

Draft minutes GRADE Working Group meeting Melbourne 26 October 2005

Attendants: David Atkins, Inge Axelsson, Jonathan Craig, Gordon Guyatt, Margaret Haugh, Sue Hill, Kent Johnson, Marissa Lassere, Ann Lethaby, Alessandro Liberati, Dianne O'Connell, Laura Maxwell, Holger Schünemann, David Tovey, Gunn Vist

1. Introduction and Agenda

We set the rules that all present would agree with the BMJ paper and we would not re-discuss any of these issues unless there were problems. Agenda was approved with the addition of partnerships of GRADE with other organisations under point 9.

2. Minutes from Rome

Correct spelling mistake on page 2.

With that change the minutes from Rome was approved.

Action: Gunn

3. Balancing benefits and downsides, making recommendations

We picked up on our discussion in Rome and discussed what a guideline developer with the profile to balance the benefits and harms would/ should do. Which are the other processes and issues that they need to work through? We worked through and amended the draft checklist / worksheet for the balancing.

Thrombolysis for acute stroke example:

In this example there are 111 events for benefits and 90 events for harm. David raised the point that given this potential for harm, even the wording of a “probably do it” would be too strong in his view. We discussed the issues around the Confidence Interval and the balancing issues and the wording we adopted in Rome (see below) for making recommendations. We reminded everyone that these are suggestions because guideline developers have asked for guidance on this issue, but that this is helpful and the different wording options would accommodate some of the concerns (suggest is better than probably do it for this recommendation).

We came back to the point that in most instances we want to make a recommendation, and only in rare instances would one not make a recommendation. We have abandoned the wording trade-off. We reassured David that the recommendation would be weak and alternatives would be equally acceptable.

David A pointed out a potential problem, the harm is subsumed under benefit because the harm is already included in those that die or become dependent. Add another footnote to include this.

Action: Holger

Confidence intervals around non-significant absolute measures – should they be given?

We agreed that even when the CI cross one, it provides useful information and we suggest that we continue to present both point estimate and CI even when there is no evidence of difference. This however is only true for natural frequencies as we agreed to present them in the GRADE profile. We would not adopt this approach for NNT where we would not provide an estimate.

Oral anticoagulation in acute stroke example:

Alessandro argued that the patients in practice are older and have higher risk than the patients included in the trials, so there is indirect evidence regarding patient population – he suggested that this be moderate evidence for the outcome major extracranial hemorrhage. We clarified that it depends on for which population we are making a recommendation. If we are making a recommendation for elderly this may be downgraded for indirectness, but we need to specify the setting and population in each GRADE profile.

Action: Holger

Equal weight of death from doing or not doing something

We also discussed the importance scale being present in the table as are intended to use to decide what goes into the table before the grading starts. Open to data driven re-evaluation at this stage. Holger sums up: First part, guideline panel should be transparent in how they make the importance rating, when it comes to balancing the benefit and downsides the panel would look carefully at the actual effect measure and becomes more data driven.

We decided before the Rome meeting that we need to give more guidance to users of GRADE about the process of balancing benefits and downsides. Therefore we discussed factors in deciding on a strong or weak recommendation.

We began by considering the table that was in the paper the ACCP published. In a prior email correspondence Andy raised concerns and suggested to add “factors that modify the expected effects in specific settings”. We agreed to add this in the table.

We worked through the ACCP article table and decided to reword a lot of the issues. We had a suggested table and were looking for examples and one sentence descriptors that David A suggested to call *recommended process*. We agreed that this would be a proper term. We worked through the table discussion in what order do we suggest that a guideline panel work through the issues: after some discussion we agreed on a sequence of steps (see table below) moving from summary of findings tables to recommendations (flowchart) and agreed on the following (we will need to work on examples):

Issue	Recommended process	Example
1. Quality of evidence (GRADE quality assessment with GRADEpro)	Strong recommendations usually require higher quality evidence for all the critical outcomes. The lower the quality of evidence the less likely becomes a strong recommendation	
2. Relative importance of the outcomes a. benefits of therapy b. harm of treatment c. burdens of therapy	Seek evidence about the relative values that patients place on outcomes and the actual value they place on them (critical, important but not critical, not important). Seek evidence about variability in preferences and values in patients and other stakeholders. It should be upfront that the relative importance of the outcomes should be included in the considerations before you make recommendations. If values and preferences vary widely a strong recommendation becomes less likely.	
3. Baseline risks of outcomes a. benefits of therapy b. harm of treatments	Consider the baseline risk for an outcome. Is the baseline risk going to make a difference? If yes, then consider making separate recommendations for different populations.	

