VICE raises the GRADE: recommendations for CAEP clinical practice guidelines

Suneel Upadhye1 · Sameer Sharif1,2 · Eddy Lang3 · Bram Rochwerg2,4

Received: 6 March 2021 / Accepted: 29 March 2021 / Published online: 20 April 2021 © The Author(s), under exclusive licence to Canadian Association of Emergency Physicians (CAEP)/ Association Canadienne de Médecine d’Urgence (ACMU) 2021

Keywords GRADE · Guidelines

Clinical scenario

You are treating a patient with fluid-refractory septic shock in your Emergency Department (ED). You recall from the 2008 Canadian Association of Emergency Medicine (CAEP) Critical Care Committee (C4) Sepsis guidelines that a vaso-pressor is the next intervention to support blood pressure. You also recollect the 2015 CAEP C4 Vasopressor and Inotrope use in Canadian ED (VICE) guidelines, which were produced using the GRADE (Grades of Recommendations, Assessment, Development and Evaluation) approach. Panelists made a “Strong” recommendation for norepinephrine as the first-line vasopressor in septic patients with refractory shock. Being unfamiliar with the GRADE framework, you are uncertain how to apply these recommendations in clinical practice.

Introduction

Clinical practice guidelines (CPGs) are “systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances” [1]. The benefits, limitations, and usefulness of CPGs have been debated in the medical literature [1]. CPGs are of greatest utility when they meet the highest standards for “trustworthiness”, as outlined in the seminal Institute of Medicine (IOM) standards [2]. The quality of Emergency Medicine (EM) CPGs has been examined, however, and found to be wanting [3].

One area of contention has been the multitude of different development frameworks that have been used amongst different author groups, leading to variable presentations of evidence and recommendations in CPGs [4]. Such variability can create confusion for readers not familiar with development methods, and limit trust and applicability in clinical decision-making. Furthermore, other limitations that impact CPG trustworthiness include incomplete stakeholder representation on CPG panels (especially patient/caregivers), panelist conflicts of interest, and non-transparent evaluation of evidence and other inputs in generating recommendations (resource needs, equity, applicability, etc.).

To address such limitations, key guidance from the IOM suggests trustworthy CPGs must clearly present the evidence quality, the strength of recommendations and clearly articulate these linked components as transparently as possible [2]. The specific tenets of the 2011 IOM standards are as follows: (1) CPGs must be based on a systematic evidence reviews; (2) be developed by a knowledgeable, multidisciplinary panel of experts and key stakeholders (including patients/caregivers); (3) prioritize patient values and preferences; (4) explicitly evaluate and manage various biases and conflicts of interest; (5) clearly explain the linkages between evidence quality and recommendations strength (for all care options and health outcomes); (6) perform scheduled updates, or when important new evidence arises [2]. These tenets are especially important in the current climate rife with misinformation stemming from the COVID-19 outbreak [5], and

Sameer Sharif
sameer.sharif@medportal.ca
1 Department of Medicine, Division of Emergency Medicine, McMaster University, Hamilton, ON, Canada
2 Department of Medicine, Division of Critical Care, McMaster University, Hamilton, ON, Canada
3 Department of Emergency Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
4 Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, ON, Canada
the variable quality of CPG’s produced early in the pandemic should not to be understated [5].

To increase CPG trustworthiness, the GRADE initiative has spent over a decade improving an intuitive framework for creating transparent evidence-based recommendations for clinician end-users. Recent research suggests that clinicians (community, academic) prefer GRADE-based recommendations in clinical decision-making, and want evidence summaries linked to CPG recommendations [6]. This editorial will briefly outline the basic elements of the GRADE framework, specifically designed for practicing EM physicians, and how GRADE guidelines should be used to inform EM practice.

How does GRADE lead to more trustworthy CPGs?

Key steps in using a trustworthy guideline to provide direction in practice are understanding both the presentation of the evidence supporting the recommendations, and the strength and implications of the recommendations themselves [7]. Guideline development involves several crucial steps: recruiting representative/multidisciplinary panels, ensuring patient engagement, managing conflict of interest, developing questions of interest using PICO (patient, intervention, comparator, outcome) format, prioritizing outcomes, performing systematic reviews of the relevant evidence, evaluating the certainty of evidence, generating recommendations after considering all factors, and finally presentation of clear guidance.

The certainty of evidence should incorporate the quality, quantity, and consistency of evidence. The GRADE approach is one of the highest scoring instruments when evaluated by the US Agency for Healthcare Research and Quality [8]; it emphasizes the link between the certainty of evidence and the recommendation, but recognizes that other factors beyond the quality of evidence contribute to the strength of a recommendation [8].

