

Title: Making the GRADE: Providing Clinical Practice Guidance to the European Association of Urology

Keywords: EAU guidelines, Methodology, GRADE, AGREE II

Word count: 994

Authors and affiliations:

Steven E. Canfield, MD, Chair Methods Committee, EAU Guidelines Office

Muhammad Imran Omar, MD, Vice-Chair Methods Committee, EAU Guidelines Office

Maria J. Ribal, MD, Chair, EAU Guidelines Office

On behalf of the EAU Guidelines Office Board

Corresponding author:

Steven E. Canfield, MD

University of Texas McGovern Medical School

6431 Fannin St.

Medical School Building 6.018

Houston, TX 77030

USA

Telephone: +1 713-500-7335

Fax: +1 713-500-7319

The importance of providing clinical guidance based on the best available knowledge cannot be overstated – the task is one of the most crucial mandates of any medical organization. The challenge has always been the changing nature of our understanding, of medical treatments, of course, but also of what defines the best available knowledge. The first definitions outlining a potential “hierarchy” of knowledge were set down in 1979 [1] by a task force which included David Sackett, the “father” of Evidence-Based Medicine. He later published the original “Rules of Evidence”, based on **consensus agreement**, which suggested *grades* of recommendations [2]. From the beginning, the concept that a hierarchy of empiric, data driven evidence should inform clinical decision making existed.

Understanding that the pure “levels” of evidence do not provide deep analysis of study quality, experts in methodology have since attempted to devise a better way to rank recommendations. One of the most widely adopted systems today is the Grading of Recommendations Assessment, Development and Evaluation (GRADE), which seeks to provide a transparent framework for moving from evidence to recommendation [3]. A core tenet of GRADE is to separate evidence quality (aka certainty) from the “strength” of a recommendation, but still acknowledge quality as a major component of that strength rating. To use GRADE properly requires a formal assessment with systematic review of the evidence underpinning each recommendation. This assumes the availability of certain types and quantities of evidence. For many topics requiring practical clinical guidance, this evidence simply does not exist, which creates a true branch point and dilemma for attempted use of the GRADE system by guideline organizations – a problem which the GRADE group does not acknowledge. Although GRADE has been implemented by many organizations worldwide, the system has its detractors, who point out there “is a central paradox: while GRADE has evolved through the evidence-based medicine movement, there is no evidence that GRADE itself is reliable.” In other words, *GRADE itself has not been validated* [4].

Over the past decade, the EAU Guidelines have endeavored to improve and incorporate the body of evidence which informs our recommendations, including by creating a Methods Committee and, when possible, following similar methods as those proposed by the GRADE group, in a modified fashion [5]. Such a process starts with synthesizing the evidence via systematic review [6]. With 21 separate guidelines, and 1865 individual recommendations historically based on the modified Oxford levels of evidence, this task must be approached in a practical manner which continues to provide clinical guidance in a familiar format to our more than 16,000 members. To date the Guidelines Office, through its innovative Associates Programme [5], has produced over 200 unique systematic reviews which have directly impacted the guidelines.

Our publicly available Guidelines Office Development Handbook [5] describes the methods currently in use, which closely follow the Cochrane Handbook and GRADE format: *From 2018 onwards, the EAU Guidelines have been using a modified GRADE approach for the grading of recommendations. Accordingly, a Summary of Evidence (SOE) table is created for each recommendation within the guidelines which addresses: 1. The overall quality of the evidence which exists for the recommendation, 2. The magnitude of the effect (individual or combined effects), 3. The certainty of the results (precision, consistency, heterogeneity and other statistical or study related factors), 4. The balance between desirable and undesirable outcomes, 5. The impact of patient values and preferences on the intervention, and 6. The certainty of those patient values and preferences. These key elements in the SOE tables are the basis which panels use to define the strength and direction of each recommendation. Panels can provide both ‘strong’ and ‘weak’ recommendations ‘for’ or ‘against’ recommending an action based on the information found in the SOE tables. The strength of each recommendation is determined by the*

balance between desirable and undesirable consequences of alternative management strategies, the quality of the evidence (including certainty of estimates), and nature and variability of patient values and preferences.

