence for both RCT and observational studies

It was discussed if reporting bias might increase confusion of internal and external validity. Spell out when publication of studies and when selective reporting of results from trials and selective reporting of outcomes is part of study limitations (internal risk of bias).

Gunn suggested changing wording from concealed randomization to allocation concealment, blinding not an issue for all cause mortality.

Paul suggested referencing the stopit trial for the early stopping issue. He also said that here is new empirical evidence by Wood, Altman regarding lack of blinding and subjective and objective outcomes. The influence of lack of follow up should also to be referenced.

The group agreed to remove section on Nof1 trials.
ITT and failure to adhere to it will be included in the abstract and article, also in Table 1, non-inferiority trials and per protocol analysis will also be included.

Action: All, send comments to Gordon on papers 5 through 9.

Articles 5 to 9 will be discussed in Rome

JCE article11, Resource use – revised

Ian took specific notes from this discussion to share with Massimo.

This paper was reported to be difficult to read because of some confusion about who the target audience is. This article is for the same audience as the other papers. It was considered that it would be helpful with more examples to more of the points. The introduction could start with the point that resource use is an outcome.

This article is about rating the quality of evidence for resource use. Discussions regarding modelling should be removed.

GRADE recommends to focus on resource use and to keep models out of the evidence profiles. Modelling in relation to decision making is a separate issue. Some modelling may be necessary when moving into recommendations.

We agreed to split this paper about resource use into two, this one that focus on the quality of the evidence, and a separate paper about recommendations where some discussion about models will be included.

Luke and Ian reminded us that NICE is using a table that looks similar to a GRADE evidence profile that incorporate models. We discussed this at a previous meeting and were clear that those tables are not GRADE evidence profiles.

Ian suggested that we use language consistently to distinguish difference between use of services and resources, to understand when it is referred to values and when to cost.

4. The research agenda for GRADE

At the Edinburgh meeting there was a suggestion to apply for funding from EU, Andy has taken the lead together with Shaun Treweek and written an application. The call is about better understanding for dissemination, presentation and implementation.

Any feedback on the proposal, please send directly to Andy.

5. The roles of mechanism

The group had a short discussion including the examples in the paper. Theories of mechanisms have been and are considered by some when there is a lack of evidence or only indirect evidence or surrogate outcomes are available. The EMB disapproval of mechanistic theories relates to lack of evidence, not the idea of mechanisms.

We may want to reconsider the issue of mechanisms in relation to generalizability.

6. Example for associations (exposure - outcome in observational studies) and how to grade the quality of evidence

We will continue to look for good examples.

7. Applying GRADE in (the osteoporosis) guideline developing

Piet reported that they set out to use GRADE for a guideline in the Netherlands, but found challenges with the use of old reviews, poor quality reviews and studies not assessed correctly.

These challenges are not challenges of GRADE – they are challenges that would be the same if you used another method! This is not a GRADE discussion.

8. Peer review/quality assurance process regarding GRADE profiles – Piet (attachment 18)

We are working on developing a GRADE evidence profile repository; this is the time to discuss peer review mechanisms and quality assurance processes.

Philipp will establish a task force together with Piet and think about the processes of how the peer review would work. Other volunteers to be included were Jan, Nancy, Elie, Gunn, Piet, Ton, Jorg (Yngve).

Action: Philipp & Piet

9. Branding of GRADE – checklist for minimal criteria of using GRADE

Tabled for discussion in Rome

10. Mixed Treatment Comparisons and GRADE

Vijay reported from initial experiences at CADTH considering mixed treatment comparisons (MTC), the question is how to apply GRADE to these data.

The group discussed some of the challenges related to MTC. It is a bayesian method that maybe should not be mixed with frequentist approach. What are the reasons for using indirect evidence when direct evidence is available? Inconsistency could you end up as no inconsistency even when there is considerable inconsistencies because the model eliminates inconsistencies. What is the validity of the MTC model? The issue is how much to downgrade, but some statisticians are arguing to not downgrade for this type of indirectness with MTC available. However, statisticians work on the assumption of a random sample and availability of a random variation of comparisons, but these are all drug trials and published based on careful decision!

There are a number of issues to consider on how to guide users and we need to work this through methodologically before we make GRADE recommendations. Gordon, Mark, Ben, Jenny and Holger volunteered to work on this. A project is ongoing with Milo Puhan from Johns Hopkins who has showed interest in this topic.

Action: Gordon & Vijay, Mark, Ben, Jenny and Holger

11. Translating GRADE into plain language

Ilkka presented the language they have chosen to present evidence statements.

We reminded ourselves of the groups' previous discussions on the issue, and the two studies lead by Elie. Some of this may also be part of Andy and Shaun's EU proposal.

Holger, Alessandro, Elie, Hilda, Gordon and Mona volunteered to work with Ilkka on this.

12. Criteria or "Do nots" for going from low or very low quality evidence to strong recommendations

13. Regret and values

The group discussed and agreed it might be relevant in the move from evidence to recommendations paper. Ben will send the note to Jeff who leads the writing on that paper.

14. Publications, workshops, applications

- Philipp is working on a grade table for European Journal of Urology and has a grant to create profiles. He is coeditor of EMUrology and will encourage authors to use GRADE
- Jorg, workshop next year in Freiburg together with Holger, Regina and Yngve
- Mark, Annals has just published a couple of papers on risk factors that could be added to the framing issues. A year or so ago the editors of Annals of Internal Medicine had show of hands on what is and what is not GRADE. The editors took it out and called it modified GRADE.
- Jan is working on a guideline on cow milk allergy, some of them with GRADE methodologies
- Ben, a paper on frozen plasma using GRADE is finally submitted. Grade grid helped to get consensus
- Elie, workshop in New York with Jan and Nancy
- Hilda, IQWIG is still not using GRADE, there are issue of directness, the selection of outcomes ahead examples were useful
- Piet will hold a workshop at GIN in November with Ton and Hans deBeer
- Ton have started a group to discuss the use of GRADE in The Netherlands
- Ilkka, use grade for recommendations
- Holger, the Institute of Medicine and practice development guideline have put together an expert panel where a presentation of grade is on the agenda. CDC for one of the immunization guideline may be using GRADE after a workshop in September. Holger will hold a session on GRADE at GIN. Four societies on respiratory will use GRADE for recommendations.

15. Future meetings

- 22-23 January 2010 Rome. Marina Davoli is the local organiser. The idea as usual is to have a local meeting before or after the meeting.
- 23 October 2010 Keystone
- Other? Florida with Philipp and Ben?