Using the GRADE evidence to decision framework (EtD) to make decisions as guideline developers

NPSO Workshop
Dublin Castle
October 17th 2018

Dr. Barbara Clyne
Dr. Miriam Bell
Ms. Shelley O’Neill
Dr. Sinéad O’Neill
Workshop Overview

• Intro to GRADE and an overview of how GRADE is used to assess the certainty of the evidence from systematic reviews

• Group work – brainstorming exercise

• Using the GRADE evidence to decision framework (EtD) to aid guideline developers to make recommendations

• Group work – Participants will become GDG members and vote on recommendations!

• GRADE in the real world – a guideline developer’s hands-on experience

• Close/Questions
What is a guideline?

“National Clinical Guidelines are systematically developed statements, based on a thorough evaluation of the evidence, to assist practitioner and service users’ decisions about appropriate healthcare for specific clinical circumstances across the entire clinical system”
(NCEC, 2018)
Guideline Development Process

1. Development of the scope of the document
2. Setting up Guideline Development Group and External Review Group
3. Management of Conflicts of Interest
4. Formulation of the questions (PICO) and choice of the patient important outcomes
5. Evidence retrieval, assessment and synthesis (systematic review(s))
   - GRADE - evidence profile
6. Formulation of the recommendations (GRADE)
   - Including explicit consideration of:
     - Benefits and harms
     - Values and preferences
     - Resource use
7. Dissemination, implementation (adaptation)
8. Evaluation of impact
9. Plan for updating

Initial guideline approval
- After completion of 1 and 2
- With draft of 4
- With plan for 3, 5-9

Final guideline approval
- After completion of 6
- With plan for 7-9
Today’s workshop – working with evidence using GRADE

For key recommendations:
• Search for and retrieve all available evidence
• Identify relevant systematic reviews
• Formally assess certainty of the evidence

GRADE provides a systematic and transparent approach to (1) assessing the certainty of the evidence and (2) making recommendations using standardised tables:
  – Summary of evidence (SOF tables)
  – The factors that affect the final recommendation (Evidence to Decision (EtD) table)
So what is GRADE?

- **GRADE**: Grades of recommendation, assessment, development and evaluation
- A common international grading system developed by experts from the WHO, NICE, Oxford CEBM, CDC, etc.
- Over 100 organisations using the system so far...
So what are we GRADING?

1. “Grading” the certainty of the evidence (from the SR)

2. “Grading” the recommendation (based on the evidence and other criteria)
1. GRADING the ‘evidence’
2. GRADING the ‘recommendations’

**1. PICO**
- Outcome: Critical
- Outcome: Critical
- Outcome: Important
- Outcome: Not important

**2. Systematic review**

**3. Guideline development**

**Formulate recommendations:**
- For or against (direction)
- Strong or conditional/weak (strength)

**By considering:**
- Quality of evidence
- Balance benefits/harms
- Values and preferences

**Revise if necessary by considering:**
- Resource use (cost)

**Randomization increases initial quality**
- High
- Moderate
- Low
- Very low

**Grade down**
- 1. Risk of bias
- 2. Inconsistency
- 3. Indirectness
- 4. Imprecision
- 5. Publication bias

**Grade up**
- 1. Large effect
- 2. Dose response
- 3. Confounders

**Grade**
- Overall quality of evidence across outcomes based on lowest quality of critical outcomes

- “We recommend using...”
- “We suggest using...”
- “We recommend against using...”
- “We suggest against using...”
1) “GRADING” the Certainty of the Evidence

In the context of making recommendations

The certainty of the evidence reflects our confidence that the estimates of an effect are adequate to support a particular recommendation.
1) “GRADING” the Certainty of Evidence – How?

- Using the GRADE approach we assess how confident we are that an estimate is close to the true effect?
- This is done for each of the main outcomes across ALL available studies

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Quality rating</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>☑️☑️☑️</td>
<td>High</td>
<td>Very confident that the true effect lies close to that of the estimate of the effect</td>
</tr>
<tr>
<td>☑️☑️☐️</td>
<td>Moderate</td>
<td>Moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different</td>
</tr>
<tr>
<td>☑️☐️☐️</td>
<td>Low</td>
<td>Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect</td>
</tr>
<tr>
<td>☑️☐️☐️</td>
<td>Very low</td>
<td>Very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of the effect</td>
</tr>
</tbody>
</table>
1) “GRADING” the Certainty of the Evidence across all Studies - Where to start?