The GRADE process allows for and may result in many more “strong” recommendations, despite low quality evidence, based on other domains. In this month’s issue of European Urology, Dahm and colleagues analyzed the adherence of EAU Guidelines to the GRADE system and report the findings that “64.0% were categorized as strong and 36.0% as weak” in their letter to the editor [7]. The GRADE group insists that “Separating the judgments regarding the quality of evidence from judgments about the strength of recommendations is a critical and defining feature of this new grading system” [8]. Nonetheless, there is an inherent bias among advocates of GRADE that low quality evidence should not lead to a strong recommendation, and guidelines which differ should be criticized. Dahm and colleagues go on to list that five of the six recommended GRADE metrics appear to be missing from all but one EAU guideline. In reality, these methods are present [5], but descriptions can only be found in the central website handbook, and evidence summaries are not yet publicly available. These issues will be resolved, and we appreciate the authors’ suggestions and interest in our process.

It should be emphasized that the most widely accepted assessment tool used to evaluate the quality of a clinical practice guideline is the AGREE II instrument, which looks at six separate process domains, including scope and purpose, stakeholder involvement, rigor of development, clarity of presentation, applicability, and editorial independence [9]. Note that production of high-quality Clinical Guidelines comprises numerous steps beyond the final recommendations. Only one step relates to methods. Furthermore, the “rigor of development” domain in AGREE II does not specify use of GRADE, or any other specific methodology, and use of GRADE itself is not a metric for guideline quality. EAU Guidelines achieve high AGREE II scores in all domains, and have been shown to outperform other globally respected guidelines when directly compared [10]. It remains our goal to ensure that every guideline produced by the EAU meets these proven standards, while keeping current with the ever-changing landscape of evidence, knowledge, and the health of our patients.

Authors' Disclosure page

Steven E. Canfield, MD, Chair, EAU Guidelines Methods Committee

Muhammad Imran Omar, MD, Vice-Chair, EAU Guidelines Methods Committee

Maria J. Ribal, MD, Chair, EAU Guidelines

No other relevant disclosures

References:

- [1] Spitzer, WO, Bayne, RD, Charron, KC, et al. The periodic health examination. Canadian Task Force on the Periodic Health Examination. *Can Med Assoc J.* 1979;121(9):1193-1254.
- [2] Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest.* 1989;95(2 Suppl):2S-4S.
- [3] Guyatt, GH, Oxman, AD, Vist, GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ.* 2008;336:924.
- [4] Kavanagh BP. The GRADE system for rating clinical guidelines. *PLoS Med.* 2009;6(9):e1000094. doi:10.1371/journal.pmed.1000094
- [5] Methodology & Policies. European Association of Urology. EAU Guidelines Handbook. <https://uroweb.org/eau-guidelines/methodology-policies>
- [6] Knoll T, Omar MI, Maclennan S, et al. Key Steps in Conducting Systematic Reviews for Underpinning Clinical Practice Guidelines: Methodology of the European Association of Urology. *Eur Urol.* 2018;73(2):290-300. doi:10.1016/j.eururo.2017.08.016
- [7] Dahm P, Cleveland B, Lauwagie A, Daniel Gonzalez-Padilla DA. Adherence of the European Association of Urology Guidelines to Grading of Recommendations Assessment, Development and Evaluation (GRADE) Methodology. *European Urology* (in press).
- [8] Schünemann HJ, Jaeschke R, Cook DJ, et al. An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. *Am J Respir Crit Care Med.* 2006;174(5):605-614. doi:10.1164/rccm.200602-197ST
- [9] Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ.* 2010;182(18):E839-E842. doi:10.1503/cmaj.090449
- [10] Dimitropoulos K, Omar MI, Chalkias A, Arnaoutoglou E, Douketis J, Gravas S. Perioperative antithrombotic (antiplatelet and anticoagulant) therapy in urological practice: a critical assessment and summary of the clinical practice guidelines. *World J Urol.* 2020;38(11):2761-2770. doi:10.1007/s00345-020-03078-2