ELSEVIER

Journal of Clinical Epidemiology 66 (2013) 726−735

Journal of Clinical Epidemiology

GRADE guidelines: 15. Going from evidence to recommendation—determinants of a recommendation's direction and strength

Jeffrey C. Andrews[a,*], Holger J. Schünemann[b,c], Andrew D. Oxman[d], Kevin Pottie[e], Joerg J. Meerpohl[f,g], Pablo Alonso Coello[h,i], David Rind[j], Victor M. Montori[k], Juan Pablo Brito[k], Susan Norris[l], Mahmoud Elbarbary[m], Piet Post[n], Mona Nasser[o], Vijay Shukla[p], Roman Jaeschke[c], Jan Brozek[b], Ben Djulbegovic[q,r], Gordon Guyatt[b,c]

[a]*Vanderbilt Evidence-based Practice Center, Vanderbilt University, #27166-719 Thompson Lane, Nashville, TN 37204-3195, USA*
[b]*Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario L8N 3Z5, Canada*
[c]*Department of Medicine, McMaster University, Hamilton, Ontario L8N 3Z5, Canada*
[d]*Norwegian Knowledge Centre for the Health Services, PO Box 7004, St. Olavs plass, Oslo 0130, Norway*
[e]*Department of Family Medicine, University of Ottawa, Ottawa, Canada*
[f]*German Cochrane Center, Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Berliner Allee 29, 79110 Freiburg, Germany*
[g]*Division of Pediatric Hematology and Oncology, Center for Pediatrics and Adolescent Medicine, University Medical Center Freiburg, Mathildenstrasse 1, 79106 Freiburg, Germany*
[h]*Iberoamerican Cochrane Center, CIBER de Epidemiología y Salud Pública, IIB, Sant Pau, Barcelona 08041, Spain*
[i]*Epidemiology and Public Health CIBER (CIBERESP), Hospital de la Sant Pau, Creu i Sant Pau, Barcelona 08041, Spain*
[j]*Harvard Medical School, Beth Israel Deaconess Medical Center, Healthcare Associates-E/Shapiro 6, 330 Brookline Aveen, Boston, MA 02215, USA*
[k]*Mayo Clinic, 200 1st SW St, Rochester, MN 55905, USA*
[l]*Department of Medical Informatics and Clinical Epidemiology, Oregon Health and Science University, Portland, OR 97239-3098, USA*
[m]*King Saud University for Health Sciences, Riyadh, Saudi Arabia*
[n]*Post Voor Zorg, Delft, The Netherlands*
[o]*Peninsula College of Medicine and Dentistry, Universities of Exeter and Plymouth, The John Bull Building, Tamar Science Park, Plymouth, PL68BU, UK*
[p]*Canadian Agency for Drugs and technologies in Health (CADTH), 600-865 Carling Avenue, Ottawa, Ontario K1S 5S8, Canada*
[q]*Division and Center for Evidence-Based Medicine and Health Outcomes Research, University of South Florida, Tampa, FL, USA*
[r]*H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA*

Accepted 3 February 2013; Published online 6 April 2013

Abstract

In the GRADE approach, the strength of a recommendation reflects the extent to which we can be confident that the composite desirable effects of a management strategy outweigh the composite undesirable effects.

This article addresses GRADE's approach to determining the direction and strength of a recommendation. The GRADE describes the balance of desirable and undesirable outcomes of interest among alternative management strategies depending on four domains, namely estimates of effect for desirable and undesirable outcomes of interest, confidence in the estimates of effect, estimates of values and preferences, and resource use. Ultimately, guideline panels must use judgment in integrating these factors to make a strong or weak recommendation for or against an intervention.

Keywords: GRADE; Quality of evidence; Strength of evidence; Guideline development; Recommendation; Evidence

1. Introduction

In prior articles in this series devoted to the GRADE approach to systematic reviews and practice guidelines, we have dealt with the process before developing recommendations, namely framing the question and choosing critical and important outcomes [1], rating the confidence in effect estimates for each outcome [2−8], dealing with resource

The GRADE system has been developed by the GRADE Working Group. The named authors drafted and revised this article. A complete list of contributors to this series can be found on the *Journal of Clinical Epidemiology* web site.

* Corresponding author. Tel.: (615) 343-5700.
E-mail address: jeff.andrews@vanderbilt.edu (J.C. Andrews).

