incorporating considerations of resource use into grading recommendations

Giovanni L. Pappagallo
The GRADE approach

- Considers
  - all factors to determine how confident we are in the results – quality of evidence
  - the evidence for each outcome in the review separately
  - magnitude of the effect
- Ensures
  - systematic process
  - transparency
**Health Care Question (PICO)**

**Systematic review**

**Studies**

- S1
- S2
- S3
- S4
- S5

**Outcomes**

- OC1: Important outcomes
- OC2
- OC3
- OC4: Critical outcomes

**Generate an estimate of effect for each outcome**

![Diagram showing the process of generating estimates and rating the quality of evidence.]

**Rate the quality of evidence for each outcome, across studies**

RCTs start with a high rating, observational studies with a low rating

**Rating is modified downward:**
- Study limitations
- Imprecision
- Inconsistency of results
- Indirectness of evidence
- Publication bias likely

**Rating is modified upward:**
- Large magnitude of effect
- Dose response
- Confounders likely minimize the effect

**Final rating of quality for each outcome:** high, moderate, low, or very low

**Rate overall quality of evidence**

(lowest quality among critical outcomes)
Health Care Question (PICO)
Systematic review

Studies
S1  S2  S3  S4  S5

Outcomes
OC1  OC2  OC3  OC4
Important outcomes
Critical outcomes

Generate an estimate of effect for each outcome

Rate the quality of evidence for each outcome, across studies
RCTs start with a high rating, observational studies with a low rating

Rating is modified downward:
- Study limitations
- Imprecision
- Inconsistency of results
- Indirectness of evidence
- Publication bias likely

Rating is modified upward:
- Large magnitude of effect
- Dose response
- Confounders likely minimize the effect

Final rating of quality for each outcome: high, moderate, low, or very low

Rate overall quality of evidence
(lowest quality among critical outcomes)
Choosing outcomes

Desirable outcomes
- lower mortality
- reduced hospital stay
- reduced duration of disease
- reduced resource expenditure

Undesirable outcomes
- adverse reactions
- the development of resistance
- costs of treatment

Recommendations must consider desirable and undesirable outcomes
Relative importance of outcomes

- Decision makers (and guideline authors) need to consider the relative importance of outcomes when balancing these outcomes to make a recommendation.
- Relative importance vary across populations.
- Relative importance may vary across patient groups within the same population.
- When considered critical - evaluate.
Generate an estimate of effect for each outcome

Rate the quality of evidence for each outcome, across studies
RCTs start with a high rating, observational studies with a low rating

Rating is modified downward:
- Study limitations
- Imprecision
- Inconsistency of results
- Indirectness of evidence
- Publication bias likely

Rating is modified upward:
- Large magnitude of effect
- Dose response
- Confounders likely minimize the effect

Final rating of quality for each outcome: high, moderate, low, or very low

Rate overall quality of evidence
(lowest quality among critical outcomes)
How to GRADE the evidence

Evidence varies from

- HIGH
- MODERATE
- LOW
- VERY LOW

The members of the Grade Working Group
Determinants of quality

5 factors that can lower quality

1. Limitations of detailed design and execution (risk of bias criteria)
2. Inconsistency (or heterogeneity)
3. Indirectness (PICO and applicability)
4. Imprecision (number of events and confidence intervals)
5. Publication bias
GRADE Evidence syntheses

• Is a summary of the key results from the systematic review for guideline panel members
  – Evidence profiles and Summary of Findings Tables

• Presents
  – the quality of the evidence
  – the magnitude of the effect
  – transparent description of judgments about evidence
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 for decisions
**Author(s):** Sturt AS, Dokubo EK  
**Date:** 2009-10-12  
**Question:** Should AZT/3TC/NVP vs AZT (+/- single dose NVP) be used for HIV-1 Infected Pregnant Women With CD4<200?  
**Settings:** Resource Limited  

