

## **Draft minutes GRADE Working Group meeting 28 October 2007, Sao Paulo**

**Present:** Elie Akl, Hilda Bastian, Jan Brozek, Jonathan Craig (half day), Patrice Dosquet, Bernard Ewigman, Gordon Guyatt, Julian Higgins (half day), Brice Kitio, Regina Kunz, Anne Lethaby, Erin Morris, Susan Norris, Andy Oxman, Vivian Robinson, Nancy Santesso, Holger Schünemann, Vijay Shukla, Gunn Vist

**Regrets:** Cindy Farquar, Mercè Marzo

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

### **2. Grading the quality of evidence when there are no events**

Two opposing views were put forward, one that there is no point in grading the quality of evidence when there is nothing to grade because there are no estimate. The other point was that if there is evidence of death or a severe outcome being a rare event, this is important information and it is possible to estimate the absolute effect.

The example of compression stockings to prevent deep vein thrombosis (DVT) on long flights was used as an illustration. The studies included approximately 3000 people and there were no deaths. It was suggested that although there is very low quality evidence (or no information) for the relative effect, there is high quality evidence for the absolute effect (unless it is downgraded for imprecision).

Gordon suggested that where the events are very few, we could base the judgement about the quality of evidence (particularly precision) on the absolute effect (making this clear in a footnote). Another option would be to use observational studies for the illustrative risk without stockings<sup>1</sup> and use the risk ratio for symptomless DVTs to estimate the risk of symptomatic DVT with stockings downgrading for indirectness (a surrogate outcome).

We need more examples, and should come back to this again.

This will have to be changed in the Handbook, and addressed in the JCE series of papers.

**Action: Holger, Gordon**

### **3. Cochrane Summary of Findings**

- how to report not measured, not pooled, not reported, 0 events

Sometimes when studies are not pooled, the range may be helpful.

Currently GRADEpro offers the following options, which determine how the SoF table is constructed: pooled, not pooled, not measured, not reported. It is not possible to edit the text that is automatically entered into some cells, as determined by these options (e.g. “See comment”, “Not estimable”). This can be a problem, for example if reviewers want to enter a

---

<sup>1</sup> Philbrick JT, Shumate R, Siadaty MS, Becker DM. Air Travel and Venous Thromboembolism: A Systematic Review. *J Gen Intern Med* 2007; 22:107–114.

range of estimates when results are not pooled. The GRADEpro team will consider ways of addressing this.

**Action: Holger, Jan**

- risk of bias and GRADE terminology

The GRADE terminology is Study limitations: None, Serious and Very serious limitations. The Cochrane risk of bias is Low, Unclear and High risk of bias. For now no changes will be made in the terminology that is used, but the relationship between these terms will be reconciled in the Handbook.

**Action: Julian**

It would have been good to add a column to the risk of bias table to make explicit judgements about whether there is a low or high risk of bias (serious or very serious limitation) when a criterion is not met, but it is too late to make this change in RevMan 5. These judgments should, nonetheless, be made explicit in footnotes of evidence profiles prepared in GRADEpro.

- presentation of standardized mean difference (SMD)

The Cochrane Handbook includes some suggestions for how to present SMDs. Alternatives include back translation to an understandable scale, conversion to a dichotomous display and the use of rules of thumb for what constitutes a small, moderate and large effect. The Cochrane Statistical Methods Group is still wrestling with how best to present SMDs. It is suggested to present the SMDs for now.

Currently SMDs are imported from RevMan into the SoF tables as “See comment”. This will need to be changed, but we first need to develop clear guidance and agree on how best to do this.

**Action: Andy, Holger, Gordon**

#### **4. Guidance for judgements about study limitations paper (4). “Rating the quality of evidence – risk of bias”**

There will be a GRADE series of 12 papers in the Journal of Clinical Epidemiology. The target audience for these is guideline developers and systematic review authors, and they are intended to provide detailed guidance about how to apply GRADE. The same text and examples will be adapted and used in GRADEpro help files.

Suggestions that were discussed included: The figure captions in the paper should be less positive and include more about alternative explanations. Something should be added about time to publication in the section on publication bias. Something should be added addressing the possibility that the lack of comprehensive search strategy in a review may increase the risk of publication bias affecting the results.

We agreed that Figure 1 (illustrating judgements about the risk of bias on a continuum) should be removed from the paper.

