Determinants Used to Justify the Strength of Recommendations among Korean Clinical Practice Guidelines

Ein-Soon Shin 1,* Ji-Eun Jang 2,* Ji-Yun Yeon 1, Da-sol Kim 1, Sung-Goo Chang 3 and Yoon-Seong Lee 4

1Research Agency for Clinical Practice Guidelines, Research Center, Korean Academy of Medical Sciences, Seoul, Korea
2Department of Medical Sciences, Ewha Womans University Graduate School, Seoul, Korea
3Department of Urology, Kyung Hee University School of Medicine, Seoul, Korea
4Department of Forensic Medicine, Seoul National University College of Medicine, Seoul, Korea

ABSTRACT

Background: A standardized systematic approach to grade evidence and the strength of recommendations is important for guideline users to minimize bias and help interpret the most suitable decisions at the point of care. The study aims to identify and classify determinants used to make judgement for the strength of recommendations among 56 Korean clinical practice guidelines (CPGs), and explore strong recommendations based on low quality of evidence.

Methods: Determinants used in the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach among 34 CPGs which have reported both strength of recommendations and level of evidence were reviewed.

Results: Five of 34 CPGs (14.7%) considered quality of evidence, benefits and harms, patients’ values and preferences, and costs. And 24 of 34 CPGs (70.6%) considered both magnitude of effect and feasibility as additional determinants. Judgement table was not widely provided for use to translate evidence into recommendations. Eighty-two of 121 recommendations (67.8%, ranged 20.0% to 100.0%) among 11 CPGs using the same judgement scheme showed ‘strong’ strength of recommendations based on low or very low quality of evidence. Among 5 paradigmatic situations that justify strong recommendations based on low or very low evidence, situation classified as ‘potential equivalence, one option clearly less risky or costly’ was 87.8% for 82 strong recommendations. Situation classified as ‘uncertain benefit, certain harm’ was 4.9%.

Conclusion: There is a need to introduce and systematize an evidence-based grading system. Using judgement table to justify the strength of recommendations and applying the 5 paradigmatic situations mentioned above is also recommended in the near future.

Keywords: Clinical Practice Guidelines; GRADE Approach; Strength of Recommendations; Quality of Evidence; Judgement Table

INTRODUCTION

To minimize bias and help interpret the most suitable decisions at the point of care for users of clinical practice guidelines (CPGs), a standardized systematic approach to grade evidence and the strength of recommendations is important. However, because several grading systems have
shown inconsistency and wide variations to grade quality of evidence and recommendations,\textsuperscript{2,3} it is difficult for guideline users to understand the message. Different grading systems include Grading of Recommendations Assessment, Development and Evaluation (GRADE), National Institute for Health and Clinical Excellence (NICE), Scottish Intercollegiate Guideline Network (SIGN), and Centre for Evidence-Based Medicine, Oxford (CEBM) are used to give a clinical recommendation in a CPG which have an influence on the clinician’s point of view and can significantly change clinician’s decisions and clinical outcomes.\textsuperscript{1}

The GRADE is an emerging consensus on rating quality of evidence and strength of recommendations. GRADE provides an explicit, comprehensive, transparent, and pragmatic approach\textsuperscript{4-6} and is increasingly being adopted by organizations worldwide.\textsuperscript{7-9} Separating the judgements regarding the quality of evidence from judgements about the strength of recommendations is a critical and defining feature of this new grading system.\textsuperscript{10,11} The GRADE approach also provides a framework to move from evidence to the recommendation.\textsuperscript{12-15} The strength of recommendations\textsuperscript{16-18} depends on estimates of magnitude of effect,\textsuperscript{19-21} estimates of values and preference and their variability,\textsuperscript{22} confidence in each of these estimates, and resource use\textsuperscript{23,24} considerations.

