The Grading of Recommendations Assessment, Development and Evaluation (short GRADE)

Lillebeth Larun
Lillebeth.larun@fhi.no

Norwegian Institute of Public Health
Division for Health Services
Department for Evidence Summaries
Plan

• Introduction to GRADE
• GRADE domains
• Summary of findings table
• Certainty of evidence
Definition of certainty of evidence

• “the ratings of the certainty of evidence reflect the extent of our confidence that the estimates of the effect are correct”

• “not rating certainty in point estimates, but rather certainty that the true effect lies in a particular range”

M. Hultcrantz et al. / Journal of Clinical Epidemiology 87 (2017) 4-13
GRADE

- assessing the certainty of the evidence
- structured and transparent
- criteria for downgrading and upgrading the certainty in evidence
<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative comparative risks* (95% CI)</th>
<th>Relative effect (95% CI)</th>
<th>No of Participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Assumed (baseline) Corresponding (intervention) risk</td>
<td>RR 1.20 (1.04 to 1.37)</td>
<td>3004 (7 studies)</td>
<td></td>
<td>low³</td>
</tr>
<tr>
<td>Final time point (3, 6, or 12 months post-intervention)</td>
<td>Vaccination status</td>
<td>55 per 100¹</td>
<td>66 per 100 (57 to 75)</td>
<td>The results for this outcome were variable, so the true result may be substantially higher or lower than this estimate</td>
<td></td>
</tr>
</tbody>
</table>

* Relative risk

¹ Based on reviews of 1 and 2 studies

² Low confidence

³ Low to very low confidence

GRADE – SoF
Evidence synthesis

Grade overall quality of evidence across outcomes based on lowest quality of critical outcomes

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

Formulate Recommendations (↓↑ | ⊕)
- "The panel recommends that ....should..." (↑↑ | ⊕)
- "The panel suggests that ....should..." (↑? | ⊕)
- "The panel suggests to not ..." (↓? | ⊕)
- "The panel recommends to not..." (↓↓ | ⊕)

Formulate Recommendations

Grade recommendations
- For or against (direction) ↓↑
- Strong or conditional/weak (strength)

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

Randomization increases initial quality

1. Risk of bias
2. Inconsistency
3. Indirectness
4. Imprecision
5. Publication bias

Grade up
Grade down

Very low
Low
Moderate
High

Summary of findings & estimate of effect for each outcome

Rate quality of evidence for each outcome

Outcome: Critical
Outcome: Important
Outcome: Not important

Outcome: Critical
Outcome: Critical

Outcome: P I C O
Start of GRADEing

• formulate question
• decide critical outcomes
• decide importance of outcome
• identify systematic reviews/studies
Framing the question

**P**: parents of preschool-aged children or expectant parents

**S**: clinics, antenatal classes, or the mother’s home

**I**: face-to-face information or educational interventions

**C**: control (no education, other education, or control not described)

**O**: outcome

---

**Objectives**

To assess the effects of face-to-face interventions for informing or educating parents about early childhood vaccination on vaccination status and parental knowledge, attitudes and intention to vaccinate.
Critical outcomes and importance
Kaufman 2018

Worksheet
Exercise 1
Selecting and rating the importance of outcomes

1. Children: vaccination status of child (i.e. vaccination status up-to-date, or receipt of one or more vaccines, as defined by study authors); outcome domain: vaccination status and behaviours
2. Parents: knowledge or understanding of vaccination; outcome domain: knowledge or understanding
3. Parents: attitudes or beliefs about vaccination; outcome domain: attitudes or beliefs
4. Parents: intention to vaccinate child; outcome domain: attitudes or beliefs
5. All categories: adverse effects; outcome domain: any

https://gdt.gradepro.org/app/handbook/handbook.html#h.1i2bwkm8zpjo
https://gradeapro.org/product/#features

Everything you need from GRADEing your systematic review, to developing your guideline and beyond

An all-in-one web solution for summarizing and presenting information for health care decision making
# GRADE categories for the certainty of a body of evidence

<table>
<thead>
<tr>
<th>Quality level</th>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>⨁⨁⨁⨁</td>
<td>We are very confident that the true effect lies close to that of the estimate of the effect.</td>
</tr>
<tr>
<td>Moderate</td>
<td>⨁⨁◯◯</td>
<td>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.</td>
</tr>
<tr>
<td>Low</td>
<td>⨁◯◯◯</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>Very low</td>
<td>⨁◯◯◯</td>
<td>We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.</td>
</tr>
</tbody>
</table>
Determinants of certainty
5 factors that can lower certainty

1. Design limitations (risk of bias)
2. Inconsistency of results (heterogeneity)
3. Indirectness (applicability)
4. Imprecision (number of events and confidence intervals)
5. Publication bias
Risk of bias assessment

Review authors’ judgement:
- Was the allocation sequence adequately generated?
- Was allocation adequately concealed?
- Was knowledge of the allocated intervention adequately prevented during the study?
  - Participants
  - Trial personnel
- Were incomplete outcome data adequately addressed?
- Are reports of the study free of suggestion of selective outcome reporting?
- Was the study apparently free of other problems that could put it at a high risk of bias?