Where to start randomised controlled trials (RCTs) and observational studies (High, moderate, low, very low)?

• RCTs start **HIGH**

• Observational studies start **LOW**

What can **lower** our confidence/certainty in the evidence?

• Risk of bias (studies conducted in a way that may lead to errors)
• Indirectness (studies conducted in a different population to the population of interest)
• Inconsistency (some studies with positive findings, some with negative findings)
• Imprecision (small sample size, low event rate)
• Publication bias (negative results and non-English studies less likely to be published)
1) “GRADING” the Certainty of the Evidence across all Studies - Where to start?

What can **increase** our confidence/certainty in the evidence?

- Large magnitude of effect (large consistent results)
- Plausible confounding would reduce the demonstrated effect or increase the effect if no effect was observed (other factors influencing results)
- Dose-response gradient (more of something makes its better/worse, e.g. vaccines and herd immunity, the more people vaccinated, the lower the risk of disease becomes)
Overall certainty of the evidence

- **Combined** rating of the certainty of the evidence across **all outcomes** considered critical for answering a health care question (e.g. mortality, admission rates)

- Can differ across outcomes. When determining the overall certainty of the evidence across outcomes:
  - Consider only ‘critical outcomes’
  - If the **certainty of the evidence is the same** for all, then this becomes the overall quality
  - If the **certainty of the evidence differs** across outcomes, the lowest quality of evidence for any of the critical outcomes determines the overall quality of evidence.
Meta-analyses of several critical and important outcomes (one PICO)

- **Hospitalization** (critical)
  - Relative Risk: 1
  - Quality: High
  - Imprecision

- **Mortality** (critical)
  - Relative Risk: 0.75
  - Quality: Moderate

- **Nausea** (important)
  - Relative Risk: 0.9
  - Quality: Low
  - Imprecision and risk of bias

- **SAE** (critical)
  - Relative Risk: 1.25
  - Quality: High

Overall Quality of Evidence:
1. Conduct SR based on a clearly defined PICO question.

2. Using the GRADE approach, assess the ‘certainty of the evidence’ for each critical outcome in your SR.

3. Present in GRADE summary of findings table.

4. Move to ‘GRADING of the recommendations’.

---

**Table 4.1 Template for a Summary of Findings table**

<table>
<thead>
<tr>
<th>Patients or population:</th>
<th>[Text]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Settings:</td>
<td>[Text]</td>
</tr>
<tr>
<td>Intervention:</td>
<td>[Text]</td>
</tr>
<tr>
<td>Comparison:</td>
<td>[Text]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Impact</th>
<th>Number of participants (Studies)</th>
<th>Quality of the evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Text]</td>
<td>[Text]</td>
<td>[Text]</td>
<td>[Text]</td>
</tr>
<tr>
<td>[Text]</td>
<td>[Text]</td>
<td>[Text]</td>
<td>[Text]</td>
</tr>
<tr>
<td>[Text]</td>
<td>[Text]</td>
<td>[Text]</td>
<td>[Text]</td>
</tr>
</tbody>
</table>

* GRADE: GRADE Working Group grades of evidence:
  
* High: We are confident that the true effect lies close to that of the estimate of the effect.
  
* Moderate: We think a true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
  
* Low: The true effect may be substantially different from the estimate of the effect.
  
* Very low: Any estimate of effect is very uncertain.

---

**Systematic review**

- Formulate question
- Select outcomes
- Rate importance
- Outcomes across studies
- Create evidence profile with GRADEpro
- Rate quality of evidence for each outcome
- Randomization increases initial quality

**Guideline development**

- For or against (direction)
- Strong or conditional/weak (strength)
- By considering:
  - Quality of evidence
  - Balance benefits/harms
  - Values and preferences

Revise if necessary by considering:
- Resource use (cost)
GROUP WORK
Brainstorming exercise
2. GRADING and ‘recommendations’

- How do we go from the evidence to recommendations?