GRADE guidance allows five situations that justify strong recommendations in the face of low or very low confidence in estimates for critical outcomes.\textsuperscript{7,25} Despite GRADE guidance’s warning against the transparent approach, World Health Organization (WHO) guidelines showed that approximately one-half of the recommendations were based on low or very low confidence.\textsuperscript{26-28} It raises concerns about whether GRADE is being optimally applied in the WHO guideline development process.

In Korea, few CPGs are trying to adopt GRADE.\textsuperscript{29-33} They considered little bit different components to grade recommendations and utilize varied grading systems. Some guidelines considered only an evidence and there was lack of information on the resource use and patients’ value and preferences mostly. There has been no review of factors or grading systems used to grade recommendations in Korea. Therefore, it is necessary to obtain a clear rationale for the adoption of the GRADE approach which is currently accepted worldwide through the analysis of the current situation, particularly the situation where strong recommendations are based on low evidence.

The study aims to identify and classify determinants used to make judgement for the strength of recommendations among 56 Korean CPGs, and explore strong recommendations based on low quality of evidence.

**METHODS**

**Data source**
To identify determinants used to make judgement for the strength of recommendations in Korea, 56 CPGs developed based on the academic societies that are listed in the Korean Medical Guideline Information Center (http://www.guideline.or.kr/) were included. Of the 56 CPGs, 27 (48.2%) were developed between 2001 and 2010, and 29 (51.8%) were developed between 2011 and the end of June 2017. Fifty-six CPGs were developed in a total of 33 academic societies; 22 of 33 academic societies (66.7%) developed 1 CPG, 6 of 33 (18.2%) developed 2 CPGs, and 5 of 33 (15.1%) developed 3 or more CPGs.
Data abstraction process
We conducted 3-step process. First, 56 Korean CPGs were classified by reporting level of evidence and/or the strength of recommendations. Second, determinants used to make judgment for the strength of recommendations among 34 CPGs were identified (Appendix 1). The following determinants used GRADE approach worldwide is included: quality of evidence, balance between benefits and harms, patients’ values and preferences, and costs (resource use). And additional determinants include absolute magnitude of effect and feasibility (Appendix 2). To collect data on feasibility, we considered two aspects, feasible to implement and whether to consider barriers. Third, to explore the strength of recommendations based on low or very low quality of evidence, 11 CPGs using the same judgement scheme (strong, weak) were reviewed. We will compare a summary on the strength of recommendations by low or very low quality of evidence with WHO CPGs. Fourth, classification by 5 paradigmatic situations that justify strong recommendations based on low or very low certainty of the evidence in GRADE approach was performed. In case of ‘life-threatening situations and potential equivalence’ and ‘one option is clearly less risky or costly,’ the strength of the recommendation in favor of the intervention can be given a strong. In case of ‘uncertain benefit with certain harm,’ ‘potential catastrophic harm,’ and ‘high similar benefits, one option potentially more risk or costly,’ the strength of the recommendation against the intervention can be given a strong (Appendix 3).

RESULTS
Classification on reporting for the strength of recommendations and/or level of evidence among 56 Korean CPGs were as follows: 18 of 56 CPGs (32.1%) did not report both strength of recommendations and level of evidence, 2 CPG (3.6%) reported level of evidence only, 2 CPG (3.6%) reported strength of recommendations only, and 34 CPGs (60.7%) reported both strength of recommendations and level of evidence (Table 1).

Determinants used in the GRADE approach and judgement table used to justify the strength of recommendations among 34 CPGs which have reported both strength of recommendations and level of evidence were as follows: 12 of 34 CPGs (35.3%) considered only quality of evidence, and 12 CPGs (35.3%) considered quality of evidence and benefits and harms. Five of 34 CPGs (14.7%) considered quality of evidence, benefits and harms, patients’ values and preferences, and costs such as GRADE approach. And 24 of 34 CPGs (70.6%) considered both magnitude of effect and feasibility as additional determinants (Table 2). Judgement table was not widely provided for use to translate evidence into recommendations. There was inconsistency between information in method part and information in background on the determinants for making judgement the strength of recommendations (Appendix 1).

