

Draft minutes GRADE Working Group meeting Freiburg, 8 October 2008

Participants: Iosief Abraha, Susanne Allander, Elie Akl, Pablo Alonso, Hilda Bastian, Heather Bennett, Jan Brozek, Massimo Brunetti, Klara Brunnhuber, Bernard Burnand, Steve Caldwell, Yngve Falck Ytter, Signe Flottorp, Bo Freyschuss, Gerald Gartlehner, Carlos Jimenez Gutierrez, Davina Ghera, Paul Glasziou, Gordon Guyatt, Robin Habour, Katharine Jones, Michelle Kho, Brice Kitio Dschassi, Regina Kunz, Stefan Lange, Monika Lelgemann, Alessandro Liberati, Joerg Meerpohl, Alessandro Montedori, Ignacio Mora, Mona Nasser, Susan Norris, Boguslawa Osinska, Andy Oxman, Silvia Pregni, Anne Rutjes, Nancy Santesso, Holger Schünemann, Vijay Shukla, Ivan Solà, Airton Stein, Jane Thomas, Gunn Vist, Vivian Welch, John Williams, James Woodcock

Regrets: Benjamin Djulbegovic, Rita Horvath, Roman Jaeschke, Nicola Magrini, Tasnim Sinuff

1. **The minutes from Rome** were approved.

2. **BMJ GRADE series**

BMJ will not make PDFs of the long versions. Pablo will make these.

Action: Pablo

The series is now open access.

BMJ is unable to do anything about all of the contributors not being listed in all of the articles in PubMed.

BMJ is not able to improve linkages between the articles in the series because of limitations in their publication system.

Copies of the BMJ articles are available in Portuguese.

JCE GRADE series

3. **JCE Article 1, Introduction and GRADE evidence tables**

There was a discussion about the usefulness of Figure 1 on how GRADE fits into the guideline development process. We agreed to keep it, but that Figure 1 needs modifications, Holger and Andy volunteered to make the modifications

Action: Andy and Holger

We agreed that it is helpful to highlight some of the other issues like selection of panels, judgement processes for importance, guiding panels and so on.

It was suggested to make a clearer distinction between rating the quality and strength of recommendations. We also discussed the need to let readers know that we are still meeting and developing the GRADE approach as the science evolves.

A suggestion was made to include something about the ethical issues related to GRADE. Airton Stein volunteered to draft a paragraph on this.

Action: Airton

4. **JCE revisions**

There were no comments on the list of articles.

We split into small groups for discussions of the remaining articles.

Article 4, Rating the quality of evidence – risk of bias

A suggestion was made to emphasize even more that the focus of the rating is for each outcome (rather than each study). The small group discussed possible problems with Cochrane risk of bias (RoB) figures, which give equal weight to each study. Holger and Julian are exploring ways of improving these figures.

Action: Holger

This small group also discussed ways in which case series could be upgraded; for example, for colon cancer screening there is no risk of puncture without colonoscopy. This evidence could potentially be upgraded based on there being a very large effect.

This small group also discussed article 6 (inconsistency) and suggested adding more examples of different types of inconsistency. Several suggestions were made for changing the title of article 4, including: “Study limitations and publication bias” and “Risk of bias and publication bias”. There was some discussion about whether we should use “risk of bias” in place of “study limitations”. It was agreed that we should continue to use both terms for a while.

Article 5, Rating the quality of evidence – imprecision

A suggestion was put forward to make the distinction between judgements about the quality of evidence for systematic reviews and guidelines clearer. There are differences in judgements about imprecision and indirectness.

There was some discussion relating to the guidance on what a clinically important difference is, the need to determine this in relationship to other outcomes, and the fact that it may differ across settings. This presents major challenges for making judgements about imprecision in systematic reviews.

It was noted that this article did not include examples where it we would downgrade two levels. Gordon asked all to send him good examples.