- **Group work**: What factors do you think need to be considered when trying to develop recommendations for National Clinical Guideline No. 9?
1.3 Scope of National Clinical Guideline

The National Clinical Guideline applies to healthcare professionals involved in the management of cancer pain. This includes Palliative Care staff, Physicians, Surgeons, General Practitioners, Pharmacists and Nursing staff in hospital, hospice and community-based settings. The guideline recommendations indicate where specialist advice should be sought. The Guideline will also be of interest to patients with cancer pain and their carers. A patient information leaflet is also available at www.hse.ie/palliativecareprogramme and www.health.gov.ie/patient-safety/ncec

The National Clinical Guideline does not apply to cancer survivors, to patients who do not have a cancer diagnosis or to the management of other forms of acute or chronic non-malignant pain. The National Clinical Guideline does not apply to children.
How do we go from the evidence to recommendations?

1. Formulate question
2. Select outcomes
3. Rate importance
   - Outcome Critical
   - Outcome Critical
   - Outcome Important
   - Outcome Not important
4. Outcomes across studies
5. Create evidence profile with GRADEpro
6. Rate quality of evidence for each outcome
   - High
   - Moderate
   - Low
   - Very low
7. Summary of findings & estimate of effect for each outcome
8. Randomization increases initial quality
   - 1. Risk of bias
   - 2. Inconsistency
   - 3. Indirectness
   - 4. Imprecision
   - 5. Publication bias
9. Systematic review
10. Grade up
   - 1. Large effect
   - 2. Dose response
   - 3. Confounders

Guideline development

1. Formulate recommendations:
   - For or against (direction)
   - Strong or conditional/weak (strength)
   - By considering:
     - Quality of evidence
     - Balance benefits/harms
     - Values and preferences
   - Revise if necessary by considering:
     - Resource use (cost)
2. Panel
3. “We recommend using…”
4. “We suggest using…”
5. “We recommend against using…”
6. “We suggest against using…”
7. Grade overall quality of evidence across outcomes based on lowest quality of critical outcomes
Evidence to decision framework (EtD)

- Criteria to determine a recommendation
  - Priority of problem
  - Benefits and harms
  - Quality of evidence
  - Values
  - Resources
  - Equity
  - Acceptability
  - Feasibility
# Criteria to determine a recommendation

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Influence on direction and strength of a recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem</td>
<td>Determined by the importance &amp; frequency of the healthcare issue (e.g. burden, risk). If the problem is of great importance a strong recommendation is more likely</td>
</tr>
<tr>
<td>Benefits &amp; harms</td>
<td>Requires an evaluation of effects of the benefits &amp; harms - desirable &amp; undesirable consequences. The greater the benefit, a strong recommendation for is more likely. The greater the harm, a strong recommendation against is more likely.</td>
</tr>
<tr>
<td>Quality of the evidence</td>
<td>The higher the quality of evidence, a strong recommendation is more likely</td>
</tr>
<tr>
<td>Values &amp; preferences</td>
<td>How important the health outcomes are to those affected, how variable this importance is, and if there is uncertainty. Greater the variability, or uncertainty, weak recommendations are more likely.</td>
</tr>
<tr>
<td>Costs (resource use)</td>
<td>The higher the costs of an intervention (i.e. more resources consumed) – strong recommendation less likely</td>
</tr>
<tr>
<td>Equity</td>
<td>The greater the likelihood to reduce health/access inequities, a strong recommendation is more likely</td>
</tr>
<tr>
<td>Acceptability</td>
<td>The greater the acceptability to all/most stakeholders, a strong recommendation is more likely</td>
</tr>
<tr>
<td>Feasibility</td>
<td>The greater the feasibility to all/most stakeholders, a strong recommendation is more likely</td>
</tr>
</tbody>
</table>
Evidence to decision framework (EtD)

**Guideline Question:** Should multi-vessel PPCI versus culprit only PPCI be used in patients with STEMI and multi-vessel coronary artery disease undergoing PPCI? – modified from Saudi Arabian guidelines

**Problem:** STEMI
**Option:** multi-vessel PPCI
**Comparison:** culprit-only PPCI
**Setting:** in hospital
**Perspective:** KSA MoH