## GRADE evidence profile

<table>
<thead>
<tr>
<th>Model</th>
<th>No of patients</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV-Transmission or Death at 1 Months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observational studies</td>
<td>7/69 (10.1%)</td>
<td>28/109 (25.7%)</td>
<td>157 fewer per 1000 (from 39 fewer to 211 fewer)</td>
<td><strong>VERY LOW</strong> CRITICAL</td>
</tr>
<tr>
<td>AZT/3TC/NVP</td>
<td>7/69 (10.1%)</td>
<td>28/109 (25.7%)</td>
<td>157 fewer per 1000 (from 39 fewer to 211 fewer)</td>
<td><strong>VERY LOW</strong> CRITICAL</td>
</tr>
<tr>
<td>AZT (+/- single dose NVP)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>-</td>
</tr>
<tr>
<td>Maternal Grade 3/4 Severe Adverse Events - not reported</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>-</td>
</tr>
<tr>
<td>Clinical, Immunologic, and Virologic Response to Anti-Retroviral Therapy - not reported</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>-</td>
</tr>
<tr>
<td>Infant Grade 3/4 Severe Adverse Events at Birth</td>
<td>4/69 (5.8%)</td>
<td>5/109 (4.6%)</td>
<td>5/109 (4.6%)</td>
<td><strong>VERY LOW</strong> IMPORTANT</td>
</tr>
<tr>
<td>Observational studies</td>
<td>4/69 (5.8%)</td>
<td>5/109 (4.6%)</td>
<td>5/109 (4.6%)</td>
<td><strong>VERY LOW</strong> IMPORTANT</td>
</tr>
<tr>
<td>AZT/3TC/NVP</td>
<td>4/69 (5.8%)</td>
<td>5/109 (4.6%)</td>
<td>5/109 (4.6%)</td>
<td><strong>VERY LOW</strong> IMPORTANT</td>
</tr>
<tr>
<td>AZT (+/- single dose NVP)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>-</td>
</tr>
<tr>
<td>Prematurity</td>
<td>14/109 (12.8%)</td>
<td>14/109 (12.8%)</td>
<td>57 fewer per 1000 (from 101 fewer to 162 more)</td>
<td><strong>VERY LOW</strong> IMPORTANT</td>
</tr>
<tr>
<td>Observational studies</td>
<td>5/69 (7.2%)</td>
<td>5/69 (7.2%)</td>
<td>5/69 (7.2%)</td>
<td><strong>VERY LOW</strong> IMPORTANT</td>
</tr>
<tr>
<td>AZT/3TC/NVP</td>
<td>5/69 (7.2%)</td>
<td>5/69 (7.2%)</td>
<td>5/69 (7.2%)</td>
<td><strong>VERY LOW</strong> IMPORTANT</td>
</tr>
<tr>
<td>AZT (+/- single dose NVP)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>-</td>
</tr>
<tr>
<td>Mother to Child Transmission at Birth</td>
<td>1/69 (1.4%)</td>
<td>9/109 (9.3%)</td>
<td>55 fewer per 1000 (from 81 fewer to 58 more)</td>
<td><strong>VERY LOW</strong> NOT IMPORTANT</td>
</tr>
<tr>
<td>Observational studies</td>
<td>1/69 (1.4%)</td>
<td>9/109 (9.3%)</td>
<td>55 fewer per 1000 (from 81 fewer to 58 more)</td>
<td><strong>VERY LOW</strong> NOT IMPORTANT</td>
</tr>
<tr>
<td>AZT/3TC/NVP</td>
<td>1/69 (1.4%)</td>
<td>9/109 (9.3%)</td>
<td>55 fewer per 1000 (from 81 fewer to 58 more)</td>
<td><strong>VERY LOW</strong> NOT IMPORTANT</td>
</tr>
<tr>
<td>AZT (+/- single dose NVP)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>-</td>
</tr>
<tr>
<td>Mother to Child Transmission at 1 Month</td>
<td>14/109 (12.8%)</td>
<td>14/109 (12.8%)</td>
<td>59 fewer per 1000 (from 3 fewer to 122 fewer)</td>
<td><strong>VERY LOW</strong> NOT IMPORTANT</td>
</tr>
<tr>
<td>Observational studies</td>
<td>2/69 (2.9%)</td>
<td>2/69 (2.9%)</td>
<td>2/69 (2.9%)</td>
<td><strong>VERY LOW</strong> NOT IMPORTANT</td>
</tr>
<tr>
<td>AZT/3TC/NVP</td>
<td>2/69 (2.9%)</td>
<td>2/69 (2.9%)</td>
<td>2/69 (2.9%)</td>
<td><strong>VERY LOW</strong> NOT IMPORTANT</td>
</tr>
<tr>
<td>AZT (+/- single dose NVP)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>-</td>
</tr>
<tr>
<td>Mother to Child Transmission Cumulative by 7 Months</td>
<td>2/69 (2.9%)</td>
<td>2/69 (2.9%)</td>
<td>2/69 (2.9%)</td>
<td><strong>VERY LOW</strong> NOT IMPORTANT</td>
</tr>
<tr>
<td>Observational studies</td>
<td>2/69 (2.9%)</td>
<td>2/69 (2.9%)</td>
<td>2/69 (2.9%)</td>
<td><strong>VERY LOW</strong> NOT IMPORTANT</td>
</tr>
<tr>
<td>AZT/3TC/NVP</td>
<td>2/69 (2.9%)</td>
<td>2/69 (2.9%)</td>
<td>2/69 (2.9%)</td>
<td><strong>VERY LOW</strong> NOT IMPORTANT</td>
</tr>
<tr>
<td>AZT (+/- single dose NVP)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>-</td>
</tr>
</tbody>
</table>