As opposed to a more complicated system of letters and numbers, the GRADE approach attempts to be clearer and more concise, summarized in the Box [6]. GRADE defines quality of evidence as confidence or certainty in estimates of prognosis, association, or intervention effects, etc., arising from a body of evidence, and rates certainty as “high, moderate, low or very low”, as outlined in the Box. In this rating system, randomized controlled trials start as high certainty (but may be rated down based on issues in any one of five domains) and observational studies start as low certainty, but may be rated down by the same five domains (balance between benefits and harms, costs/resources, patients values and preferences, feasibility, acceptability, and impact on health equity), or up if there are aspects of the literature that might increase certainty (for example dose response or a large magnitude of effect) (Box 3) [9]. For instance, an evidence base in which all the randomized controlled trials show a consistent effect is more believable than if some show benefit and some show harm; the latter circumstance might be rated down for inconsistency (akin to statistical heterogeneity). Subsequently, the costs, resources, values and preferences, feasibility, acceptability, balance of benefits and harms, and certainty of evidence are incorporated into generating recommendations, using the Evidence-to-Decisions decision process (facilitated by use of free online GradePRO or MAGIC software; available at https://gradePRO.org and https://app.magicapp.org). Recommendations are characterized as strong or weak (also synonymous with conditional) (Box). Given this increased clarity of presentation and increased trustworthiness, prehospital emergency guideline author groups have embraced the use of GRADE methods [10]. The recently created Society of Academic Emergency Medicine GRACE (Guidelines for Reasonable and Appropriate Care in ER) groups have also begun using GRADE methodology to create new CPG products (S. Upadhye, personal communication, manuscript in preparation).

How should clinicians respond to different GRADE recommendations? Strong recommendations recognize that almost all well-informed patients would choose to follow the clinical guidance. Strong recommendations are used to inform standards of care and quality markers. Although strong recommendations would apply to almost all individuals, no recommendation is meant to be absolutely dogmatic and there will always be exceptions. A weak recommendation recognizes that there will be more variability in practice, and despite the fact that the majority of well-informed patients would choose to follow the clinical guidance, there may still be a large minority that would not (Box). Weak recommendations require more nuanced shared decision-making between patients and clinicians. For physicians, they should recognize that different choices will be appropriate for different patients; patients must be helped to reach a management decision consistent with their values and preferences [6].

The GRADE system also requires transparent summary presentation of the evidence that was used to inform the recommendations [9]. A summary of findings table lists the key outcome measures for interventions as prioritized by panel members and patients, corresponding relative and absolute effect estimates, GRADE certainty of evidence rating, and plain language summaries (Box). Of note, although both relative and absolute effect estimates are presented, GRADE suggests that absolute estimates and associated 95% confidence intervals should be the focus when making recommendations. The summary of findings table is aimed at a broad audience looking for a quick summary snapshot of key information needed to make decisions. An evidence profile
is an expanded version of the summary of findings and includes a detailed assessment of key factors determining the certainty of evidence, useful for review authors, guideline development panels and stakeholders who are looking to better understand the evidence that was used to generate the recommendations (Box).

Box. Key features of GRADE framework

Advantages of GRADE compared to other ranking schema

1. Clearly separated evidence quality assessments and recommendation strengths.
2. Focus on patient-important rather than surrogate outcomes.
3. Explicit detailed published criteria for rating evidence quality up or down.
4. Transparent process of moving from evidence to recommendations with explicit published guidance on factors to consider.
5. Clear acknowledgement of values and preferences.
6. Clear pragmatic interpretation of “strong” vs “weak” recommendations for various end-users (clinicians, patients, policy-makers).
7. Useful in systematic reviews, CPGs, and health technology assessments.

Quality of evidence (confidence or certainty in effects estimates) and definitions

High quality = very confident that the true effect lies close to that of the estimate of the effect.
Moderate quality = moderately confident in the effect estimate (true effect likely to be close to effect estimate, but possible that it is substantially different).
Low quality = limited confidence in effect estimate (true effect may be substantially different from effect estimate).
Very low quality = very little confidence in effect estimate (true effect likely to be substantially different from effect estimate).

Factors influencing quality of evidence evaluations

Decreased quality of evidence: risk of bias, inconsistent results, indirectness of evidence, imprecision of effect estimates, publication bias.
Increased quality of evidence: larger magnitude of effect, dose–response gradients, plausible confounding (which may reduce a demonstrated effect).

