Going from evidence to recommendations

Gordon H Guyatt, Andrew D Oxman, Regina Kunz, Yngve Falck-Ytter, Gunn E Vist, Alessandro Liberati, Holger J Schünemann and for the GRADE Working Group

BMJ 2008;336:1049-1051
doi:10.1136/bmj.39493.646875.AE

Updated information and services can be found at:
http://bmj.com/cgi/content/full/336/7652/1049

These include:

References
This article cites 4 articles, 2 of which can be accessed free at:
http://bmj.com/cgi/content/full/336/7652/1049#BIBL
2 online articles that cite this article can be accessed at:
http://bmj.com/cgi/content/full/336/7652/1049#otherarticles

Rapid responses
One rapid response has been posted to this article, which you can access for free at:
http://bmj.com/cgi/content/full/336/7652/1049#responses
You can respond to this article at:
http://bmj.com/cgi/eletter-submit/336/7652/1049

Email alerting service
Receive free email alerts when new articles cite this article - sign up in the box at the top left of the article

Correction
A correction has been published for this article. The correction is available online at:
http://bmj.com/cgi/content/full/336/7658/0-b

Notes

To order reprints follow the "Request Permissions" link in the navigation box
To subscribe to BMJ go to:
http://resources.bmj.com/bmj/subscribers
RATING QUALITY OF EVIDENCE AND STRENGTH OF RECOMMENDATIONS

GRADE: going from evidence to recommendations

The GRADE system classifies recommendations made in guidelines as either strong or weak. This article explores the meaning of these descriptions and their implications for patients, clinicians, and policy makers.

This is the third of a series of five articles describing the GRADE approach to developing and presenting recommendations for management of patients. In it, we deal with how GRADE suggests clinicians should interpret the strength of a recommendation.

What do we mean by the strength of a recommendation?
The strength of a recommendation reflects the extent to which we can be confident that the desirable effects of an intervention outweigh the undesirable effects. Desirable effects of an intervention include reduction in morbidity and mortality, improvement in quality of life, reduction in the burden of treatment (such as having to take drugs or the inconvenience of blood tests), and reduced resource expenditures. Undesirable consequences include adverse effects that have a deleterious impact on morbidity, mortality, or quality of life or increase use of resources.

Previous grading systems have used up to nine categories of strength of recommendations. The GRADE system has only two categories—although in this article we will characterise them as strong and weak, guideline panels may choose different words to characterise the two categories of strength. When using GRADE, panels make strong recommendations when they are confident that the desirable effects of adherence to a recommendation outweigh the undesirable effects. Weak recommendations indicate that the desirable effects of adherence to a recommendation probably outweigh the undesirable effects, but the panel is less confident.

Strong and weak recommendations provide specific guidance
GRADE's binary classification of strength of recommendations provides clear direction to patients, clinicians, and policy makers. The implications of a strong recommendation are:

• For patients—most people in your situation would want the recommended course of action and only a small proportion would not; request discussion if the intervention is not offered
• For clinicians—most patients should receive the recommended course of action
• For policy makers—the recommendation can be adopted as a policy in most situations.

The implications of a weak recommendation are:

• For patients—most people in your situation would want the recommended course of action, but many would not
• For clinicians—you should recognise that different choices will be appropriate for different patients and that you must help each patient to arrive at a management decision consistent with her or his values and preferences
• For policy makers—policy making will require substantial debate and involvement of many stakeholders.

As clinicians become more aware of variability in patients’ values and preferences, they are turning to structured decision aids to facilitate the decision-making process. A strong recommendation indicates that use of a decision aid is unnecessary—almost all informed patients will make the same choice. A weak recommendation indicates that a decision aid could be useful.

Managers of healthcare systems are becoming increasingly interested in ensuring the quality of care. Guidelines help managers to differentiate practices that constitute quality of care from others that are discretionary. GRADE provides clear guidance on these matters. The management options associated with strong, but not with weak, recommendations are candidates for quality criteria. When a recommendation is weak, discussing with patients and families the relative merits of the alternative management strategies may become a quality criterion.

Four key factors determine the strength of a recommendation
The first key determinant of the strength of a recommendation is the balance between the desirable and undesirable consequences of the alternative management strategies, on the basis of the best estimates of those consequences (table 1). Consider, for instance, the use of antenatal steroids in women destined to deliver an infant prematurely. Administration of steroids to mothers decreases the risk of infant respiratory distress syndrome with minimal side effects, inconvenience, and costs. Advantages of steroid administration hugely outweigh the disadvantages, indicating the appropriateness of a strong recommendation.

