

Ranking the Evidence

Choosing therapies to treat patients with severe sepsis and septic shock requires an organized approach to evaluating the evidence. The Sepsis Resuscitation and Management Bundles were derived from the [2008 Surviving Sepsis Campaign Guidelines](#) which incorporated the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) system approach to evaluating the literature.

The GRADE approach is a structured system for rating quality of evidence and grading strength of recommendation in clinical practice (1-4). The SSC Steering Committee and individual authors collaborated with GRADE representatives to apply the GRADE system to the SSC guidelines revision process. The representatives of GRADE system were directly involved in all discussions and deliberations amongst the guidelines committee members as to grading decisions.

The GRADE system is based on a sequential assessment of the quality of evidence, followed by assessment of the balance between benefits versus risks, burden, and cost and, based on the above, development and grading of a management recommendations (2-4). Keeping the rating of quality of evidence and strength of recommendation explicitly separate constitutes a crucial and defining feature of the GRADE approach. This system classifies quality of evidence as high (Grade A), moderate (Grade B), low (Grade C), or very low (Grade D). Randomized trials begin as high quality evidence, but may be downgraded due to limitations in implementation, inconsistency or imprecision of the results, indirectness of the evidence, and possible reporting bias (see table 1). Examples of indirectness of the evidence include: population studied, interventions used, outcomes measured, and how these relate to the question of interest. Observational (non-randomized) studies begin as low-quality evidence, but the quality level may be upgraded on the basis of large magnitude of effect. An example of this is the quality of evidence for early administration of antibiotics.

Table 1. Determination of the Quality of Evidence

- **Underlying methodology**

- A RCT
- B Downgraded RCT or upgraded observational or cohort studies
- C Well-done observational or cohort studies with controls
- D Case series or expert opinion

- **Factors that may decrease the strength of evidence**

1. Poor quality of planning and implementation of available RCTs suggesting high likelihood of bias.
2. Inconsistency of results (including problems with subgroup analyses).
3. Indirectness of evidence (differing population, intervention, control, outcomes, comparison).
4. Imprecision of results.
5. High likelihood of reporting bias.

- **Main factors that may increase the strength of evidence**

1. Large magnitude of effect [direct evidence, relative risk (RR) >2 with no plausible confounders]
2. Very large magnitude of effect with RR >5 and no threats to validity (by two levels)
3. Dose response gradient

*RCT= randomized controlled trial; RR=relative risk

Table 2. Factors Determining Strong vs. Weak Recommendation

| What should be considered | Recommended Process |
|--|--|
| Quality of evidence | The lower the quality of evidence the less likely a strong recommendation |
| Relative importance of the outcomes | If values and preferences vary widely, a strong recommendation becomes less likely |
| Baseline risks of outcomes | The higher the risk, the greater the magnitude of benefit |
| Magnitude of relative risk including benefits, harms, and burden | Larger relative risk reductions or larger increases in relative risk of harm make a strong recommendation more or less likely respectively |
| Absolute magnitude of the effect | The larger the absolute benefits and harms, the greater or lesser likelihood respectively of a strong recommendation |
| Precision of the estimates of the effects | The greater the precision the more likely is a strong recommendation |
| Costs | The higher the cost of treatment, the less likely a strong recommendation |

The GRADE system classifies recommendations as strong (Grade 1) or weak (Grade 2). The grade of strong or weak is considered of greater clinical importance than a difference in letter level of quality of evidence. The committee assessed whether the desirable effects of adherence will outweigh the undesirable effects, and the strength of a recommendation reflects the group's degree of confidence in that assessment. A strong recommendation in favor of an intervention reflects that the desirable effects of adherence to a recommendation (beneficial health outcomes, less burden on staff and patients, and cost savings) will clearly outweigh the undesirable effects (harms, more burden and greater costs). A weak recommendation in favor of an intervention indicates that the desirable effects of adherence to a recommendation probably will outweigh the undesirable effects, but the panel is not confident about these tradeoffs-- either because some of the evidence is low-quality (and thus there remains uncertainty regarding the benefits and risks) or the benefits and downsides are closely balanced. While the degree of confidence is a continuum and there is a lack of a precise threshold between a strong and a weak recommendation, the presence of important concerns about one or more of the above factors makes a weak recommendation more likely. In the actual guidelines paper, a "strong" recommendation is worded as "we recommend" and a weak recommendation as "we suggest."

The implications of calling a recommendation "strong" are that most well-informed patients would accept that intervention, and that most clinicians should use it in most situations. There may be circumstances in which a "strong" recommendation cannot or should not be followed for an individual patient because of that patient's preferences or clinical characteristics which make the recommendation less applicable. It should be noted that

being a "strong" recommendation does not automatically imply standard of care. For example, the strong recommendation for administering antibiotics within one hour of the diagnosis of severe sepsis, although desirable, is not necessarily the actual observed standard of care seen in current practice. The implication of a "weak" recommendation is that although a majority of well-informed patients would accept it (but a substantial proportion would not), clinicians should consider its use according to particular circumstance.

1. Guyatt G, Schünemann H, Cook D, Jaeschke R, Pauker S. Applying the grades of recommendations for antithrombotic and thrombolytic therapy: The seventh ACCP conference of antithrombotic and thrombolytic therapy. *Chest* 2004; 126:179S-187S

2. GRADE working group. Grading quality of evidence and strength of recommendations. *BMJ* 2004; 328:1490-1498

3. Guyatt G, Gutterman D, Baumann MH, Addrizzo-Harris D, Hylek EM, Phillips B. Grading strength of recommendations and quality of evidence in clinical guidelines: report from an American College of Chest Physicians task force. *Chest* 2006; 129(1):174-181

4. An Official ATS Statement: Grading the Quality of Evidence and Strength of Recommendations in ATS Guidelines and Recommendations. Holger J. Schünemann, Roman Jaeschke, Deborah J. Cook, et al on behalf of the ATS Documents Development and Implementation Committee. *Am J Respir Crit Care Med* Vol 174. pp 605–614, 2006