

Reviewing Evidence Using GRADE

Why GRADE?

The GRADE (Grading of Recommendations Assessment, Development and Evaluation) system has advantages over other systems including the current system used by ICSI. Advantages include:

- Developed by a widely representative group of international guideline developers
- Explicit and comprehensive criteria for downgrading and upgrading quality of evidence ratings
- Clear separation between quality of evidence and strength of recommendations that includes a transparent process of moving from evidence evaluation to recommendations
- Clear, pragmatic interpretations of strong versus weak recommendations for clinicians, patients and policy-makers
- Explicit acknowledgement of values and preferences and
- Explicit evaluation of the importance of outcomes of alternative management strategies.

GRADE is becoming widely accepted in the guideline writing community. Currently over 60 organizations have adopted the GRADE methodology.

How is Literature Searched?

In the GRADE process, evidence is gathered related to a specific topic or PICO (population, intervention, comparison, outcome) question. Systematic reviews are utilized first. Further literature is incorporated including randomized control trials, observational studies etc. The evidence addresses the same population, intervention, comparisons and outcomes.

How is Quality of Evidence Handled?

In the GRADE methodology, systematic reviews provide the strongest evidence as they summarize one or more well-designed and well-executed randomized control trials (RCTs) yielding consistent directly applicable results. Additional strong evidence can also come, under unusual circumstances, from observational studies yielding very large effects. The moderate strength category is populated by randomized trials with important limitations and by exceptionally strong observational studies. Observational studies, and on occasion RCTs with multiple serious limitations, will fill the low quality evidence category. This categorization follows the principle that all relevant clinical studies provide evidence, the strength of which varies.

The quality of each piece of evidence within the topic is first classified as high, moderate or low according to factors that include the study design, the consistency of the results, and the directness of the evidence. These pieces of evidence then become the “body” of evidence around that topic. This “body” of evidence is then given an overall quality rating.

How is Quality Defined?

Category	Quality Definitions
High Quality Evidence	Further research is very unlikely to change our confidence in the estimate of effect
Moderate Quality Evidence	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
Low Quality Evidence	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate or any estimate of effect is very uncertain.

Note: Transition steps incorporating GRADE methodology include the following:

- All existing Class A (RCT's) studies have been considered as High quality evidence unless specified differently by a work group member.
 - All existing Class B, C and D studies have been considered as Low quality evidence unless specified differently by a work group member.
 - All existing Class M and R studies are identified by study design OR giving a low quality rating.
- However, systematic Reviews are assessed using GRADE methodology.
- All new literature considered by the work group for this revision has been assessed using GRADE methodology.

What To Do When Quality Of Evidence Differs Across Outcomes?

ICSI will provide a single rating of quality of evidence for every recommendation. Recommendations, however, depend on evidence regarding a variety of outcomes. Thus, it may occasionally be necessary to report a single evidence grade when the quality of evidence differs across important outcomes. Benefit outcomes, for instance decreased mortality, are often supported by high quality evidence whereas only moderate or low quality evidence supports harm outcomes. In such instances, we suggest that editors should consider whether toxicity endpoints are crucial to the decision regarding the optimal management strategy. If they are, one must rate the overall quality of the evidence (the “body”) according to the studies that address toxicity. If not, the overall rating of the evidence is based on the evidence regarding benefit.

What is the Difference in the Strength of Recommendations?

Recommendation language should include who should do what to whom under which circumstances. They should be supported by the quality of evidence. Within the annotations, the following should occur:

- Discussion of benefits and harms
- Values used by the work group in making recommendation
- Discussion of difference of opinion, if any, within in the work group or why differing from other guidelines (if applicable).

How is Strength Defined?

GRADE classifies recommendations as strong or weak. 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.

Factor	Comment
Costs (resource allocation)	The higher the costs of an intervention—that is, the greater the resources consumed—the lower the likelihood that a strong recommendation is warranted
Balance between desirable and undesirable effects	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
Quality of evidence	The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted
Values and preferences	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
Costs (resource allocation)	The higher the costs of an intervention—that is, the greater the resources consumed—the lower the likelihood that a strong recommendation is

What Are Other Ways To Think About Strength?

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

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.

How Do Strength and Quality Correlate?

Category	Quality Definitions	Strong Recommendation	Weak Recommendation
High Quality Evidence	Further research is very unlikely to change our confidence in the estimate of effect	The work group is confident that the desirable effects of adhering to this recommendation outweigh the undesirable effects. This is a strong recommendation for or against. This applies to most patients.	The work group recognizes that the evidence, though of high quality, shows a balance between estimates of harms and benefits. The best action will depend on local circumstances, patient values or preferences.
Moderate Quality Evidence	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate	The work group is confident that the benefits outweigh the risks, but recognizes that the evidence has limitations. Further evidence may impact this recommendation. This is a recommendation that likely applies to most patients.	The work group recognizes that there is a balance between harms and benefit, or that there is uncertainty about the estimates of the benefits and harms of the proposed intervention that may be affected by new evidence. Alternative approaches will likely be better for some patients under some circumstances.
Low Quality Evidence	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate or any estimate of effect is very uncertain.	The work group feels that the evidence consistently indicates the benefit of this action outweighs the harms. This recommendation might change when higher quality evidence becomes available.	The work group recognizes that there is significant uncertainty about the best estimates of benefits and harms. Very weak recommendation, other alternatives may be equally reasonable.