Implications of strong vs weak recommendations for end-users [ED example]

Strong recommendations for intervention Patient perspective = all or almost all patients in this situation would want the recommended intervention (few would reject); [ED: inhaled steroids in asthma exacerbations will help prevent relapses].
Clinician perspective = most patients should be encouraged to use this intervention, time consuming shared decision-making unnecessary [prescribe inhaled steroids for asthma patients discharged from ED].
Quality assurance perspective = recommendation adherence could be used as a quality assurance performance indicator; failure to adhere should require documentation to follow rationale [ED chart review of steroid scripts for discharged asthmatics].

Weak recommendations for interventions Patients = most fully informed patients would follow suggested intervention, but many would not [antibiotics not needed for acute otitis media in ED for first 48 h].
Clinicians = in most instances requires shared decision-making, including communication of consequences of alternative management strategies to the patient may require self-education of evidence [avoiding antibiotics for otitis media in ED].
Quality assurance = ED charting of clinician/patient discussions in such scenarios could be used as a quality assurance measure.

Future directions for CAEP CPG author groups

The evolving landscape of CPG methodology presents new opportunities for CAEP Interest Groups. The VICE guidelines, produced by the CAEP C4 group, represent the first use of GRADE for a CAEP-produced CPG, and a welcome addition to the CAEP CPG library. An increasing variety of training opportunities should make it easier for CAEP author groups to acquire GRADE methods skills. These include attending training workshops, the online INGUIDE training/certification (available at https://inguide.org), and using the Guidelines International Network [GIN]-McMaster CPG checklists (multiple languages) available at https://cebr grade.mcmaster.ca/guidecheck.html). Finally, there is parallel need to develop specific training around patient engagement in CPG development panels, prioritizing CPG topic questions in alignment with ED patient values and preferences,
and dissemination strategies linked to meaningful quality improvement of care in the ED setting. All such activities can be readily integrated in GRADE-based CPG product development.

**Resolution of clinical scenario**

With increased understanding of GRADE, you gain confidence that applying a “strong” recommendation in favour of using norepinephrine as the first-line vasopressor will help your patient in this scenario. A norepinephrine infusion is initiated and the patient’s blood pressure starts to normalize.

**Conclusion**

Trustworthy guidelines require rigorous transparent development in offering clear recommendations to end-users. The GRADE approach offers explicit processes to evaluate evidence, present summary findings, and create meaningful recommendations for different consumers—all of which can inform shared clinical decision-making (Fig. 1). Emergency medicine guideline author groups should be trained in using GRADE tools, and create guideline products addressing important patient-oriented questions in order to improve ED clinical care.

**Declarations**

**Conflict of interest** None.
References

1. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. BMJ. 1999;318(7182):527–30.

2. Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice G. (2011) In: Graham R, Manchester M, Miller Wolman D, Greenfield S, Steinberg E, editors. Clinical Practice Guidelines We Can Trust, Washington (DC): National Academies Press (US) Copyright 2011 by the National Academy of Sciences. All rights reserved.

3. Zupon A, Rothenberg C, Couturier K, Tan TX, Siddiqui G, James M, et al. An appraisal of emergency medicine clinical practice guidelines: Do we agree? Int J Clin Pract. 2019;73(2):e13289.

4. Upshur REG. Are all evidence-based practices alike? Problems in the ranking of evidence. Can Med Assoc J. 2003;169(7):672.

5. Dagens A, Sigfrid L, Cai E, Lipworth S, Cheng V, Harris E, et al. Scope, quality, and inclusivity of clinical guidelines produced early in the covid-19 pandemic: rapid review. BMJ. 2020;369:m1936.

6. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924–6.

7. Harrison MB, Légaré F, Graham ID, Fervers B. Adapting clinical practice guidelines to local context and assessing barriers to their use. Can Med Assoc J. 2010;182(2):E78.

8. Shekelle P, Woolf S, Grimshaw JM, Schünemann HJ, Eccles MP. Developing clinical practice guidelines: reviewing, reporting, and publishing guidelines; updating guidelines; and the emerging issues of enhancing guideline implementability and accounting for comorbid conditions in guideline development. Implement Sci. 2012;7(1):62.

9. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ. What is “quality of evidence” and why is it important to clinicians? BMJ. 2008;336(7651):995–8.

10. Brown KM, Macias CG, Dayan PS, Shah MI, Weik TS, Wright JL, et al. The development of evidence-based prehospital guidelines using a GRADE-based methodology. Prehosp Emerg Care. 2014;18(Suppl 1):3–14.