Action: All

Article 6, Rating the quality of evidence - inconsistency and indirectness & Combining direct and indirect evidence

There are different types of indirectness. Paul has been trying to clarify the terminology that we use to reduce confusion about different types of indirectness. The small group discussed how to deal with judgements about inconsistency and downgrading when there are different types of evidence (e.g. direct and indirect, different types of indirect evidence). When it is clear that one type of evidence is higher quality than other types, only the higher quality evidence should be included in an evidence profile or SoF. Other evidence can, if appropriate, be comment on in footnotes or an appendix.

We discussed how to deal with empirical evidence of a mechanism (surrogate or intermediary outcomes) when it provides supporting evidence (that is consistent with the direct evidence for a patient-important outcome) or reduces our confidence (because it is inconsistent with direct evidence). It was agreed that plausibility (theoretical mechanisms) is in the eye of the beholder and cannot easily be used as a criterion for the quality of evidence, although this is an underlying assumption for undertaking a review or developing a guideline. There are, for example, arguments that homeopathic effects are not plausible (based on our understanding of chemistry) and that it is plausible (based on laboratory studies that demonstrate purported mechanisms). We discussed downgrading evidence when the mechanism is not known, but agreed that this would be difficult to use as a criterion.

We discussed whether the quality of evidence should be downgraded if there is inconsistent evidence of different types (e.g. from trials and observational studies). It was suggested to add a paragraph to the article that addresses downgrading for inconsistency from different types of evidence.

Action: Gordon

We discussed coherence and the sigmoidoscopy example in the article. Because there are other plausible explanations for the difference in distal versus proximal cancers, the argument for upgrading the quality of evidence for this example (a single case control study) was considered weak. There was also concern about a criterion for upgrading based on coherence might be prone to misuse; i.e. people are quite clever at finding biological explanations to support almost any finding. Additionally, we often have a limited understanding of mechanisms. We agreed that coherence should not be used to upgrade.

We need more examples of different types of indirectness. All were asked to send these to Gord.

Action: All

Article 7, Upgrading and summarizing the quality of evidence

Article 8, Preparing summary of findings (SoF) tables & Guidance for presenting continuous outcomes

Article 10, Special challenges - resource use

These papers were not discussed in any of the small groups.

5. JCE article 9, Special challenges – diagnostic tests

It was noted that there still are methodological issues that need further work, including more detailed guidance regarding assessment of the risk of bias, assessment of the risk of publication bias, and judgements about imprecision and inconsistency. Many of these issues are similar to the challenges we have providing detailed guidance for grading the quality of evidence for treatments (or other interventions) and it was argued that the framework as far as it has been developed is helpful, despite these challenges and more limited experience with grading the quality of evidence for diagnostic tests.

It was suggested that the paper should address situations where a new test is more accurate than the reference test. Many good suggestions were also made at the meeting we had in Rome, which still need to be incorporated into the article, which has not yet been revised.

Action: Holger et al.

6. Good practice recommendations

It was suggested that we should call “good practice” or “motherhood” recommendations “Recommendations that are not graded”.

We need examples and guidance regarding when it is appropriate not to grade a recommendation and when it is inappropriate not to grade. It was suggested that a reason should always be given for each recommendation, including ones that are not graded. This can help panels to think through why they are making each recommendation.

Action: All

7. Grading the evidence for comorbidities

We discussed the paper on grading the evidence for comorbidities by James and Klara. It was felt that the issues that were addressed were a specific example of general challenges, that the

suggested approach for upgrading evidence for co-morbid conditions was not needed, and that the arguments were not compelling. It was suggested that the approach that is generally advocated for subgroups would work better; i.e. assuming that the overall evidence applies and downgrading the evidence if there are reasons for doing so (compelling arguments why the overall evidence should be considered indirect).

A general concern was raised about adaptations of GRADE, such as the one proposed in this paper, being published by members of the GRADE Working Group. We reconfirmed earlier discussions that modifications, such as the one proposed in this paper, should be discussed with the Working Group and that there should be a consensus regarding acceptable modifications that the Working Group supports (e.g. such as the ACCP decision to use three instead of four grades of quality). It was agreed to add a statement about modifications of GRADE to our web pages, such as: "Some organizations in their enthusiasm to use GRADE, have modified the GRADE approach. We recommend against such modifications because the elements of the GRADE process are interlinked, because modifications may confuse some evidence and guideline users, and because such changes compromise the goal of a single system with which clinicians, policy-makers and patients can become familiar."