**Background and Objective:** Many patients with STEMI have multi-vessel disease with significant stenosis in arteries other than the culprit vessel. The strategy of treating all significant lesions at the time of PPCI has the advantage of complete revascularization thus potentially decreasing future cardiac events. The management of the additional lesions at the time of PPCI is controversial. This question addresses whether multi-vessel PPCI versus culprit only PPCI should be performed in patients with STEMI and multi-vessel coronary artery disease.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Judgements</th>
<th>Research evidence</th>
<th>Additional considerations</th>
</tr>
</thead>
</table>
| **Problem** | Is there a problem priority? | ○ No<br>○ Probably no<br>○ Uncertain<br>○ Probably yes<br>● Yes<br>○ Varies | The Saudi population has been reported to show a high burden of cardiovascular risk factors and early manifestation in a younger cohort in comparison, for example, to European populations – e.g. mean age 58 years (SD +/- 12.9) in SPACE registry study (Khan 2014, AlHabib et al 2011).<br>The SPACE registry study of 5055 acute coronary syndrome patients admitted to 17 hospitals in Saudi Arabia between December 2005 and December 2007 reported that 41.5% had STEMI (AlHabib et al 2011). Additionally, the GULF RACE-2 registry study conducted in 65 hospitals from 6 Arabian Gulf countries (including Saudi Arabia) between October 2008 and June 2009 reported that of 7930 patient enrolled 45.6% had STEMI, and 1-year mortality in STEMI patients was 11.5% (AlHabib et al 2012). | }
Evidence to decision framework (EtD)

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Judgements</th>
<th>Research evidence</th>
<th>Additional considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What is the overall certainty of this evidence?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No included studies</td>
<td></td>
<td><strong>Low</strong></td>
<td>No evidence specific to KSA identified in literature search for patients’ values and preferences.</td>
</tr>
<tr>
<td>Very low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td>Panel members, including patient representative noted that there is no important variability or uncertainty in patients’ values and preferences in the KSA setting and that patients would place a high value on reduction in mortality.</td>
</tr>
<tr>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **Is there important uncertainty about how much people value the main outcomes?** |            |                   |                           |
| No important uncertainty of variability       |            |                   |                           |
| Possibly important uncertainty or variability |            |                   |                           |
| Probably no important uncertainty of variability |            |                   |                           |
| No known undesirable                           |            |                   |                           |

| **Are the desirable anticipated effects large?** | |                   |                           |
| No                                            |            |                   |                           |
| Probably no                                   |            |                   |                           |
| Uncertain                                     |            |                   |                           |
| Probably yes                                  |            |                   |                           |
| Yes                                           |            |                   |                           |
| Varies                                        |            |                   |                           |

| **Are the undesirable anticipated effects small?** | |                   |                           |
| No                                            |            |                   |                           |
| Probably no                                   |            |                   |                           |
| Uncertain                                     |            |                   |                           |
| Probably yes                                  |            |                   |                           |
| Yes                                           |            |                   |                           |
| Varies                                        |            |                   |                           |

| **Are the desirable effects large relative to undesirable effects?** | |                   |                           |
| No                                            |            |                   |                           |
| Probably no                                   |            |                   |                           |
| Uncertain                                     |            |                   |                           |
| Probably yes                                  |            |                   |                           |
| Yes                                           |            |                   |                           |
| Varies                                        |            |                   |                           |

**Summary of findings:** multi-vessel PPCI compared to culprit only PPCI in patients with STEMI and multi-vessel coronary artery disease undergoing PCI.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Without multi-vessel PPCI</th>
<th>With multi-vessel PPCI</th>
<th>Difference (95% CI)</th>
<th>Relative effect (RR) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality - long term</td>
<td>46 per 1000 (27 to 77)</td>
<td>27 fewer per 1000 (from 4 more to 46 fewer)</td>
<td>RR 0.63 (0.37 to 1.05)</td>
<td></td>
</tr>
<tr>
<td>Reinfarction</td>
<td>25 per 1000 (13 to 48)</td>
<td>42 fewer per 1000 (from 19 fewer to 54 fewer)</td>
<td>RR 0.37 (0.19 to 0.71)</td>
<td></td>
</tr>
<tr>
<td>Revascularization</td>
<td>192 per 1000 (50 to 102)</td>
<td>121 fewer per 1000 (from 90 fewer to 142 fewer)</td>
<td>RR 0.37 (0.28 to 0.53)</td>
<td></td>
</tr>
</tbody>
</table>

| Contrast Induced Nephropathy | 17 per 1000 (3 to 33) | 5 fewer per 1000 (from 15 fewer to 15 more) | RR 0.55 (0.16 to 1.80) |

| Additional considerations | |                   |                           |
| No data available on harms. Panel noted that there are 3 potential harms of multi-vessel PPCI: (i) using more contrast, (ii) every PCI carries risk (iii) time of procedure is extended. | |                   |                           |
## Evidence to decision framework (EtD)

<table>
<thead>
<tr>
<th>Are the resources required small?</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
</tr>
<tr>
<td>Probably no</td>
</tr>
<tr>
<td>Uncertain</td>
</tr>
<tr>
<td>Probably yes</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Varies</td>
</tr>
</tbody>
</table>

No evidence identified specific to KSA – Panel members considered costs and resource use when comparing multi vessel vs. culprit only PCI.