**Action: Gordon**

**5. Guidance for judgements about inconsistency and indirectness paper (6)  
“Rating the quality of evidence – inconsistency and indirectness”**

It was suggested that the guidance for considering confidence intervals should be clarified, emphasising that the importance of differences in results depends on where the estimates are in relationship to no difference (and when the differences would lead to different recommendations). This should be illustrated in a figure. The problem with focusing on inconsistencies in relative effects when absolute effects are considered in making recommendations or decisions was also discussed.

**Action: Gordon**

**6. Report on status of BMJ series of articles**

All five papers have been submitted, as both long and short versions. The comments made by peer reviewers about the first paper were very positive. Peer reviewers' comments on the other papers had not yet been received. Andy will contact BMJ about progress on the other papers.

**Action: Andy**

**7. Revised outlines for new series of articles + authorship**

A draft outline was previously circulated and revised. We need lots of examples, particularly ones that illustrate specific challenges or difficulties that people have.

**Action: All**

Both Susan and Jan think that having a visual of how it all fits together is helpful, they have made suggestions that they will send to Gordon.

**Action: Susan & Jan**

It was suggested that the first paper of the new series start with an easy overview and plan for the whole series. Andy suggested putting a Summary of Findings (SoF) table in into the first paper – to show the endpoint at the beginning. For the very beginners at grading with GRADE, they will be directed to the BMJ series and those will not be repeated too much.

We had a discussion about the 9 point scale for importance of outcomes. These are value judgements and the strength of it is that it is very helpful to panels for reaching consensus about the importance. And the importance scale has worked very well.

A suggestion was also made to adapt some from Holgers paper on how people come to consensus,- how to get the panel through 25 odd guidelines in a couple of days?

Authorship proposal for the new series

In the first of the papers, the GRADE working group in by-line and then list everybody else who have contributed to the development. The following papers with everyone who has made contribution in the by-line on behalf of the GRADE working group. This proposal was accepted, we will ask JCE for approval of this.

**Action: Gordon**

We agreed on a low threshold for being an author of papers in the series. The criterion to become a member of the GRADE Working Group is to say that you want to. Anyone who

contributes to a paper, including by providing an example that is used, will be an author listed on the by-line.

**Action: All**

**8. Should guidance for upgrading based on the magnitude of effect be different for HR and OR than for RR?**

We agreed that this should be made transparent in the JCE article that addresses upgrading.

**Action: Gord**

**9. Ungraded recommendations**

We agreed that we should include some guidance about ‘good practice’ guidelines in the JCE series. It may, however, be a challenge to define these. Jonathan argued that this is potentially a major issue because ‘good practice’ recommendations may be regarded as important and some may use this as a loop hole to make recommendations without looking at or basing them on evidence (when they should). He suggested we should stay away from endorsing ‘good practice’ recommendations. Andy argued that this problem needs to be addressed explicitly because it is already widely done, sometimes appropriately and sometimes not. We agreed to look at a sample of good practice recommendations as a basis for developing guidance.

We need examples of ‘good practice’ guidelines.

**Action: All**

**10. Practicalities and shortcuts – guidance on acceptable and unacceptable modifications**

Some people that want to use GRADE find it too complicated or too time consuming and they make short cuts and modifications. We discussed what the minimum standards are that must be met for us to agree that GRADE was used. ACCP lumped low and very low quality and we previously discussed and agreed that that is OK. NICE uses GRADE for the quality of evidence but not for recommendations. We previously discussed whether producing an evidence profile should be considered a minimum standard. According to the minutes from Bilbao, which were attached to the agenda: “We discussed and agreed that for organisations to report that they use GRADE they should produce an evidence profile.” There was however disagreement with this and some recalled that we agreed that the word “ideally” should have been inserted between should and produce.

We agreed that in order to state that GRADE was used, the GRADE definitions of quality of evidence and strength of recommendations should have been used. Moreover, the quality criteria that we suggest should have been used without substantive modifications, deleting or adding criteria.

We agreed to review examples of modifications that people have made as a basis for discussing ones that we would find acceptable and unacceptable. A simple search for examples of uses of GRADE with modifications was suggested.

**Action: Gunn**

## **11. Evidence profiles for observational studies**

Holger explained the attachment and the problem with preparing evidence profiles for case control studies, where there is not a comparison between groups that do and do not get the intervention. Additionally, guideline panels may want information about the included studies that is difficult to incorporate in evidence profiles.

It was suggested that the odds ratio from case control studies can be used the same as other estimates of relative effects to derive illustrative risks for relevant assumed risks without the intervention. It would be confusing and difficult to use different formats for evidence profiles that include different types of studies.