Action: Yngve

This issue will be discussed in article 14 in the JCE series.

Action: Holger, Gordon et al.

Because there was limited time to discuss this paper it was agreed that it could be put on the agenda for the next meeting and discussed further.

Action: James and Klara

8. Grading health promotion guidelines

We discussed whether we should respond to an article (attached to the agenda) on grading of evidence for health promotion. This was of concern to WHO since the third author works at WHO and concerns have been raised about using GRADE because much of the documentation regarding health promotion is low quality evidence based on the GRADE approach. No one at the meet felt compelled to write a response.

9. Cochrane Summary of Findings (SoF) tables

There is a large demand for training for Cochrane SoF tables. A group working with Holger is coordinating training efforts and will be looking for people with experience and methodological expertise able to provide support in different regions.

Action: Holger

10. GRADE Profiler

Minor changes have been made in updated versions, but no substantial changes have been made in the software since the previous GRADE Working Group meeting in Rome.

Regina noted that it is very helpful to have paper copies of the help file at workshops.

11. Database of evidence profiles

Yngve has made the database accessible at www.gradeprofiles.org. The database is searchable. Importable files can be downloaded and edited, but cannot be resubmitted. Linking can be arranged. We would like to make PDF files or some other format available to ensure that the profiles can be easily viewed in the correct format. Metadata (that is not already included in GRADEpro) would be helpful, but would require more work when profiles are uploaded.

If the database is going to be open access, we need effective quality control mechanisms. We should negotiate with GIN for them to put their examples in the database and for GIN to agree to open access. An option might be to have a short-term protected area where profiles can be uploaded prior to quality control mechanisms and moving them into an open access area.

Regina will discuss this with GIN and Holger will discuss this with Cochrane (Lorne Becker).

Action: Regina, Holger

All should send examples to Yngve, and give Yngve feedback

Action: All

12. GRADE website

Currently anyone wanting to join the Working Group can simply email Yngve, be added to the discussion list, and become a member of the Working Group. It was suggested that we might want a more transparent method, but nothing was decided. We had a brief discussion of forums, blogs, public and members areas, and Facebook. For many of these we would need an active moderator. It was agreed that these options warrant further discussion at the next meeting.

13. Publications, workshops, applications

- Ben is helping the American Association for Blood Banking (AABB) and using GRADE in a guideline on FFP. He is trying to introduce the production of a GRADE evidence profile as a pre-requisite for students to graduate from an EBM course.
- Rita has been asked by Clinical Chemistry to write an article based on the GRADE BMJ diagnostic article. She is involved in an American diabetes diagnostic guideline group for NACB (in collaboration with ADA) and trying to convince them to use GRADE.
- Nicola has started to use GRADE for the application of new drugs for the WHO essential medicines list. They are running 3 and 2 days introductory courses.
- Alessandro is involved with several guideline groups using GRADE to make recommendations regarding anticancer drugs and drug addiction. GRADE will be used in a national project on recommendations for anticancer drugs. He is also collaborating with WHO. He is organizing a training workshop for a network of people in 2009.
- Jan is writing a series of three papers on GRADE for Allergy.
- Signe gave a GRADE workshop at the GIN conference together with Regina and Gunn, which was a success.
- Michelle is studying GRADE in her PhD thesis
- Bo reported that SBU in Sweden is in the process of switching to GRADE.
- Gordon reported on possible plans to translate GRADE materials into Japanese.
- Holger and Yngve led a one-day course and workshop for AHRQ. Other facilitators in that workshop included Ben, Mark, John, Signe, Elie and Gunn.

There were more reports after this, but Gunn had to leave and asked Michelle to take notes, which have not yet been incorporated in the minutes..

14. Future meetings

- SIGN will host a meeting in Edinburgh 7th & 8th May 2009.
- Cochrane Colloquium 11 to 14 October 2009, Singapore full day meeting the day before or the day after. Aim for the 15th October.
- Amsterdam or Utrecht January 2010.