When advantages and disadvantages are closely
The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted. The more values and preferences vary, or the greater the uncertainty in values and preferences. Given that alternative management strategies will always have advantages and disadvantages, and thus a trade-off exists, how a guideline panel values benefits, risks, and inconvenience is critical to the strength of any recommendation.

Consider the subject of preventing strokes in patients with atrial fibrillation. Warfarin, relative to no antithrombotic therapy, reduces the risk of stroke by approximately 65% but increases the risk of severe gastrointestinal bleeding. Devereaux and colleagues asked 63 physicians and 61 patients how many serious gastrointestinal bleeds they would tolerate in 100 patients and still be willing to prescribe or take warfarin to prevent eight strokes (four minor and four major) in 100 patients. Figure 1 shows the results. Whereas physicians gave a wide diversity of responses, most patients placed a high value on avoiding a stroke and were ready to accept a bleeding risk of 22% to reduce their chances of having a stroke by 8%. Even among patients, however, diversity in values and preferences was apparent; a few patients were ready to accept only a small risk of bleeding. These data suggest that only in patients at high risk of stroke would a strong recommendation for warfarin be warranted.

Contrast this with the decision faced by pregnant women with deep venous thrombosis. Warfarin treatment between the sixth and 12th week of pregnancy puts women’s unborn infants at risk of relatively minor developmental abnormalities. The alternative, heparin, eliminates the risk to the child. The benefit, however, comes with disadvantages of pain, inconvenience, and cost. Clinicians’ experience is that women overwhelmingly place a high value on preventing fetal complications. Thus, despite its disadvantages, a strong recommendation for heparin substitution is warranted.

The final determinant of the strength of a recommendation is cost. Cost is much more variable over time and geographical areas than are other outcomes. Drug costs tend to plummet when patents expire, and charges for the same drug differ widely across jurisdictions. In addition, the resource implications vary widely. For instance, a year’s prescription of the same expensive drug may pay for a single nurse’s salary in the United States and 30 nurses’ salaries in China.

Thus, although higher costs reduce the likelihood of a strong recommendation in favour of an intervention, the context of the recommendation will be critical. In considering resource allocation, guideline panels must therefore be specific about the setting to which a recommendation applies.

**Strong recommendations may not be important from all perspectives**

If the consequences of the choice are relatively unimportant, some patients may not bother with even strong recommendations. This is particularly likely if they are faced with many new drugs or many suggestions to change their lifestyle.

When setting priorities, governments and public health officials must also consider factors beyond the strength of a recommendation. These include the prevalence of the health problem, considerations of equity, and the potential for improvement in quality of care, all of which will have an impact on the population health gain of an intervention.

**Determinants of strength of recommendation**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance between desirable and undesirable effects</td>
<td>The larger the difference between the desirable and undesirable effects, the higher the likelihood that a strong recommendation is warranted. The narrower the gradient, the higher the likelihood that a weak recommendation is warranted.</td>
</tr>
<tr>
<td>Quality of evidence</td>
<td>The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted.</td>
</tr>
<tr>
<td>Values and preferences</td>
<td>The more values and preferences vary, or the greater the uncertainty in values and preferences, the higher the likelihood that a weak recommendation is warranted.</td>
</tr>
<tr>
<td>Costs (resource allocation)</td>
<td>The higher the costs of an intervention—that is, the greater the resources consumed—the lower the likelihood that a strong recommendation is warranted.</td>
</tr>
</tbody>
</table>
SUMMARY POINTS

The strength of a recommendation reflects the extent to which we can be confident that desirable effects of an intervention outweigh undesirable effects.

GRADE classifies recommendations as strong or weak.

Strong recommendations mean that most informed patients would choose the recommended management and that clinicians can structure their interactions with patients accordingly.

Weak recommendations mean that patients’ choices will vary according to their values and preferences, and clinicians must ensure that patients’ care is in keeping with their values and preferences.

Strength of recommendation is determined by the balance between desirable and undesirable consequences of alternative management strategies, quality of evidence, variability in values and preferences, and resource use.

Endpiece

Two kinds of surgery

Surgical operations are of two kinds—those that benefit the patient and those that kill him.

Abu al-Qasim Khalaf bin ‘Abbas el-Zahrawi, also known as Albucasis (940-1013)

Submitted by Munier Hossain, staff grade orthopaedic surgeon, Ysbyty Gwynedd, Bangor