Background

- The GRADE system is becoming widely used in guideline development.
- The full GRADE evidence profiles of clinical study characteristics and summary of findings are comprehensive and detailed and may not be appropriate for all guidelines.
- It is also necessary to consider ease of readability for the guideline developer and user of the final guideline.
- Consequently, different guidelines have adopted different formats to present GRADE evidence profiles.
Purpose

• To evaluate the variations of GRADE evidence profiles presented in NICE clinical guidelines, as well as guidelines produced by other developers.

Methods

• Preliminary search: NICE published guidelines that used GRADE from 2005 to 2010. A checklist was used to evaluate the variations of GRADE evidence profiles, including:
  – where the evidence profiles were presented in the guideline
  – formats of the evidence profiles in relation to study designs
  – data included in the evidence profiles
  – styles of the footnotes
  – others

• (Next step) A web-based search will be carried out to identify published guidelines (in English language) from other developers that have used GRADE.
Standard GRADE profiles

- From GRADEprofiler software
- Aimed to capture all information as comprehensive as possible?
- What were the rationales for the design originally?

Preliminary results

NICE clinical guidelines

- From 2001 to May 2010 = 99 published clinical guidelines by NCCs
- From 2005 to May 2010 = 76 published clinical guidelines by NCCs
- Guidelines that used GRADE (by May 2010) = 17
Preliminary results

Apart from the standard format, there are many other ways of presenting evidence profiles

1) Where and Format (1):

Substantial variations among NICE clinical guidelines

- 9 used separate ‘summary of evidence profiles’ (portrait) in the full guideline, with full GRADE profiles (7 x landscape; 2 x split-portrait) in the appendices.
- 2 used full GRADE profiles (split-portrait) in the full guideline.
- 1 used modified full GRADE profiles (portrait) in the full guideline.

Preliminary results

1) Where and Format (2):

- 1 only cross-referred the GRADE profiles number in the full guideline, with full GRADE profiles (landscape) in the appendices.
- 3 only presented full GRADE profiles (landscape) in the appendices, with no discussion in the full guideline.
- 1 only used ‘summary of evidence profiles’ (portrait) in the full guideline, with no full GRADE profiles in the appendices.
Preliminary results

2) Clinical areas and study design:
   - All were on intervention questions, only 2 were on diagnostic accuracy (Glaucoma and Lower Urinary Tract Symptoms)

3) Styles of the footnotes:
   - Most guidelines have footnotes listed below individual profiles, only 3 guidelines have added extra column in the profiles to document the footnotes.

4) Data included in profiles:
   - 2 types of evidence profiles:
     - Summary of GRADE profiles (in the full guideline),
     - Full GRADE profiles (in appendices, OR split or modified in the full guidelines)

Data included varied across different guidelines
   - Summary of GRADE profiles:
     - summary of findings
     - overall quality of the evidence (without information on the quality assessment).
   - Full GRADE profiles:
     - various formats, but captured all details from GRADE. Some also incorporated quality assessment from other system such as SIGN.
Some examples

1) Summary of evidence profiles (portrait) in the full guideline

<table>
<thead>
<tr>
<th>Study IDs</th>
<th>Psychological risk factors or sub-threshold symptoms</th>
<th>Trauma/infant/neonatal (psychological trauma/operative delivery)</th>
<th>Postnatal decline/length of stay</th>
<th>PTSD criterion A</th>
<th>Overall result (all studies)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Overall quality of evidence</th>
<th>Moderate</th>
<th>Low</th>
<th>Low</th>
<th>Low</th>
<th>Moderate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression below threshold</td>
<td>RR = 0.7 (0.54, 0.91) (K=7; n=933)</td>
<td>RR = 1 (0.56, 1.79) (K=2; n=1064)</td>
<td>RR (random effects) = 0.55 (0.17, 1.75) (K=2; n=230)</td>
<td>RR = 0.96 (0.55, 1.67) (K=1; n=103)</td>
<td>RR (random effects) = 0.79 (0.62, 1.02) (K=12; n=230)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attrition rate treatment group versus standard care</th>
</tr>
</thead>
<tbody>
<tr>
<td>24% versus 25% (K=6; n=1258)</td>
</tr>
</tbody>
</table>

Notes: RR= relative risk (95% confidence interval); K = number of trials contributing to the summary statistic; n = number of participants; ‘RMnppr nn.nn’ refers to the relevant forest plot in Appendix 20.
1) Summary of evidence profiles (portrait) in the full guideline