Domains from the Cochrane Handbook: http://handbook.cochrane.org/
Risk of bias summary for a body of evidence
«All studies had limitations in design. We downgraded evidence where the contributing studies were at high or unclear risk of bias for sequence generation (Jackson 2011), or allocation concealment (Hu 2017; Usman 2009; Usman 2011; Wood 1998)"
Inconsistency - Heterogeneity

No overlap, same direction of effect

Overlap, different directions of effect

No overlap, different directions of effect
Inconsistency

Exercise 3

https://gdt.gradepro.org/app/handbook/handbook.html#h.g2dqzi9je57e
“We also downgraded the certainty of the evidence for vaccination status, due to inconsistency, which was clear from the high level of statistical heterogeneity.”
Indirectness

- Differences in population (applicability)
- Differences in interventions (applicability)
- Differences in outcomes measures (surrogate outcomes)
- Indirect Comparisons

https://gdt.gradepro.org/app/handbook/handbook.html#h.w6r7mtvq3mjz
“We downgraded the certainty of the evidence for attitudes due to indirectness, because the specific measures for this outcome (perceived diseases severity and vaccine benefits) were only part of what could be measured, and what was relevant to parents and other decision makers.”
Imprecision

Ischemic stroke point estimate and confidence interval

Threshold if side effects, toxicity and cost minimal, NNT = 200. Entire confidence interval to left of threshold, do not rate down for imprecision

Threshold if side effects and toxicity appreciable, NNT = 100. Confidence interval crosses threshold, rate down for imprecision

NNT: Number-Needed-to-Treat

Fig. 1. Rating down for imprecision in guidelines: thresholds are key.
Imprecision

- Exercise 4

https://gdt.gradepro.org/app/handbook/handbook.html#h.ygojbnr1bi5y
Kaufman 2018

«The certainty of the evidence for intention to vaccinate and adverse effects was downgraded for imprecision, due to the wide confidence intervals for the included studies.»
Publication bias is a systematic under-estimation or an over-estimation of the underlying beneficial or harmful effect due to the selective publication of studies. Confidence in the combined estimates of effects from a systematic review can be reduced when publication bias is suspected, even when the included studies themselves have a low risk of bias.
"We did not downgrade any outcomes for issues of publication bias."
The Summary of Findings tables

• Is a summary of the key findings from the systematic review for users

• Presents
  – the quality of the evidence
  – the magnitude of the effect
  – reasons behind decisions
### Summary of Findings for the Main Comparison

**Face-to-face interventions directed to parents for informing or educating parents about early childhood vaccination, as compared with control**

**Patient or population:** parents of preschool-aged children or expectant parents  
**Settings:** clinics, antenatal classes, or the mother’s home  
**Intervention:** face-to-face information or educational interventions  
**Comparison:** control (no education, other education, or control not described)

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative comparative risks (95% CI)</th>
<th>Relative effect (95% CI)</th>
<th>No of Participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vaccination status</td>
<td></td>
<td>RR 1.20 (1.04 to 1.37)</td>
<td>3004 (7 studies)</td>
<td>☋�� ○ low²</td>
<td></td>
</tr>
<tr>
<td>Final time point (3, 6, or 12 months post-intervention)</td>
<td>55 per 100¹</td>
<td>66 per 100 (57 to 75)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ Assumed (baseline) Corresponding (Intervention) risk  
² GRADE criteria for outcome: low quality due to risk of bias, inconsistency, indirectness, imprecision, and publication bias.
Table: GRADE's approach to rating quality of evidence (aka confidence in effect estimates)

For each outcome based on a systematic review and across outcomes (lowest quality across the outcomes critical for decision making)

<table>
<thead>
<tr>
<th>1. Establish initial level of confidence</th>
<th>2. Consider lowering or raising level of confidence</th>
<th>3. Final level of confidence rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study design</strong></td>
<td><strong>Initial confidence in an estimate of effect</strong></td>
<td><strong>Reasons for considering lowering or raising confidence</strong></td>
</tr>
<tr>
<td>Randomized trials →</td>
<td>High confidence</td>
<td>↓ Lower if</td>
</tr>
<tr>
<td>Observational studies →</td>
<td>Low confidence</td>
<td>↑ Higher if*</td>
</tr>
</tbody>
</table>

*upgrading criteria are usually applicable to observational studies only.

http://www.gradeworkinggroup.org/
Learning material - examples

- Cochrane interactive learning
  http://training.cochrane.org/interactivelearning
- GRADE https://cebgrade.mcmaster.ca/
- Equatornetwork http://www.equator-network.org/
- Testing treatments
  http://www.testingtreatments.org/
Peer review