<table>
<thead>
<tr>
<th>Is the incremental cost small relative to the net benefits?</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
</tr>
<tr>
<td>Probably no</td>
</tr>
<tr>
<td>Uncertain</td>
</tr>
<tr>
<td>Probably yes</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Varies</td>
</tr>
</tbody>
</table>

No evidence identified specific to KSA – Panel members considered whether the incremental costs (when comparing multi vessel vs. culprit only PCI) are small relative to benefits.

One cost-consequence analysis was identified in the NICE systematic review, based on the HELP-AMI RCT (Di Mario 2004) conducted in Italy. The analysis found that based on the small sample size of 69 patients culprit-only PCI was more costly over 12 months than immediate multi-vessel PCI; incremental cost of £1412 more per patient (p=0.325). The trial found no significant differences in clinical outcomes between culprit-only and multi-vessel PCI in patients who had multi-vessel disease.

Given the probably large anticipated desirable effects but additional cost of the procedure, and no data specific to the KSA setting, the panel judged the cost-effectiveness as ‘probably yes’ to ‘yes’.
## Evidence to decision framework (EtD)

<table>
<thead>
<tr>
<th>Equity</th>
<th>What would be the impact on health inequities?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image" alt="Table" /></td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Options" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Acceptability</th>
<th>Is the option acceptable to key stakeholders?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image" alt="Table" /></td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Options" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feasibility</th>
<th>Is the option feasible to implement?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image" alt="Table" /></td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Options" /></td>
</tr>
</tbody>
</table>
Strength of recommendations

• Reflects extent to which a guideline development group is confident that desirable effects of an intervention outweigh undesirable effects, or vice versa, across the range of patients for whom the recommendation is intended

• GRADE has two categories:
  – Strong
    • “we recommend” or “clinicians should”
  – Weak/conditional
    • “we suggest” or “we conditionally recommend” or “clinicians may”
<table>
<thead>
<tr>
<th></th>
<th>Strong Recommendation</th>
<th>Weak Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>For patients</strong></td>
<td>Most individuals in this situation would want the recommended course of action and only a small proportion would not.</td>
<td>The majority of individuals in this situation would want the suggested course of action, but many would not.</td>
</tr>
<tr>
<td><strong>For clinicians</strong></td>
<td>Most individuals should receive the recommended course of action. Adherence to this recommendation could be used as a quality criterion or performance indicator. Formal decision aids are not likely to be needed to help individuals make decisions consistent with their values and preferences.</td>
<td>Recognize that different choices will be appropriate for different patients, and that you must help each patient arrive at a management decision consistent with her or his values and preferences. Decision aids may well be useful helping individuals making decisions consistent with their values and preferences. Clinicians should expect to spend more time with patients when working towards a decision.</td>
</tr>
<tr>
<td><strong>For policy makers</strong></td>
<td>The recommendation can be adapted as policy in most situations including for the use as performance indicators.</td>
<td>Policy making will require substantial debates and involvement of many stakeholders. Policies are also more likely to vary between regions. Performance indicators would have to focus on the fact that adequate deliberation about the management options has taken place.</td>
</tr>
</tbody>
</table>
Participants exercise – using Mentimeter

- Workshop participants – you will now act as guideline developers and vote for the following:

1. Whether to make a recommendation that education will form part of the National Early Warning System (NEWS) National Clinical Guideline No. 1 [Yes, No]

2. And if so, whether or not to make a ‘strong’ or ‘weak’ recommendation by voting on each of the EtD criteria (including the certainty of the evidence)

3. Using your phones, please go to the following link: www.menti.com and enter the code: 58 64 92
Participants exercise – using Mentimeter – Background info

NEWS is an early warning system used in adult patients in acute settings (e.g. hospitals) whereby vital signs are measured frequently (including blood pressure, heart rate, respiratory rate, etc.) with the view to detecting deterioration. These measurements are plotted on a standardised observation chart with a scoring system (0-3) which is colour-coded.