c. burdens of therapy	The higher the baseline risk, the higher the magnitude and the more likely a strong recommendations	
4. Magnitude of relative risk a. benefits (reduction in RR) b. harms (increase in RR) c. burden	Consider the relative magnitude of the net effect. Large relative effects will lead to a higher likelihood of a strong recommendation if the balance of benefit, harms and burden go in the same direction. If they go in opposite directions and the relative magnitude of effects is large, the recommendation is more likely to be weak	
5. Absolute magnitude of the effect a. benefits b. harms c. burden	Large absolute effects are more likely to lead to strong recommendation.	
6. Precision of the estimates of the effects a. - benefits of therapy b. - harms of treatments c. - burdens of therapy	The greater the precision the more likely is a strong recommendation	
7. Factors that modify effects in specific settings/Local factors that may affect translating of the evidence into practice	The more similar setting and patients for which you are making a recommendation to where you obtained the evidence from, the more likely is a strong recommendation	
8. Costs	Consider that important benefits should come at a reasonable cost. The higher the incremental cost, all else being equal, the less likely is a strong recommendation in favour of an intervention	

4. Definition of strength of recommendation

The definition includes confidence that adherence to a recommendation will do more good than harm. This implies that it is impossible to make a strong recommendation when weak evidence.

Gordon reports that UpToDate had difficulties with the definition. He brought up two examples that are compellingly difficult, Holger's interpretation is that they want to make strong recommendations and that a solution to the problem would be to make a weaker recommendation. Sue agreed. We decided that we are not changing the definition at this point.

We were reminded that it is important to separate the quality of evidence and strength of recommendation. In Rome, we agreed to avoid probably and we confirmed that we still agreed with this decision.

Strong: Do it / Don't it = we recommend (should)

Weak: Probably do it / Probably don't do it = we suggest (might)

Translators do whatever they want anyway. But we should make suggestions and consistent translations for them to use if they want.

5. Quality of evidence for single RCTs

We agreed to postpone discussing this until the next meeting.

6. Cochrane Summary of findings (SoF) table and GRADE profiles

6.1. How should continuous effects be presented?

Julian Higgins joined the meeting and started by suggesting that the SoF table should be presented using more words to inform the information. Julian recommended that the raw data from the meta-analysis remain in the table

Clinical relevant cut point can be used to transform continuous data to dichotomous, but often only the means are available. There are several options for dealing with continuous outcomes but no solution is perfect right now.

1. SMD
2. meta-study
3. chose cut off and convert each study to risks and re do meta-analysis based on risks for each study

It was proposed that clinicians do not understand SMD well and that it is difficult to translate into something more interpretable. The other two options (dichotomization) are the ones we discussed in more detail. Julian points out that #2 and #3 will only work for normal distribution. Julian report from correspondence with Andy that he thinks that it is too early for these approaches (#2, #3) and that Julian agrees. He has worked on an example which turned out to be similar in terms of the results.

If one was to choose one of the approaches, one would need to have indication for normal distribution before starting this process. We discussed whether this assumption can be made and whether we would not be at too high a risk to be off with the estimates. Julian said that he is indeed not certain enough, that we need more examples (from GRADE) and we agreed to not go that route for now.

Additionally David A points out the challenge/ difficulty of deciding where is the cut point on each scale, which needs deciding before the conversion

We agree to continue using SMD for continuous results.

It was suggested that this would make a nice master student project (2-3 months, Cam summer students in epidemiology or stats post doc). Julian or others will look into this.

Action: all

6.2. Summary of findings table for Cochrane reviews pilot study

Gunn presented results from the pilot testing of the SoF table where Cochrane review authors has prepared SoF tables from their reviews. 20 reviews have been completed and submitted to us with associated SoF tables! Feedback has been collected through the answers in a questionnaire about the GRADEpro software and instructions that we administered after completion of the SoF table.

The pilot was effective trouble-searching and we received a lot of good suggestions for improvements of both the software and the instructions / help file. Results in short is that the participants spent an average of 4 additional hours on the review to prepare the SoF table and

that 85% thought that the SoF table would improve accessibility of the main results of their review.

Thank you to the members of the GRADE Working Group who were contact person for review authors during this process: David, Andy, Jacek, Katharine, Paul, Yngve, Jane, Signe, Francoise, Helena, Jeff, Holger and Gunn.

The Cochrane Applicability and recommendations methods group want to be involved in / make the recommendation on how the SoF table should end up looking.

7. GRADEpro tool. Introduction, development and integration with RevMan

Holger demonstrated how to use GRADEpro. And he informed about the discussions regarding future integration of GRADEpro into RevMan for the next version (5).