<table>
<thead>
<tr>
<th>Table 1. Classification of reporting for the strength of recommendations and/or level of evidence among 56 Korean CPGs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporting items</td>
</tr>
<tr>
<td>Reported quality or level of evidence only</td>
</tr>
<tr>
<td>Reported strength of recommendations only</td>
</tr>
<tr>
<td>Reported neither strength of recommendations nor quality or level of evidence</td>
</tr>
<tr>
<td>Reported both strength of recommendations and quality or level of evidence</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

CPG = clinical practice guideline.
Strength of recommendations based on low quality of evidence among 11 Korean CPGs using the same judgement scheme (strong, weak or conditional) were as follows: 82 of 121 recommendations (67.8%, ranged 20.0% to 100.0%) showed ‘strong’ strength of recommendations based on low or very low quality of evidence. Six of 11 CPGs (54.5%) made judgement ‘strong’ for more than 70% of recommendations based on low or very low quality of evidence (Table 3). The determinants used to make judgement varied among 11 CPGs (Appendix 1). Table 4 showed comparison a summary on the strength of recommendations by low or very low quality of evidence. Sixty-seven point eight % of recommendations was given ‘strong’ recommendations based on low or very low quality of evidence in 11 Korean CPGs, follows 67.4% for 44 CPGs of 20 countries, 53.0% for 43 CPGs of WHO, and 48.4% for 8 CPGs (human immunodeficiency virus [HIV] and tuberculosis [TB] only) of WHO. It was ranged from 57.8% to 74.5% for low quality of evidence and ranged from 27.6% to 57.8% for very low quality of evidence.

Classification by 5 paradigmatic situations that justify strong recommendations based on low or very low certainty of the evidence is shown in Table 5. Situation classified as ‘potential equivalence, one option clearly less risky or costly’ was 87.8% for 82 strong recommendations; 86.8% (66 of 76) in low quality of evidence and 100% (6 of 6) in very low quality of evidence. Situation classified as ‘uncertain benefit and certain harm’ was 4.9% (4 of 82 recommendations), and 7.3% (6 of 82) was classified as other situation including no intervention.
DISCUSSION

Determinants used to justify the strength of recommendations among Korean CPGs were identified and classified. Five different grading systems each considered different factors; 1) only quality of evidence was considered, 2) quality of evidence and benefits and harms were considered, 3) quality of evidence, benefits and harms, and patients' values and preferences were considered, 4) quality of evidence, benefits and harms, and cost were considered, and 5) quality of evidence, benefits and harms, patients' values and preferences, and cost were considered. We also considered additional factors including absolute magnitude of effect and feasibility. GRADE approach has been adopted by 5 of 34 CPGs (14.7%). And 24 of 34 CPGs (70.6%) considered both magnitude of effect and feasibility as additional determinants (Table 2). However, we could not find judgement table with transparent explanation and evidence to justify the strength of recommendations among Korean CPGs. Template and judgement table by recommendation has been provided by WHO (Appendix 2). It is a summary table for translating evidence of effectiveness into recommendations. It will be helpful to the guideline end users at the point of care if guideline developers can present a judgement table to clearly explain what factors are considered important.

If CPGs do not provide a structured judgement table, it would require laborious review of the method and/or background for every recommendation listed to discover which determinants were considered. Also, there were some discrepancies between information in method part
and information in background on the determinants for making judgement the strength of recommendations (Table 2). To overcome these shortcomings, we should adopt a systematic and transparent GRADE approach which is used by many organizations and associations worldwide. We also reviewed current situation regarding strong recommendations based on low or very low quality of evidence among Korean CPGs. Lower percentage of ‘strong’ recommendations in very low quality of evidence (31.6% vs. 57.8% in 20 countries) is shown in Table 4.