1 All denominators represent an available case analysis.  
2 Small numbers of events.  
3 Hepatic severe adverse events were not provided at this timepoint, the authors provided hematologic severe adverse events only.  
4 Confidence interval includes the null.  
5 Although a large effect, there are few events from one small observational study.
**QUESTO 1:** Nei pazienti affetti da carcinoma renale metastatico variante istologica a cellule chiare, pretrattati con inibitori di VEGF/VEGFr, il trattamento con **Nivolumab** può essere preso in considerazione in alternativa a un ritrattamento con inibitore di VEGF/VEGFr o un inibitore di mTOR?

**RACCOMANDAZIONE:**
Nei pazienti affetti da carcinoma renale metastatico variante istologica a cellule chiare pretrattati con un VEGF-TKI, il trattamento con Nivolumab dovrebbe essere preso in considerazione come approccio terapeutico di prima scelta.

Forza della raccomandazione: **Positiva Forte**

**Qualità delle Evidenze** (confronto vs Everolimus): **Bassa**. **Outcome** di beneficio: **Bassa**; **Outcome** di danno: **Bassa**

**Motivazioni/Commenti al rapporto Beneficio/Danno** (Confronto vs Everolimus):
**Outcome** di beneficio: 27% riduzione relativa del rischio di decesso, 22% riduzione relativa del rischio di progressione, miglioramento dei parametri di qualità della vita, della stomatite, della nausea e della anemia rispetto ad Everolimus.

**Outcome** di danno: maggiore incidenza di prurito (14%) rispetto ad everolimus (9.8%). Fatigue sovrapponibile

**Votazione rapporto Beneficio/Danno**

<table>
<thead>
<tr>
<th>Favorabile</th>
<th>Incerto (favorevole)</th>
<th>Incerto (sfavorevole)</th>
<th>Sfavorevole</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Votazione Forza della Raccomandazione**

<table>
<thead>
<tr>
<th>Positiva forte</th>
<th>Positiva debole</th>
<th>Negativa debole</th>
<th>Negativa forte</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Implicazioni per le ricerche future:**
Necessità di uno studio randomizzato vs Axitinib a progressione da una prima linea di trattamento con VEGFr-TKI; verifica dell’esistenza di marcatori predittivi di efficacia.
Resource use (cost) and

- Cost may be considered just another potentially important outcome – like mortality, morbidity, and quality of life – associated with alternative ways of managing patient problems.
Resource use (cost) and

- Cost may be considered just another potentially important outcome – like mortality, morbidity, and quality of life – associated with alternative ways of managing patient problems.

- The GRADE recommends that important differences in resource use should be included in SoF tables:
Balancing benefits and downsides

RATING QUALITY OF EVIDENCE AND STRENGTH OF RECOMMENDATIONS

GRADE: Incorporating considerations of resources use into grading recommendations

Guideline panellists have differing opinions on whether resource use should influence decisions in individual patients. As medical care costs rise, resource use considerations become more compelling, but panellists may find dealing with such considerations challenging.

Gordon H Guyatt

GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence


Journal of Clinical Epidemiology 66 (2013) 140–150
Consider Vs not consider resource use

• Guideline panels may or may not consider resource use in their judgments about the direction and strength of recommendations.

• Reasons for not considering resource use include:
  ✓ a lack of reliable data,
  ✓ the intervention is not useful and the effort of calculating resource use can be spared,
  ✓ the desirable effects so greatly outweigh any undesirable effects that resource considerations would not alter the final judgment,
  ✓ they have elected to leave resource considerations up to other decision makers.
Consider Vs not consider resource use

• Guideline panels may or may not consider resource use in their judgments about the direction and strength of recommendations.

• Reasons for not considering resource use include:
  - A lack of reliable data,
  - The intervention is not useful and the effort of calculating resource use can be spared,
  - The desirable effects so greatly outweigh any undesirable effects that resource considerations would not alter the final judgment,
  - They have elected to leave resource considerations up to other decision makers.

