Journal of Clinical Epidemiology 68 (2015) 597–600

Journal of Clinical Epidemiology

COMMENTARY

Guideline panels should not GRADE good practice statements

Gordon H. Guyatt[a,b,*], Holger J. Schünemann[a,b], Benjamin Djulbegovic[c], Elie A. Akl[a,d]

[a]*Department of Clinical Epidemiology and Biostatistics, McMaster University, 1200 Main St. West, Hamilton, Canada L8S 4K1*
[b]*Department of Medicine, McMaster University, 1200 Main St. West, Hamilton, Canada L8S 4K1*
[c]*Department of Internal Medicine, Division of Evidence-based Medicine and Health Outcomes Research, University of South Florida, 4202 East Fowler Avenue, Tampa, FL 33612, USA*
[d]*Department of Internal Medicine, American University of Beirut, Cario Street, Riad-El-Solh Beirut, Beirut 1107 2020, Lebanon*

Accepted 14 December 2014; Published online 31 December 2014

In the first article in JCE's series presenting guidance for the application of grades of recommendation, assessment, development and evaluation (GRADE) methodology [1], we identified a number of limitations associated with the GRADE approach. One of these limitations related to a category of recommendations that guideline panels may feel are important but that are not appropriate for rating the certainty of the evidence (synonyms: confidence in estimates, quality of the evidence). Because, for such recommendations, a formal rating of certainty is inappropriate, they fall outside the domain of the standard GRADE process.

That article did not place the description of this category of recommendations in a prominent place. Perhaps as a consequence, our informal experiences with guideline panels, and two formal assessment, suggest that most guideline panels applying GRADE are unaware of good practice statements. The purpose of the present editorial to clarify the issue and to provide a more prominent exposition that will increase awareness and appropriate use.

In the original article, we described what we called "an ill defined set of recommendations" labeled as "motherhood statements" or "good practice recommendations"—here, we will refer to them as "good practice statements." Perhaps the best way to understand the sort of statement to which we are referring is to consider a number of examples: please look now at the Box 1 that presents recommendations that would optimally be characterized as good practice statements.

In our initial discussion of such recommendations, we struggled how guideline panels could best recognize these situations when it may be inadvisable to apply formal GRADE methodology. We suggested that it was obvious that such recommendations would do substantially more good than harm (or vice versa) and that therefore no one would consider doing a study to definitively establish the answer to the implicit question.

We made an additional suggestion that we now believe is the best way to recognize recommendations that should not be graded but characterized as good practice statement. Before presenting that suggestion, we will consider how guideline panels have typically dealt with good practice statements. Panels using GRADE to address these issues offer strong recommendations with the evidence classified as warranting low or very low certainty (low confidence or low quality evidence). Such recommendations are not uncommon: indeed, in a systematic examination of Endocrine Society recommendations, of 121 strong recommendations based on low or very low certainty evidence (discordant recommendations), investigators classified 43 (36%) as good practice statements [2]. Furthermore, in a similar examination of World Health Organization recommendations, of 160 discordant recommendations, 29 (18%) were classified as good practice statements [1,3].

Is it true that the evidence supporting all these statements warrants low or very low certainty? Clearly, it is not. If one asked panellists recommending these clinical behaviors if they are confident that the behaviors will result in more desirable than undesirable consequences, they would invariably answer in the affirmative.

Their response (ie, implicitly expressing moderate or high certainty in estimates of effect), in the face of formally classifying evidence as low or very low quality, is clearly contradictory and highlights a common misunderstanding of GRADE methodology. In the absence of randomized trials—indeed, in the absence of any formal studies addressing the question of interest—guideline panels believe that they should classify evidence as low or very low quality. In doing so, they have not grasped GRADE's definition of quality of evidence as confidence in estimates of effect.

Disclaimer: This article represents the views of the authors and does not constitute official GRADE working group guidance.

Conflict of interest: None of the authors has any financial conflict of interest to disclose. All authors are members of the GRADE working group.

* Corresponding author. Tel.: 905-525-9140x22900; fax: 905-524-3841.
E-mail address: guyatt@mcmaster.ca (G.H. Guyatt).

http://dx.doi.org/10.1016/j.jclinepi.2014.12.011

Box 1 Examples of good practice statement previously mistakenly presented as GRADEd recommendations

For patients with congenital adrenal hyperplasia, we recommend monitoring patients for signs of glucocorticoid excess [5].

Triage (ie, take different courses of action for low vs. higher pretest probability) people with tuberculosis symptoms [6].

Health services should be made available, accessible, and acceptable to sex workers based on the principles of avoidance of stigma, nondiscrimination, and the right to health [7].

In patients presenting with heart failure, initial assessment should be made of the patient's ability to perform routine/desired activities of daily living [8].