0895-4356/$ - see front matter
http://dx.doi.org/10.1016/j.jclinepi.2013.02.003

use [9], rating the confidence in effect estimates across outcomes [10], and creating an evidence profile and a Summary of Findings table [11−13]. The immediately previous article described GRADE's approach to classifying the strength and direction of recommendations and discussed the implications of strong and weak recommendations, and the options for presentation and wording [14]. The present article presents GRADE's approach to moving from evidence to recommendations. As we did in the previous article, we will refer to guideline developers as ''the panel.''

1.1. Globalizing evidence and localizing decisions

The pithy summary by Eisenberg [15] on the relationship between evidence and recommendations, ''globalize the evidence, localize the decisions,'' provides fundamental guidance for those working to produce evidence-based recommendations [15]. Summaries of evidence regarding alternative management strategies from the medical literature should ideally be very similar, no matter the site of the application of the recommendation.

Rating of confidence in estimates of effect (quality of evidence) may, however, differ for a variety of reasons. First, desirable and undesirable outcomes may be valued differently, leading to different thresholds of acceptability. This could lead to different judgments regarding imprecision, as we have highlighted in the article in this series dealing with imprecision [5].

Second, differences in values and preferences could lead to differences in the overall balance of desirable and undesirable outcomes and the rating of confidence in estimates: an outcome judged as critical by one panel (and thus included in the rating of overall confidence in estimates) may be judged important but not critical by another (and thus not included in the overall rating).

Finally, ratings of confidence may also differ as a result of uncertainties in the risk profile of untreated populations (baseline risk). We may be very confident of baseline risk in one setting but not at all confident in another. This could lead to rating down confidence in estimates for indirectness.

Continued rapid uptake of GRADE by organizations that produce systematic summaries of evidence will greatly facilitate the production of transparent evidence summaries. If organizations work together to produce summaries, there will be an enormous gain in efficiency [16]—even if, in the end, judgments about confidence in estimates will differ across settings, for reasons described in the preceding paragraphs.

We now turn to a systematic presentation of the determinants of direction and strength of recommendations.

2. Determinants of direction and strength of recommendations

GRADE has identified six determinants of the direction and strength of recommendations, namely the magnitude of estimates of effect of the interventions on important outcomes, confidence in those estimates, estimates of typical values and preferences, confidence in those estimates, variability of values and preferences, and resource use. In the presentation here, we will present these six determinants in four domains. We package magnitude of effect and typical values and preferences together with the label balance of desirable and undesirable consequences or ''trade-offs.'' We also include uncertainty regarding typical values, and variability in values, in a single domain (Table 1).

Alternative groupings may work better, depending on the circumstances. We believe that the approach we present here is best for presenting the rationale for the recommendations to the guideline consumer audience. In developing recommendations, panels may want to keep all six determinants separate or group the three values and preferences determinants together.

Ultimately, guideline panels must integrate these six determinants to make a strong or weak recommendation for or against an intervention. Table 2 illustrates how the elements of the GRADE framework for moving from evidence to recommendations can be applied in making strong and weak recommendations, and Table 3 provides an example of the application in the management of chronic obstructive pulmonary disease.

2.1. Trade-offs between desirable and undesirable consequences of alternative management strategies

When we consider the balance between desirable and undesirable outcomes (''trade-offs''), we are considering two domains. The first is our best estimates of the magnitude of desirable effects and the undesirable effects. If a guideline panel has adhered to the GRADE process, they will find the best estimates of effect in the evidence profiles that they have prepared or accessed.

The second element that determines the balance among desirable and undesirable outcomes is the typical values that patients—or a population—apply to those outcomes. This can be otherwise conceptualized as the relative preferences for those outcomes—and thus the term we generally use, values and preferences (Box 1).

Ideally, to inform estimates of typical patient values and preferences, guideline panels will conduct or identify systematic reviews of relevant studies of patient values and preferences [18]. Given the paucity of empirical examinations of patients' values and preferences, however, well-resourced guideline panels will usually complement such studies with consultation with individual patients and patients' groups. The panel should discuss whose values these people represent, namely representative patients, a defined subset of patients, or representatives of the general population.

For example, the Canadian Collaboration for Immigrant and Refugees Health (CCIRH) guidelines sought to advance understanding of immigrant patient perspectives in