Table 8 GRADE profile – outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. of studies</th>
<th>Design</th>
<th>Intervention</th>
<th>Control</th>
<th>Relative risk</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of antibiotics</td>
<td>1</td>
<td>RCT</td>
<td>Delayed 27/62 (43%)</td>
<td>Immediate 54/81 (89%)</td>
<td>0.49 (0.36, 0.66)</td>
<td>Moderate</td>
</tr>
<tr>
<td>Temperature (°C) (day 3)</td>
<td>1</td>
<td>RCT</td>
<td>Mean score [n, C]; delayed = 36.7, immediate = 36.9 (analysis of comparison not provided)</td>
<td>Moderate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptom scores* (day 3)</td>
<td>1</td>
<td>RCT</td>
<td>Mean score: delayed = 5.4, immediate = 5.1 (analysis of comparison not provided)</td>
<td>Moderate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belief AB are effective</td>
<td>1</td>
<td>RCT</td>
<td>Delayed 51/87 (76%)</td>
<td>Immediate 47/82 (76%)</td>
<td>1.00 (0.82, 1.21)</td>
<td>Moderate</td>
</tr>
<tr>
<td>Patient satisfaction* (day 3)</td>
<td>1</td>
<td>RCT</td>
<td>Delayed 64/87 (96%)</td>
<td>Immediate 58/82 (94%)</td>
<td>1.02 (0.93, 1.10)</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

* Intervention = delayed antibiotics
* Control = immediate antibiotics

2) Full GRADE profiles (landscape, in appendices)
Some examples

2) Full GRADE profiles (landscape, in appendices)

### Evidence Summary: antidepressants review

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Evidence details</th>
<th>GRADE assessment</th>
<th>Study Quality</th>
<th>Direction</th>
<th>Imprecision</th>
<th>Inconsistency</th>
<th>Reporting</th>
<th>GRADE</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>09/09/2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2</strong> Full GRADE profiles (landscape, in appendices)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Table 1: Comparison of tricyclics versus placebo

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Evidence details</th>
<th>GRADE assessment</th>
<th>Study Quality</th>
<th>Direction</th>
<th>Imprecision</th>
<th>Inconsistency</th>
<th>Reporting</th>
<th>GRADE</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>09/09/2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2</strong> Full GRADE profiles (landscape, in appendices)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Some examples

3) Full GRADE profiles (split-portrait, in appendices)

![Table and Graph]

Some examples

4) Full GRADE profiles (split-portrait, in Full guideline)

![Table and Graph]
Some examples

5) Full GRADE profiles (modified-portrait, in Full guideline)

Table 19 GRADE profiles – carbamazepine as monotherapy for neuropathic pain

<table>
<thead>
<tr>
<th>No. of studies</th>
<th>Design</th>
<th>Treatment</th>
<th>Placebo</th>
<th>Relative risk (95% CI)</th>
<th>[ARR]</th>
<th>[NNTB, 95% CI]</th>
<th>Limitations</th>
<th>Other considerations</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 (PSP) (MixNP)</td>
<td>RCT</td>
<td>Carbamazepine</td>
<td>Placebo</td>
<td>1.31 (0.80, 2.15)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (PSP)</td>
<td>RCT</td>
<td>Any adverse effect</td>
<td>Placebo</td>
<td>1.86 (1.04, 3.30)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Very low</td>
</tr>
</tbody>
</table>

- N = No serious; S = Serious; VS = Very serious
- N/A = not applicable; PSP = post stroke pain; MixNP = mixed neuropathic pain
- Categorical scales for patient-reported global improvement/impression of change were dichotomised for analysis. For example, 'at least moderate improvement' on a 6-item scale, 'at least good improvement' on a 5-item scale or 'much or very much improved' on the patient’s global impression of change (PGIC) scale were the cut-offs for dichotomisation.
- GDG consensus: Total number of events (positive trends) less than 300, downgrade quality by 1 level.
- GDG consensus: If there is only 1 study with total number of adverse effects less than 100, the GDG decided that the quality should be graded as ‘very low’.

1 Leijon and Boivie (1989)
2 Nicol (1969)

Next step

- To expand the study to clinical guidelines developed by other developers

One example: BMJ Clinical Evidence: Diabetes: foot ulcers and amputations (Dereck Hunt, 2007)
Preliminary discussion

- Substantial variations (even just within NICE guidelines)

**Some key comments and questions:**

Firstly,
- Why and how the standard GRADE profiles were derived? What were the rationales behind it?

From the preliminary results:
- Why different formats and styles were adopted by developers?
  - Portrait vs landscape – better flow in a printed document?
- Do Full GRADE profiles need to be in the Full guideline? Or the Summary of findings & overall quality sufficient?
  - Clinical guidelines are already TOO LONG to read?
  - Full GRADE profiles vs Summary of evidence profiles – Do the readers really want to know the details of quality assessment (details on GRADE criteria)

---

Preliminary discussion

- As long as appropriate details were captured, does the format of the profiles really matter?
- What are the readers’ point of views?

**Next step:**

To find answers for the questions raised!

Collect information on:
- Evidence profiles used by other guideline developers
- The rationales behind the original standard GRADE profiles (GRADE Working Group)
- Reasons and rationales behind different adapted evidence profiles from different guideline developers
- Reader’s point of views
Still ongoing project........

More results in the near future

Thank you

Toni Tan: toni.tan@nice.org.uk
Faisal Siddiqui: faisal.siddiqui@nice.org.uk
Judith Thornton: judith.thornton@nice.org.uk