Alexander et al.27,28 explained the reason why WHO guideline developers make strong recommendations inconsistent with GRADE guidance. The main reason was limitations in their understanding and optimal application of GRADE. To utilize GRADE approach consistent with GRADE guidance, it requires training of guideline developers or panelists and formal processes to maximize adherence to GRADE principles.7 GRADE guidance present five paradigmatic situations that justify strong recommendation based on low or very low certainty of the evidence (Appendix 3).25,27 Among 5 paradigmatic situations, situation classified as ‘potential equivalence, one option clearly less risky or costly’ was 87.8% for 82 strong recommendations. Situation classified as ‘uncertain benefit and certain harm’ was 4.9% (4 of 82). There was no recommendation classified as ‘life-threatening situation’ on the strong recommendations in favor of the intervention, and there was no recommendation classified as ‘potential catastrophic harm,’ and ‘high similar benefits, one option potentially more risk or costly’ on the strong recommendations against the intervention. Especially, when guideline developers have to make a decision the strong recommendation based on low evidence, it will be better to consider the five paradigmatic situations mentioned above.

Limitations for carrying out this study are as follow: First, since only 56 CPGs in the Korean Medical Guideline Information Center are analyzed among all Korean CPGs, there was lack of representative. At the end of 2013, 115 CPGs were reported to have been developed by 45 societies and organizations.40 Second, 34 Korean CPGs were using different grading systems. Therefore, we performed data abstraction on the 11 CPGs using the same judgement scheme (strong, week or conditional), because we could not explore judgement of strong recommendation based on low or very low level of evidence for all CPGs included in the study. Third, when extracting the factors considered in the recommendation grading, we classified them based on the information provided in the method part. However, there were many cases in which the background part of the recommendation actually provided information that considers other factors. We summarized the two kinds of information in Appendix 1 because there are differences according to whether classification of factors considered in grading recommendation is based on information of method part or information of background part.

In conclusion, there were several grading systems with wide variations to grade quality of evidence and strength of recommendations, and strong recommendations based on low or very low quality of evidence were very common among Korean CPGs. There is a need to introduce and systematize an evidence-based grading system. It is also necessary to aggressively review, apply, and disseminate the worldwide GRADE approach that grades recommendations in consideration of important factors including quality of evidence, benefits and harms, patients’ value and preferences, and cost in Korea. Before applying and disseminating the GRADE approach on rating quality of evidence and strength of recommendations in Korea, it is necessary to study what external and internal barriers are to use this grading system in advance. Judgement table was not widely provided for use to translate evidence into recommendations, and the five paradigmatic situations were not used against low evidences to decide strength of recommendations to be made.
REFERENCES


Determinants for Strength of Recommendations


22. Fraenkel L. Incorporating patients’ preferences into medical decision making. Med Care Res Rev 2013;70(1 Suppl):80S-93S.


https://jkms.org

https://doi.org/10.3346/jkms.2018.33.e79

8/13


### Appendix 1. Data abstraction on the determinants used to make judgement for the strength of recommendations

#### Guidelines (n = 34)

<table>
<thead>
<tr>
<th>Guidelines (n = 34)</th>
<th>Quality of evidence</th>
<th>Benefits and harms</th>
<th>Patients’ values and preferences</th>
<th>Cost (resource use)</th>
<th>Absolute magnitude of effect</th>
<th>Feasibility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Part of method</td>
<td>Part of background</td>
<td>Part of method</td>
<td>Part of background</td>
<td>Part of background</td>
<td>Part of method</td>
</tr>
</tbody>
</table>