• Prepare a short introduction to what the review is about
• Positive feedback
• Points to consider improving
• Preparation time apr. 45 min
• Each feedback 14 min, max 7 min by opponent
• Group starts at 12.00
• Done by 13.30
<table>
<thead>
<tr>
<th>Name</th>
<th>Tentative research question</th>
<th>Tutor</th>
<th>Opponent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nazar</td>
<td>Breath analysis as a test for cancer screening: A systematic review and meta-analysis</td>
<td>Lillebeth</td>
<td>Israel</td>
</tr>
<tr>
<td>Israel</td>
<td>What is the prevalence of hearing loss among noise-exposed workers in Africa?</td>
<td>Lillebeth</td>
<td>Olive</td>
</tr>
<tr>
<td>Paul</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olive</td>
<td>Severe illness among infants, a systematic review</td>
<td>Lillebeth</td>
<td>Mari</td>
</tr>
<tr>
<td>Mari</td>
<td>To what degree do patients adhere to anticoagulation treatment in secondary prevention after cardioembolic stroke? Which factors may affect the level of adherence?</td>
<td>Lillebeth</td>
<td>Nazar</td>
</tr>
<tr>
<td>Name</td>
<td>Tentative research question</td>
<td>Tutor</td>
<td>Opponent</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------</td>
<td>-----------</td>
</tr>
<tr>
<td>Melf</td>
<td>What do we know about the cost effectiveness of SMS reminders to increase adherence to preventive therapies in sub-Saharan Africa?</td>
<td>Eva</td>
<td>Tekle</td>
</tr>
<tr>
<td>Tekle</td>
<td>Antimicrobial peptides are emerging topics of interest to fight against antimicrobial drug resistance. The bacterial toxin anti-toxin (TA) system, is less studied yet its promising results attracted attention of experts in the area. “So, what is the TA system in bacteria: its biology, mechanism of action, possibilities and implication in public health intervention” will be my topic during the systematic review course</td>
<td>Eva</td>
<td>Bezawit</td>
</tr>
<tr>
<td>Airgecho</td>
<td>Antimicrobial peptides are emerging topics of interest to fight against antimicrobial drug resistance. The bacterial toxin anti-toxin (TA) system, is less studied yet its promising results attracted attention of experts in the area. “So, what is the TA system in bacteria: its biology, mechanism of action, possibilities and implication in public health intervention” will be my topic during the systematic review course</td>
<td>Eva</td>
<td>Bezawit</td>
</tr>
<tr>
<td>Bezawit</td>
<td>Traditional healers contribution to disease control in a pastoralist community: systematic review.</td>
<td>Eva</td>
<td>Martina</td>
</tr>
<tr>
<td>Martina</td>
<td>Which instruments are used for measuring pain after stroke in the literature?</td>
<td>Eva</td>
<td>Melf</td>
</tr>
<tr>
<td>Reiten</td>
<td>Traditional healers contribution to disease control in a pastoralist community: systematic review.</td>
<td>Eva</td>
<td>Martina</td>
</tr>
<tr>
<td>Name</td>
<td>Tentative research question</td>
<td>Tutor</td>
<td>Opponent</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td>-------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Soheir</td>
<td>The Prevalence of Non-communicable Disease Risk Factors in East Africa Adults</td>
<td>Geir</td>
<td>Erik</td>
</tr>
<tr>
<td>Erik Oftedahl</td>
<td>metaanalyse eller review om assosiasjonen mellom sivilstatus og selvmord</td>
<td>Geir</td>
<td>Thandile</td>
</tr>
<tr>
<td>Thandile</td>
<td>Iron supplementation in children with severe acute malnutrition</td>
<td>Geir</td>
<td>Ingrid Kristine</td>
</tr>
<tr>
<td>Ingrid Kristine</td>
<td>To assess the effect of the Prosigna Breast Cancer Prognostic Gene Signature Assay (Prosigna test) for predicting disease recurrence in breast cancer patients.</td>
<td>Geir</td>
<td>Soheir</td>
</tr>
<tr>
<td>Name</td>
<td>Tentative research question</td>
<td>Tutor</td>
<td>Opponent</td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>Binyam</td>
<td>Does behavior change communication applied in developing text message intervention in maternal and child health?</td>
<td>Hilde</td>
<td>Berit</td>
</tr>
<tr>
<td>Josien</td>
<td>What are essential clinical decision making skills for Caesarean section in Low and Middle income countries</td>
<td>Hilde</td>
<td>Binyam</td>
</tr>
<tr>
<td>Berit</td>
<td>Is Postpartum depression a determinant when evaluating models of maternal care in developing countries?</td>
<td>Hilde</td>
<td>Josien</td>
</tr>
</tbody>
</table>