Holger argued for including the number of events and total number of participants for the comparison groups, despite problems with these sometimes being misused and causing confusion. He will do some user testing. For now, it was agreed to stick with the format for summary of findings tables that was developed based on user testing for Cochrane summary of findings.

**Action: Holger**

## **12. Upgrading RCTs**

We have suggested that the quality of evidence should only be upgraded because of a large effect when there are at least two studies without serious limitations, focusing on observational studies. If an RCT has a serious limitation in relationship to randomisation it could be considered equivalent to an observational study, in which case it might be appropriate to upgrade because of a large effect. However, there is a risk of people inappropriately upgrading and people should be reminded that upgrading is infrequently justified for either observational or randomised studies. We agreed that it should be possible to upgrade randomised trials and that there should be a pop-up warning when people elect to upgrade in GRADEpro.

**Action: Holger, Jan**

## **13. GRADE profiler**

Holger reported that the programming is going well and a version ready for use is available. The help files are not complete, the software is still being tested, and some changes are still needed, included those discussed during this meeting. People are encouraged to use GRADEpro and provide feedback.

**Action: All**

It is planned to release RevMan 5 and the new version of GRADEpro in March 2008.

## **14. GRADE website**

When searching the web using “GRADE Working Group” in Google, our web pages come up on top. They come up third when searching for “grade”!

The latest version of GRADEpro needs to be moved onto the GRADE web pages.

**Action: Holger, Jan**

The members' pages will need to be revised in the future. For now they are not a problem and it is OK that the password is freely shared.

#### **15. Database of evidence profiles**

We have been waiting for the new version of GRADEpro to be ready before investing more time in developing a database of evidence profiles. Andy suggested a possible strategy for developing an initial set of quality assured evidence profiles in connection with launching of summary of findings tables in Cochrane reviews. He will explore this further.

**Action: Andy**

#### **16. Workshops: Europe, North America**

We need to come up with a way of offering more workshops in Europe, including to Cochrane systematic review authors. We need funding for this. We also need to develop better ways of communicating requests for workshops to members of the Working Group.

Some workshop material is available on the web pages. This should be updated.

**Action: Andy**

It was pointed out that people who are not members of the Working Group offer GRADE workshops, and agreed that there is nothing we can do about that, but hope they get it right.

#### **17. Publications and applications**

- Susan, Neurological Association told her that they did not like GRADE and will not use it. NIH is updating cholesterol and obesity guideline (previously consensus based) and Susan is hoping to get involved. There is a chapter in NIH Methods book about GRADE.
- Patrice, several French societies know GRADE and some are considering starting to use it.
- Holger, the European Respiratory societies has decided to use GRADE.
- Brice, workshop by Anaesthesiology about GRADE. Wants to translate it into French
- Hilda, IQWIG is planning to get people in to make GRADE evidence profiles for them, and they are planning to cooperate with the Finnish for developing patient information using GRADE. Guidelines for HTA are now recommending GRADE (without modifications).
- Jan, nothing new relating to GRADE lately in Poland. A lot of adaptation of other guidelines rather than producing their own. Starting work with a pragmatic GRADE approach with profiles when there is a systematic review available.
- Elie, is doing systematic reviews on different ways of reporting and presenting results, he hopes to have results ready by the next meeting.
- Andy, GRADE evidence profiles are being developed to decisions about maternal child health in low and middle-income countries as part of an EC funded project (SUPPORT).
- Holger, GOLD guideline producers are starting to use GRADE. Infectious disease workshops. NIH workshop. European Urology Society is officially adopting GRADE and has announced this in the Journal.
- Alessandro has published a paper in the Journal of Clinical Oncology about there use of GRADE for oncology guidelines.

- Regina, KDIGO global initiative has struggled with GRADE, has a paper in the American Journal of Gastroenterology with Yngve, and also one in a Haematology Journal.
- Gordon, KDIGO also provides excellent examples of what not to do.
- Gunn, Cochrane CAM field wants SoF tables.

#### **18. Future meetings**

- January 10-11, 2008, Washington DC, focus on resource use
- May 6-7, 2008, Rome, focus on diagnostic tests
- October 2008, Freiburg, before or after the Cochrane Colloquium

The current chair of GIN wants more formal collaboration with GRADE. The GIN meetings are often just before the Cochrane Colloquium (this year their meeting is in Helsinki, just before the Cochrane Colloquium in Freiburg). Regina and Andy are discussing with GIN the possibility of their providing support for GRADE meetings in conjunction with the annual GIN meeting.

**Action: Regina, Andy**

Thank you to Alvaro and the organisers of the Colloquium for providing the room for the meeting and refreshments!