**A. Determinants information provided in a methodology part:**

1. Quality or level of evidence only (n = 12, 35.3%)
   - Treatment Guideline for Diabetes (2007)
     - Quality or level of evidence
   - Korea Breast Cancer Society Practice Recommendations of Breast Cancer (2008)
     - Quality or level of evidence
   - Diagnostic Guideline of Ulcerative Colitis (2009)
     - Quality or level of evidence
   - Diagnostic Guideline of Crohn’s Disease (2009)
     - Quality or level of evidence
   - Clinical Practice Guideline for Stroke Rehabilitation in Korea 2009 (2009)
     - Quality or level of evidence
     - Quality or level of evidence
   - Intravenous Infusion Nursing Practice Guideline (2012)
     - Quality or level of evidence
   - Clinical Practice Guideline for Stroke Rehabilitation in Korea (2012)
     - Quality or level of evidence
   - Practical Guidelines for the Surgical treatment of Gallbladder Cancer (2014)
     - Quality or level of evidence
     - Quality or level of evidence
     - Quality or level of evidence
     - Quality or level of evidence
   - Korean Guideline for Sexually Transmitted Infections (2016)
     - Quality or level of evidence

**B. Determinants information provided in methodology part:**

2. Quality or level of evidence and benefits and harms (n = 12, 35.3%)
   - Evidence Based Medicine Guideline for Post-traumatic Stress Disorder (2008)
     - Quality or level of evidence and benefits and harms
   - Acute Inflammatory Facial Nerve Paralysis (2010)
     - Quality or level of evidence and benefits and harms
     - Quality or level of evidence and benefits and harms
   - Clinical Practice Guideline for CAPD Peritonitis (2012)
     - Quality or level of evidence and benefits and harms
     - Quality or level of evidence and benefits and harms
   - Guideline for Intervention Recanalization of Lower Extremity Artery (2014)
     - Quality or level of evidence and benefits and harms
   - Guidelines for initial management and transfer of burns at emergency department (2015)
     - Quality or level of evidence and benefits and harms
   - Evidence-based Recommendations for Hypertension in Primary Care (2016)
     - Quality or level of evidence and benefits and harms
   - Evidence-based Recommendations for Type 2 Diabetes in Primary Care (2016)
     - Quality or level of evidence and benefits and harms
   - Diagnosis and Treatment of Lower Extremity Deep Vein Thrombosis: Korean Practical Guideline (2016)
     - Quality or level of evidence and benefits and harms
   - Evidence-based Clinical Imaging Guidelines (2016)
     - Quality or level of evidence and benefits and harms
   - Evidence-based Recommendations for Dyslipidemia in Primary Care (2017)
     - Quality or level of evidence and benefits and harms

(continued to the next page)
### Appendix 1. Data abstraction on the determinants used to make judgement for the strength of recommendations

#### Guidelines (n = 34)

<table>
<thead>
<tr>
<th>Four GRADE components</th>
<th>Two additional components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of evidence</td>
<td>Absolute magnitude of effect*</td>
</tr>
<tr>
<td>Benefits and harms</td>
<td>Feasibility</td>
</tr>
<tr>
<td>Patients’ values and preferences</td>
<td>Feasible to implement</td>
</tr>
<tr>
<td>Cost (resource use)</td>
<td>Consider barriers</td>
</tr>
<tr>
<td>Part of method</td>
<td>Part of background</td>
</tr>
<tr>
<td>Part of method</td>
<td>Part of background</td>
</tr>
<tr>
<td>Part of method</td>
<td>Part of background</td>
</tr>
<tr>
<td>Part of method</td>
<td>Part of background</td>
</tr>
<tr>
<td>Part of method</td>
<td>Part of background</td>
</tr>
</tbody>
</table>

#### C. Determinants information provided in methodology part: ③ Quality or level of evidence, benefits and harms, and patient’s values and preferences (n = 3, 8.8%)

- Korean Clinical Practice Guideline for Gastric Cancer (2012)[^36]
  - Quality or level of evidence: ★
  - Benefits and harms: ★
  - Patients’ values and preferences: ★

- Korean Clinical Practice Guidelines for Common Bile Duct Stones (2013)[^37]
  - Quality or level of evidence: ★
  - Benefits and harms: ★
  - Patients’ values and preferences: ★

- Korean Clinical Practice Guideline for Benign Prostate Hyperplasia (2015)[^38]
  - Quality or level of evidence: ★
  - Benefits and harms: ★
  - Patients’ values and preferences: ★