8. Use of GRADE by other organisation

- ACP are considering GRADE and USPSTF – they will decide on one
 - Dough McRory want to see examples from GRADE for this
- UpToDate is officially going to start grading the recommendation in February 2006 and use a modified version of GRADE (the ACCP version)
- ACCP endorsed the amended version with collapsing of low and very low quality and will use it from now on for all of their guideline projects
- American Endocrine Society is using it – the first guidelines using the approach are being published
- EBEuro urology continues to work with GRADE in order to build up a collaboration; Holger will present at a meeting to them on Dec 3 in Rome and is working on a grading guide and handbook for them
- Surviving Sepsis Campaign will use the GRADE approach (very likely) – Roman has drafted a paper based on the ACCP paper but not collapsing the low and very low quality level. Roman will work with this initiative
- ML may push with the Endocrinology association, Royal Australasia association
- Swedish guideline groups use the Oxford approach, not sure if interested (Inge Axelson)
- BMJ Clin Evidence (Holger went up to Clinical Evidence to speak to the editorial team in September) are trying it out
- J Bone and Joint Surgery – David Tovey wrote an editorial for the UK edition of the journal suggesting GRADE while the North American Edition supported a different approach – we need to clarify
- Sue Hill, tried GRADEpro but was discouraged by footnotes, otherwise liked it. Moving to include several of the issues.
- WHO have officially endorsed it “but are not using it”, they have a long way to go, and need the new improved version.
- Presentation of guidelines and recommendations, Holger also covered within the WHO project that Andy is heading
- Dianne O’Connell is involved with the metastatic prostate cancer guidelines and will try there, will also have to try Philippa Middeltons version
- Margaret Haugh, France political in house but interested, National Cancer institute.
- Kunnskapssenteret in Norway endorse the use of GRADE

- American Thoracic Society – we published an editorial in their three journals and officially suggested GRADE. Holger will chair a new documents development and implementation committee that will supervise the guideline development projects at ATS hopefully adopting GRADE; Roman and Gordon are members of the committee
- A Neurology S, committee documents
- Holger held a workshop at NICE (July 2005), and they are still interested; have tried the GRADE approach. Feedback was, that they want software – which some groups will now try

9. Partnership with other organisations

GRADE has established a good reputation; more and more interest is mounting in its activities

The chair of GIN, Günter Ollenschlaeger approached Holger to meet with him in Melbourne. He indicated that GIN wants to be associated and engage in a preferred partnership between GIN and GRADE. The problem is that GRADE is not institutionalized and so the question is with whom GIN would partner.

If there was a way to partner, there are benefits and downsides for GRADE from a partnership (several for GIN):

GRADE might benefit as new guideline groups would then be introduced to GRADE (GIN would be endorsing GRADE). However, GRADE has tried to be open and inclusive whereas GIN has a membership fee and allows for profit organizations to join (none has joined yet).

We discussed our options:

- 1) Become an organization
- 2) Become an official part of Cochrane (with which we are already affiliated) – the natural home would be the applicability and recommendations method group
- 3) Join GIN

We considered the potential benefits of these decisions (options 1 – 3 above):

- 1) Could apply for grants and would be able to form partnerships
- 2) Would have structure to operate from. GRADE would be facing less critics within Cochrane
- 3) We would have a home

We considered the potential downsides of each of these decisions.

- 1) too much bureaucracy and cost (registration and meetings we would have to hold) that we do not have resources for
- 2) Cochrane has critics – this could hinder us in advancing further (e.g. in US). We would have to obey the steering group. Cochrane in itself is not an organization
- 3) We would be under control of GIN. GIN is free to promote GRADE without partnership or closer affiliations anyways.

Any affiliation has the risk of the organization we partner with not living up to our standards. If we partner we will have to make it a term agreement. For example, for two years. If we partner with GIN we would ask for funds to fund one meeting a year. GRADE would have a representative on the GIN steering committee (although Margaret mentioned that this is not as easy as it seems; e.g. GIN would have to officially approve and it may be a no-voting membership. We would also have to select someone.

We decided that 1) is not a good option and that we will table this for follow-up discussion for another meeting in Lyon and that all members should read this and come up with an opinion and other arguments.

We favored option 2, but also think that we could remain as what we are. We will – after the Lyon meeting – respond to GIN.

10. Website (www.gradeworkinggroup.org).

Thank you to Yngve. Everyone thought this looked very nice.

Everyone send feedback

Action: all

11. Publications

- Editorial in ACP Journal Club – coming out
- ACCP publication including GRADE in Chest – coming out
- Editorial for BMJ - Holger and Nicola have started to work on this; we could provide an update similar to what we have in ACP JC
- Paper on diagnosis – coming from Holger for Lyon
- Urology paper – Holger in the works
- Surviving sepsis campaign – not sure whether official publication but one is out there
- UpToDate – grading paper – probably for the UpToDate website
- Summary of Findings tables – Gunn will write this paper

12. Future meetings

Lyon, France 7 & 8 December

We requested that everyone looks for funding. David Atkins suggested to submit an application for funding to AHRQ – Holger and Gordon to explore

In connection with the Cochrane Colloquium in Dublin? If so, we decided to have the GRADE meeting at end or before (not during the Colloquium as experience have shown this not to work well)

Action: all

13. Any other business

Diagnostic tests, quality criteria, reporting bias, suggestion to invite John Deeks to Lyon for discussing diagnostics.

Action: Margaret