Guideline panellists considering good practice statements have failed to make the connection that their high level of certainty in net benefits would mandate a corresponding rating of high quality, and they are therefore mistaken in classifying the evidence as low or very low quality. Good practice statements typically represent situations in which a large body of indirect evidence, made up of linked evidence including several indirect comparisons, strongly supports the net benefit of the recommended action.

Although indirectness often results in diminished certainty in effects, this is not always the case. An amusing example repeatedly used is the difference in outcome when one does or does not use a parachute when jumping from an aircraft. Panels should consider making good practice statements when, without a formal literature search, they are confident that indirect evidence is at or near this level of certainty in the net benefit of the intervention. Furthermore, panels might reasonably consider making good practice statements when it would be an onerous and unproductive exercise to collect the indirect linked evidence supporting the recommendations.

Why are we confident that it is wise to monitor patients with congenital adrenal hyperplasia for glucocorticoid excess? The reason is that relevant symptoms and signs appear not infrequently, that patients will suffer if clinicians fail to recognize these signs, and that clinical action can ameliorate the problem.

It would be possible to accumulate and summarize the relevant evidence. There have been no randomized trials or observational studies that have directly compared monitoring to no monitoring of glucocorticoid excess in patients with congenital adrenal hyperplasia—thus, we have no direct evidence. The panel could, nevertheless, build a case for the benefits through indirect evidence. They could collect all the reports of the adverse consequences of glucocorticoid excess. They could then collect the evidence that supports the usefulness of the relevant symptoms, signs, and laboratory tests in the diagnosis of glucocorticoid excess. Then, they could collect and summarize the evidence of the benefits of the candidate management strategies. Finally, they could describe how they link these three bodies of evidence to make the case for their high level of certainty regarding the net benefits of monitoring for glucocorticoid excess. The case for the good practice statement is the poor use of time in collecting and summarizing the relevant evidence.

To turn to another of our examples, why are we confident that it is wise to triage every patient with symptoms that might even remotely suggest possible tuberculosis? By triage, the guideline developers mean isolation and investigation of patients with suspected tuberculosis for only those patients with a sufficiently high pretest probability. The reason we are confident in the advisability of triage is that failure to do so—that is, fully investigating every individual with symptoms even remotely suggestive of tuberculosis rather than restricting investigation to those with a higher pretest probability—will lead to over investigation and wasteful use of scarce health resources.

Why the confidence in providing appropriate health services to sex workers? First, confidence is based on an underlying value we place in equitable access to health care. Second, because a large number of health care interventions do more good than harm sex workers will therefore have better health if they have access to services.

In each case, although there is a great deal of evidence supporting the recommended behaviors, teasing out the nature of this evidence would be challenging and a waste of time and energy. Given that time and energy is typically at a high premium in the guideline development exercise, their expenditure in turning good practice statements into GRADEd recommendations (strong recommendations based on high or moderate not low or very low certainty) is likely to be inadvisable.

1. Reservations regarding good practice statements

A word of caution is required: good practice statements may be subject to abuse. They potentially allow guideline panels to issue strong recommendations that may be unwarranted (which guideline panels seem prone to do [2,3]) and to do so without the intellectual work that formally applying the GRADE process demands. Furthermore, judgments about what are incontestable net benefits are inevitably subjective. Thus, good practice statements represent a temptation, and panels should therefore use them sparingly.

We would suggest that guideline panels explicitly address the following issues before they make good

Table 1. Questions guideline panels considering good practice statement should ask themselves

i)	Is the statement clear and actionable?
ii)	Is the message really necessary?
iii)	Is the net benefit large and unequivocal?
iv)	Is the evidence difficult to collect and summarize?
v)	If a public health guideline, are there specific issues that should be considered (eg, equity)
vi)	Have you made the rationale explicit?
vii)	Is this better to be formally GRADEd?

practice statements (Table 1). First, as with all recommendations, good practice statements should be clear, specific—including specification of the population of interest—and actionable. For instance, in the statement in the Box 1 regarding congenital adrenal hyperplasia, the associated text should specify the frequency and nature of the monitoring, and the action to be taken should the clinician identify signs of glucocorticoid excess.

Note that, if what is meant by monitoring is multiple additional visits to the physician specifically to check for glucocorticoid excess, whether such monitoring is beneficial and not simply a waste of resources would be called into question. As a result, this would no longer be a good practice statement. This highlights the necessity for very clearly specifying the intervention and alternative in best practice statements—which, when clearly specified, may in fact warrant formal GRADE appraisal.

Second, the message should be necessary: that is, without the guidance, clinicians might fail to take the appropriate action. Is it really plausible that clinicians who are the target audience for the guideline and who look after patients with congenital adrenal hyperplasia will fail to monitor for signs of glucocorticoid excess? If the answer is that it is not plausible, there is no need for the good practice statement.

Third, the proposed course of action should be feasible in the context considered, and it should be associated with minimal harm and cost: in other words, from the patient's point of view, the net benefit should be large and unequivocal. Furthermore, the intervention should not be associated with excessive opportunity cost—that is, panellists should consider what other, possibly more useful, interventions might be jeopardized by instituting the proposed course of action.