#### D. Determinants information provided in methodology part: ④ Quality or level of evidence, benefits and harms, and cost (n = 2, 5.9%)

- KASL Clinical Practice Guidelines: Management of Chronic Hepatitis B (2011)[^34]
  - Quality or level of evidence: ★
  - Benefits and harms: ★
  - Patients’ values and preferences: ★

- Cough Guideline 2014 (2014)[^38]
  - Quality or level of evidence: ★
  - Benefits and harms: ★
  - Patients’ values and preferences: ★

#### E. Determinants information provided in methodology part: ⑤ Quality or level of evidence, benefits and harms, patient’s values and preferences, and cost (n = 5, 14.7%)

- Practice Guideline for Gynecologic Cancer Version 2.0 (2010)[^39]
  - Quality or level of evidence: ★
  - Benefits and harms: ★
  - Patients’ values and preferences: ★

- Revision of the Guidelines for Chronic Obstructive Pulmonary Disease (2013)[^40]
  - Quality or level of evidence: ★
  - Benefits and harms: ★
  - Patients’ values and preferences: ★

- Korean Clinical Practice Guideline for Colon and Rectal Cancer v1.0 (2012)[^41]
  - Quality or level of evidence: ★
  - Benefits and harms: ★
  - Patients’ values and preferences: ★

- Revision of the Guidelines for Chronic Obstructive Pulmonary Disease (2014)[^42]
  - Quality or level of evidence: ★
  - Benefits and harms: ★
  - Patients’ values and preferences: ★

- Korean Guidelines for Asthma, revised (2014)[^33]
  - Quality or level of evidence: ★
  - Benefits and harms: ★
  - Patients’ values and preferences: ★

---

GRADE = Grading of Recommendations Assessment, Development and Evaluation.

*Marked Δ: cited existing meta-analysis or systematic review, ○: analyzed, performed meta-analysis or systematic review.
## Appendix 2. A template and sample of judgement table to justify the strength of recommendations

<table>
<thead>
<tr>
<th>Judgement sample</th>
<th>Vitamin D supplementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention</td>
<td>The higher the quality of the evidence, the stronger the recommendation. However, when “low” or “very-low” quality, consider more carefully the other criteria below in deciding the strength of the recommendation.</td>
</tr>
<tr>
<td>Quality of the evidence</td>
<td>☐ High  ☐ Moderate  ☒ Low  ☐ Very low</td>
</tr>
<tr>
<td>Values and preferences</td>
<td>This refers to values placed by health workers, policy-makers, patients and other stakeholders on the intended outcomes of interventions. If there is wide variability between values and preferences of various stakeholders, it is less likely to have a strong recommendation.</td>
</tr>
<tr>
<td>☒ No significant variability  ☐ Significant variability</td>
<td></td>
</tr>
<tr>
<td>Absolute magnitude of effect</td>
<td>This refers to the potential of the intervention to have large effects. The effects can be enhanced by combining with other interventions. Consider what are the possible associations (or “bundles”) that will enhance the effect. The larger the potential effects and for longer periods of time, the more likely to have a strong recommendation.</td>
</tr>
<tr>
<td>☐ Large effect in the long term  ☒ Small effect for short duration</td>
<td></td>
</tr>
<tr>
<td>Balance of benefits versus disadvantages</td>
<td>Benefits should consider the intended effects of the intervention. Disadvantages should consider the potentially negative effects of the intervention, as well as the unintended effects. The less potentially negative effects, the more likely to have a strong recommendation.</td>
</tr>
<tr>
<td>☐ Benefits clearly outweigh disadvantages  ☒ Benefits and disadvantages are balanced  ☐ Disadvantages clearly outweigh benefits</td>
<td></td>
</tr>
<tr>
<td>Resource use</td>
<td>The resource needed for implementing the recommendation may comprise financial resources, human resources, and infrastructure or equipment. Ideally, the benefits of the intervention should come at reasonable, affordable and sustainable costs. One should consider that capital costs, such as for infrastructure development, even if initially high, may yield benefits in the long run. The higher the incremental or recurrent costs, all other things being equal, the less likely it is to have a strong recommendation.</td>
</tr>
<tr>
<td>☒ Less resource intensive  ☐ More resource intensive</td>
<td></td>
</tr>
<tr>
<td>Feasibility</td>
<td>All interventions require political commitment and wide stakeholder engagement as a prerequisite. In addition, “technical” feasibility requires functional organizational and institutional structures necessary to manage, follow through, and monitor the implementation of the recommendation. The elements of technical feasibility vary widely by country or context, but if these elements are likely to be functional in a wide variety of settings, the more likely is to have a strong recommendation.</td>
</tr>
<tr>
<td>☒ Yes, globally  ☐ Yes, conditionally</td>
<td></td>
</tr>
<tr>
<td>Overall ranking</td>
<td>Strength of the recommendation.</td>
</tr>
<tr>
<td>☐ Strong recommendation  ☒ Weak recommendation</td>
<td></td>
</tr>
<tr>
<td>Conclusion about recommendation direction</td>
<td>☐ In favour of the intervention  ☒ Against the intervention</td>
</tr>
</tbody>
</table>