Any derangements in vital signs (a score of 3 in a single vital sign, or a total score of 7+) should trigger healthcare workers to call for help (escalate care).
Participants exercise – using Mentimeter – Background info

- The **NEWS** guideline is currently being updated. A systematic review (SR) on the effectiveness of various early warning score-based educational interventions has been completed.

- For each of the SRs primary outcomes, GRADE was used to assess the certainty of the evidence, which was found to be:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>☠ ☠ ☠ ☠</td>
<td>Very low</td>
</tr>
</tbody>
</table>

- The summary of findings table is presented on the next slide.
The effectiveness of educational interventions in detecting physiological deterioration in adult (non-pregnant) patients in acute health care settings

**Patient or population:** nurses, doctors, other health care professionals  
**Setting:** Varied (hospital, university simulation lab)  
**Intervention:** educational interventions (including virtual or mannequin-based simulation, validated education programmes such as COMPASS®, hospital specific educational interventions) delivery either face-to-face or blended (online component).  
**Comparison:** another educational intervention, or no educational intervention

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Impact</th>
<th>No of participants (studies)</th>
<th>Certainty of the evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>All 5 studies demonstrated an increase in knowledge post-educational intervention.</td>
<td>270 (5 studies including 3 RCTs, 2 before-after studies)</td>
<td>☒☐☐☐ VERY LOW ab,c</td>
</tr>
<tr>
<td>Performance/confidence</td>
<td>All 7 studies demonstrated an increase in clinical performance or self-confidence post-educational intervention.</td>
<td>371 (7 studies including 4 RCTs and 3 before-after studies)</td>
<td>☒☐☐☐ VERY LOW ab,c,d</td>
</tr>
</tbody>
</table>

The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval.

GRADE Working Group grades of evidence
- **High certainty:** We are very confident that the true effect lies close to that of the estimate of the effect
- **Moderate certainty:** We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
- **Low certainty:** Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
- **Very low certainty:** We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

**Explanations**
- a. High or unclear risk of risk: no blinding of participants and personnel in three RCTs,  
- b. High risk of confounding in observational studies,  
- c. Single centre study -may not be generalisable to other settings,  
- d. Small sample size
VOTING TIME

Would you make a recommendation that:

“There should be an educational programme around the National Early Warning Score (NEWS)”?

YES or NO

www.menti.com and enter the code: 58 64 92
VOTING TIME

If YES, going through each of the EtD criteria vote:

1) Is the problem a priority?
2) How substantial are the desirable anticipated effects (benefits)?
3) How substantial are the undesirable anticipated effects (harms)?
4) What is the certainty of the evidence (very low)
5) Is there important uncertainty about or variability in how much people value the main outcomes (values)
6) Does the balance between desirable and undesirable effects favour the intervention or the comparison?
7) How large are the resources required (costs)?
8) What would be the impact on health equity?
9) Is the intervention acceptable to key stakeholders?
10) Is the intervention feasible to implement?

Strength of the recommendation: Strong or weak?

www.menti.com and enter the code: 58 64 92
NCG No. 1 National Early Warning System (NEWS):
Guideline Development Group and the GRADE Process
What we did...

- Reviewed existing NEWS recommendations x 60 (in-depth familiarity was of benefit when reviewing the evidence)
- Streamlined existing recommendations where possible through amalgamation/omission
- Reviewed relevant sections of systematic review and SoF tables from HRB-CICER
- Drafted GRADE EtD framework (Education)
- Draft EtD (Education) reviewed by GDG, discussed, debated and agreed by consensus
- New recommendations for education drafted
- Justification for strength and direction of recommendation debated, agreed and documented (GRADE)
- GDG reviewed, discussed and agreed draft new recommendations
Grade assessment: Is the problem a priority? **YES**

- **GDG Rationale:** Patients and service users expect that healthcare professionals (HCPs) are educated to use EWSs.

- Regulatory bodies expect that EWS education is provided and that HCPs undertake such training.

- Audit, reports, critical incident analyses and focus group findings strongly suggest education should be provided for HCPs on EWSs.