The fourth criterion, that evidence should indeed be difficult to collect and summarize, is an issue of opportunity cost: is the guideline panel's limited time and energy better spent on other efforts to maximize the guideline's methodological quality and overall trustworthiness? With regard to this criterion, consider the following recommendation: women with severe hypertension during pregnancy should receive treatment with antihypertensive drugs. A guideline panel issued this as a strong recommendation based on very low-quality evidence [4]. If the panel really did believe the quality of evidence was very low (ie, they were very uncertain there was net benefit), they should not have made a strong recommendation.

Is it possible, however, that the panel actually was sure there were benefits (ie, they really believed the evidence warranted high certainty) and was misapplying GRADE in the certainty judgment? If so, should this recommendation be transformed into a good practice statement?

The answer is that it should not. Presumably, the panel's logic starts with the fact that we have evidence warranting high certainty that, in nonpregnant individuals, treating severe hypertension over long periods of time results in important benefits in morbidity and mortality. This evidence is easy to find and summarize. The panel is then presumably deducing that treatment of pregnant individuals over shorter periods of time may also reduce long-term morbidity and possible mortality. The certainty that is warranted by this deduction might be a matter of debate but should be made explicit. If only low or very low certainty is warranted, a weak recommendation is appropriate. In any case, the recommendation requires a formal application of the GRADE approach.

Fifth, although the principles enunciated here apply to all guidelines, additional considerations may be required for public health guidelines intended for global audiences. Such considerations may include the cultural and ethical standards of particular populations.

Finally, given the subjective nature of the judgment that appreciable net benefit from the recommended behavior is incontestable, the rationale for that judgment should be explicit. Earlier in this article, we have provided such rationales for three of the best practice statements in the Box 1. The fourth might be "the relation between physiological measures and patients' function in heart failure is weak. Patients value their function highly, and management should be tailored to optimizing function. Without an inquiry into function, such tailored management will not be possible." The explicit statement of the rationale for the belief in benefit allows that judgment to be open to question.

2. Conclusion

We suggest that guideline panellists can best understand GRADE principles and apply these principles to the recognition of recommendations that warrant good practice statements rather than rigorous application of GRADE, by asking themselves how certain they are in estimates of effect. When they have a high level of certainty in these estimates based on the previously mentioned principles, they will also be confident that the associated clinical actions will do more good than harm, or vice versa. There will be instances in which they indeed have a high level of certainty in estimates and that high level of certainty is based on a large body of linked evidence. Because that evidence is not well described or published, formally accumulating and summarizing the evidence will be a poor use of their

time and energy. Under such circumstances, they could forego the formal GRADE process and issuing a formal GRADEd recommendation and instead make a good practice statement. In doing so, they should make clear to their audience how their good practice statements differ from formal GRADEd recommendations.

Finally, panels should be cautious and sparing in their use of good practice statements, carefully considering the necessity for the statement, making explicit their rationale, and seriously considering the possible merit of a formal GRADE assessment of the indirect linked evidence and the extent of the indirectness.

Acknowledgments

The authors thank Andy Oxman, Susan Norris, and Roman Jaeschke for their insightful suggestions on a draft version of this article and Paul Alexander for identifying the World Health Organization recommendations we have used as examples.

References

[1] Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383−94.

[2] Brito JP, Domecq JP, Murad MH, Guyatt GH, Montori VM. The Endocrine Society guidelines: when the confidence cart goes before the evidence horse. J Clin Endocrinol Metab 2013;98:3246−52.

[3] Alexander PE, Bero L, Montori VM, Brito JP, Stoltzfus R, Djulbegovic B, et al. World Health Organization recommendations are often strong based on low confidence in effect estimates. J Clin Epidemiol 2014;67:629−34.

[4] WHO. Guideline: WHO recommendations for prevention and treatment of pre-eclampsia and eclampsia. Geneva, Switzerland: World Health Organization; 2011.

[5] Speiser PW, Azziz R, Baskin LS, Ghizzoni L, Hensle TW, Merke DP, et al. Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab 2010;95:4133−60.

[6] WHO. Guideline: WHO policy on TB infection control in health-care facilities, congregate settings and households. World Health Organization; 2009.

[7] WHO guidelines on the prevention and treatment of HIV and other sexually transmitted infections for sex workers and their clients in low and middle income countries. World Health Organization. Geneva, Switzerland: World Health Organization; 2012.

[8] Hunt SA, Abraham WT, Chin MH, Feldman AM, Francis GS, Ganiats TG, et al. 2009 focused update incorporated into the ACC/AHA 2005 Guidelines for the Diagnosis and Management of Heart Failure in Adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines: developed in collaboration with the International Society for Heart and Lung Transplantation. Circulation 2009;119:e391−479.