1This recommendation was made strong against the intervention despite of the low quality of evidence due to the fact that some participants expressed concerns about the limited evidence on safety of vitamin D supplementation during pregnancy. The guideline development group also noted that several studies were ongoing on this topic which may lead to a change in the evidence base in the future.

### Determinants for Strength of Recommendations

#### Appendix 3. Five situations that justify strong recommendations based on low or very low certainty of the evidence in GRADE approach\(^{35,27}\)

<table>
<thead>
<tr>
<th>Paradigmatic situation</th>
<th>Certainty of the evidence for health outcomes (quality of evidence)</th>
<th>Balance of benefits and harms</th>
<th>Values and preferences</th>
<th>Resource considerations</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life-threatening situation</td>
<td>Low or very low certainty</td>
<td>Immaterial (very low to high)</td>
<td>Intervention may reduce mortality in a life threatening situation; adverse events not prohibitive</td>
<td>A very high value is placed on an uncertain but potentially life preserving benefit</td>
<td>Small incremental cost (or resource use) relative to the benefits justify the intervention</td>
</tr>
<tr>
<td>Example 1.</td>
<td></td>
<td></td>
<td>Indirect evidence from seasonal influenza suggests that patients with avian influenza may benefit from the use of oseltamivir (low certainty of the evidence). Given the high mortality of the disease and the absence of effective alternatives, the WHO made a strong recommendation in favor of the use of oseltamivir rather than no treatment in patients with avian influenza.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example 2.</td>
<td></td>
<td></td>
<td>Fresh frozen plasma or vitamin K in a patient receiving warfarin with elevated INR and an intracranial bleed. Only low-quality evidence supports the benefits of limiting the extent of the bleeding.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncertain benefit, Low or very low certain harm</td>
<td>High or moderate</td>
<td>Possible but uncertain benefit; substantial established harm</td>
<td>A much higher value is placed on the adverse events in which we are confident than in the benefit, which is uncertain</td>
<td>High incremental cost (or resource use) relative to the benefits may not justify the intervention</td>
<td>Strong recommendation against</td>
</tr>
<tr>
<td>Example 1.</td>
<td></td>
<td></td>
<td>In patients with idiopathic pulmonary fibrosis, treatment with azathioprine plus prednisone offers a possible but uncertain benefit in comparison with no treatment. The intervention, however, is associated with a substantial established harm. An international guideline made a recommendation against the combination of corticosteroids plus azathioprine in patients with idiopathic pulmonary fibrosis.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example 2.</td>
<td></td>
<td></td>
<td>Head-to-toe CT/MRI screening for cancer. Low-quality evidence of benefit of early detection but high-quality evidence of possible harm and/or high cost (strong recommendation against this strategy).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potential equivalence, one option clearly less risky or costly</td>
<td>Low or very low</td>
<td>High or moderate</td>
<td>Magnitude of benefit apparently similar, though uncertain for alternatives; we are confident in less harm or cost for one of the competing alternatives</td>
<td>A high value is placed on the reduction in harm</td>
<td>Strong recommendation for less harmful/less expensive</td>
</tr>
<tr>
<td>Example 1.</td>
<td></td>
<td></td>