- While certainty of evidence is very low, all studies showed improvement in outcomes following educational interventions (performance, confidence and knowledge; documentation of observations; communication and collaboration between nurses and physicians).

- This finding is corroborated in areas of research relevant to EWSs (Lockey et al. (2018) Impact of adult advanced cardiac life support course participation on patient outcomes – a systematic review and meta-analysis. *Resuscitation*, Vol. 129, pp 48-54) where quality of evidence is very low but studies demonstrate improved outcomes from participation in education.
Desirable effects: how substantial are the desirable anticipated effects? **Moderate**

- **GDG Rationale**: Anticipated desirable effects probably large but cannot say as there is low evidence quantifying the effect of educational interventions.

- GDG agreed to revisit this question if more evidence comes to light or GDG members have additional comments following GDG meeting.

NEWS GDG October 2018
Patient Values and Preferences

GDG: **Probably no important uncertainty or variability**

• Is there important uncertainty about or variability in how much people value the main outcomes?

• **GDG rationale:** It can be assumed that the majority of people would value improved healthcare professional knowledge and performance of EWSs
So what did the GDG decide?

• Is the problem a priority? **YES**
• How substantial are the desirable anticipated effects (benefits)? **Moderate**
• How substantial are the undesirable anticipated effects (harms)? **Small**
• What is the certainty of the evidence **Very low**
• Is there important uncertainty about or variability in how much people value the main outcomes (values) **Probably not**
• Does the balance between desirable and undesirable effects favour the intervention or the comparison? **Favours intervention**
• How large are the resources required (costs)? **Large costs**
• What would be the impact on health equity? **Probably ✓**
• Is the intervention acceptable to key stakeholders? **Yes**
• Is the intervention feasible to implement? **Probably yes**

**Strength of the recommendation:** **Strong**
Most useful features of GRADE Process?

- Structure (SoFs, EtDs)
- Prompts (EtD sections and questions)
- GRADE ‘Conclusion section’ (justification, subgroup considerations, implementation considerations, monitoring and evaluation and research priorities)
- Very useful in assisting GDG in making rationale for recommendations based on factors other than certainty of evidence
Biggest challenge?

- Becoming familiar with GRADE software
- Becoming familiar with the GRADE process which was new to GDG members
- Absence of high quality evidence for most of our clinical questions
- Making the case for a strong recommendation in the absence of evidence certainty
Acknowledgements

• Funding from the Health Research Board (HRB) Ireland, for HRB-CICER (Grant No: HRB-CICER-2016-1871)

• NPSO Committee for the opportunity to deliver this workshop
Further reading and resources


• Cochrane Training: http://training.cochrane.org/path/grade-approach-evaluating-quality-evidence-pathway

• GRADE BMJ series: 2008;336;924-926


• Journal of Clinical Epidemiology GRADE series: http://www.jclinepi.com/content/jce-GRADE-Series

• GRADE HANDBOOK (in particular Chapters 4,5,6) http://gdt.guideddevelopment.org/app/handbook/handbook.html
GRADE Online Learning Modules

These online learning modules are designed to help guideline developers and authors of systematic reviews learn how to use the GRADE approach to grade the evidence in systematic reviews, to create Summary of Findings Tables and GRADE Evidence Profiles, and move from evidence to making recommendations.

There are two sets of modules:
1. For authors of Cochrane systematic reviews and other systematic reviewers
2. For World Health Organization (WHO) guideline developers and other guideline developers

Each module covers a specific topic related to GRADE and can be viewed in sequence as presented here or, depending on your learning needs, viewed in any order. Most modules are approximately 20 minutes or less and can be started and stopped at any time, and restarted at the same point at which you stopped.

Viewing is self-directed and anonymous. However, we welcome feedback. Please contact us with any comments.

- Introduction to GRADE and Summary of Findings Tables
- Choosing a comparison and outcomes for the Summary of Findings Table
- Assessing Risk of Bias
- WHO guidelines and GRADE: An overview summary (60 minutes)
- WHO guidelines and GRADE: Introduction
- Formulating questions and choosing outcomes

https://cebgrade.mcmaster.ca/index.html
GRADE Working Group

Welcome to the GRADE working group

From evidence to recommendations – transparent and sensible

http://www.gradeworkinggroup.org/
Guideline Development Tool

https://GRADEpro_GDT_software/
Questions? Comments?