A guideline panel may legitimately ignore considerations of resource use, and make recommendations solely on the basis of other advantages and disadvantages of the alternatives being considered.
Cost may be considered just another potentially important outcome – like mortality, morbidity, and quality of life – associated with alternative ways of managing patient problems.

The GRADE recommends that important differences in resource use should be included in SoF tables:

- identify items of resource use that are potentially important to decision makers
Identifying potentially important resource use

• Clearly state the viewpoint (perspective) from which the recommendations are being made:
  ✓ a broad viewpoint that includes all important health care, non-health care, and patient and informal caregiver resources, regardless of who pays for them
  Vs
  ✓ limit considerations of resource use (and costs and relative efficiency) to those resources that incur a cost to the health and social care system.

• GRADE suggests that a broad perspective is desirable
Resource use (cost) and GRADE

• Cost may be considered just another potentially important outcome – like mortality, morbidity, and quality of life – associated with alternative ways of managing patient problems.

• The GRADE recommends that important differences in resource use should be included in SoF tables:
  ✓ identify items of resource use that are potentially important to decision makers
  ✓ find evidence for the difference in resource use between the options being compared
  ✓ present resource use in natural units
Resource implications considered

• Evidence profiles and SoF tables should always present resource use, not just monetary values (as monetary values for the same resource will vary depending on setting).

• It is recommended that natural units be used to estimate resource use.
  ✓ (e.g., days in hospital, minutes of clinician time)

• A balance sheet should inform judgments about whether the net benefits are worth the incremental costs.

• Formal economic modeling may – or may not – be helpful
Formal economic modelling

- Formal economic modeling results in cost per unit benefit achieved:
  - e.g., cost per natural unit, cost per quality adjusted life year gained, cost and benefits valued in monetary values
- High probability of being flawed or biased
- GRADE working group recommends not including cost-effectiveness or cost-utility models in evidence profiles.
Resource use (cost) and GRADE

- Cost may be considered just another potentially important outcome – like mortality, morbidity, and quality of life – associated with alternative ways of managing patient problems.

The GRADE recommends that important differences in resource use should be included in SoF tables:

- identify items of resource use that are potentially important to decision makers
- find evidence for the difference in resource use between the options being compared
- present resource use in natural units
- rate the confidence in effect estimates

---


Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria

Silvia Evers, Mariëlle Goossens, Henrica de Vet, Maurits van Tulder, André Ament
Cost may be considered just another potentially important outcome – like mortality, morbidity, and quality of life – associated with alternative ways of managing patient problems.

The GRADE recommends that important differences in resource use should be included in SoF tables:

- identify items of resource use that are potentially important to decision makers
- find evidence for the difference in resource use between the options being compared
- present resource use in natural units
- rate the confidence in effect estimates

None of published checklist specifically constructed to assess the quality of a body of evidence as defined by GRADE – *that is, the confidence in estimates of effect*.
Resource use (cost) and GRADE

Cost may be considered just another potentially important outcome – like mortality, morbidity, and quality of life – associated with alternative ways of managing patient problems.

The GRADE recommends that important differences in resource use (cost) should be included in SoF tables:

- identify items of resource use that are potentially important to decision makers
- find evidence for the difference in resource use between the options being compared
- present resource use in natural units
- rate the confidence in effect estimates

Confidence for effect estimates for each important/critical economic outcome to be appraised explicitly using the same criteria as for health outcomes.
Determinants of quality

5 factors that can lower quality

1. Limitations of detailed design and execution (risk of bias criteria)
2. Inconsistency (or heterogeneity)
3. Indirectness (historical, geographical, or publication bias)
4. Imprecision (random error, confidence intervals)
5. Publication bias

- difference in disease severity
- differences in amount of attention, ancillary treatment, diagnostic investigation
- incomplete outcome data

Biased estimates of resource use
• RND trials unrepresentative of routine practice
• Older studies (technologies & organization)
• Teaching / research-based hospitals Vs nonteaching hospitals

5 factors that limit the strength of evidence (risk of bias)
1. Limitations (risk of bias)
2. Inconsistency (heterogeneity)
3. **Indirectness** *(PICO and applicability)*
4. Imprecision *(number of events and confidence intervals)*
5. Publication bias
Clinical trials often underpowered to detect differences in resource use
Resource use (cost) and attitudes differ as to whether costs should influence a doctor’s decision about treating individual patients.