<td>Low-certainty evidence suggests that initial Helicobacter pylori eradication in patients with early-stage extradural marginal zone (MALT) B-cell lymphoma results in similar rates of complete response in comparison with the alternatives of radiation therapy or gastrectomy, but with high confidence of less harm, morbidity, and cost. Consequently, UpToDate made a strong recommendation in favor of H. pylori eradication rather than radiotherapy in patients with MALT lymphoma.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example 2.</td>
<td></td>
<td></td>
<td>H. pylori eradication in patients with early stage gastric MALT lymphoma with H. pylori positive. Low-quality evidence suggests that initial H. pylori eradication results in similar rates of complete response compared with the alternatives of radiation therapy or gastrectomy; high-quality evidence suggests less harm/morbidity,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High similar benefits, one option potentially more risky or costly</td>
<td>High or moderate Low or very low</td>
<td>Established that magnitude of benefit is similar for alternative management strategies; best (though uncertain) estimate is that one alternative has appreciably greater harm</td>
<td>A high value is placed on avoiding the potential increase in harm</td>
<td>High incremental cost (or resource use) relative to the benefits may not justify one of the alternatives</td>
<td>Strong recommendation against the intervention with possible greater harm</td>
</tr>
<tr>
<td>Example 1.</td>
<td></td>
<td></td>
<td>In women requiring anticoagulation and planning conception or in pregnancy, high certainty of the evidence suggests similar effects of different anticoagulants. However, indirect evidence (low certainty of the evidence) suggests potential harm to the unborn infant with oral direct thrombin (e.g., dabigatran) and factor Xa inhibitors (e.g., rivaroxaban, apixaban). The AT9 guidelines recommended against the use of such anticoagulants in women planning conception or in pregnancy.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example 2.</td>
<td></td>
<td></td>
<td>Hypertension in women planning conception and in pregnancy. Strong recommendations for labetalol and nifedipine and strong recommendations against ACE inhibitors and ARBs; all agents have high-quality evidence of equivalent beneficial outcomes, with low-quality evidence for greater adverse effects with ACE inhibitors and ARBs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potential catastrophic harm</td>
<td>Immaterial (very low to high)</td>
<td>Low or very low</td>
<td>Potential important harm of the intervention, magnitude of benefit is variable</td>
<td>A high value is placed on avoiding potential increase in harm</td>
<td>High incremental cost (or resource use) relative to the benefits, may not justify the intervention</td>
</tr>
<tr>
<td>Example 1.</td>
<td></td>
<td></td>
<td>In men with androgen deficiency, testosterone supplementation likely improves quality of life. Low-certainty evidence suggests that testosterone increases cancer spread in patients with prostate cancer. The US Endocrine Society made a recommendation against testosterone supplementation in patients with prostate cancer.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example 2.</td>
<td></td>
<td></td>
<td>Testosterone in males with or at risk of prostate cancer, high-quality evidence for moderate benefits of testosterone treatment in men with symptomatic androgens deficiency to improve bone mineral density and muscle strength. Low-quality evidence for harm in patients with or at risk of prostate cancer.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

GRADE = Grading of Recommendations Assessment, Development and Evaluation, WHO = World Health Organization, INR = international normalized ratio, CT/MRI = computed tomography/magnetic resonance imaging, MALT = mucosa-associated lymphoid tissue, ACE = angiotensin-converting enzyme, ARB = angiotensin receptor blocker.

https://doi.org/10.3346/jkms.2018.33.e79

https://jkms.org