

COMMENTARY

## Challenges in guideline methodology

Sharon E. Straus\*, Sasha Shepperd

*Department of Medicine, St. Michael's Hospital, Toronto, Ontario, Canada M5B 1W8*

Accepted 19 January 2011

Accompanying the international focus on optimizing quality of care and reducing practice variation there has been an explosion in the development of clinical practice guidelines to help clinicians, patients, and policy makers make decisions that are consistent with the evidence. For example, a quick search of the National Guidelines Clearinghouse reveals that there are more than 400 guidelines on heart failure [1]. It could be argued that rather than reducing practice variation, this plethora of guidelines has become a source of variation in care. Further complicating this situation for clinicians and policy makers who are trying to apply these various guidelines is the inconsistent approach to grading evidence and making recommendations by guideline developers. A systematic review in 2002 found more than 100 systems for rating the quality of evidence [2]. A more recent comparison of some of these instruments found that different recommendations arise with use of these various approaches [3]. The result—guidelines addressing the same clinical topic are produced by different organizations resulting in different recommendations, leading to confusion for those trying to implement the guidelines [4,5].

Grading of Recommendations Assessment, Development and Evaluation (GRADE) is an attempt to provide clarity through transparent (and hopefully reliable) methods to rate the quality of evidence and strength of recommendations [6]. To date, it has been adopted both by systematic review authors and by many guideline producers, including the newly reconstituted Canadian Task Force on Preventive Health Care [7]. In this issue of the Journal, we introduce a new series that attempts to provide a “users’ guide” to GRADE to facilitate use by guideline developers and those using systematic reviews. The series will also include commentaries by authors who have used GRADE in various settings and contexts [8,9]. Bedford et al. [8] and English and Opiyo [9] highlight some of the advantages and challenges with its use, including the need for extensive training and

resources to use GRADE reliably, confusion around how to use it for studies outside of those focusing on therapy or diagnosis, and the persistent need to make “scientific value judgments” about a body of evidence.

The GRADE Working Group should be congratulated for their efforts on developing an international collaborative process for advancing the methods of assessing evidence. Similarly, there have been other international groups, such as Appraisal of Guidelines Research and Evaluation (AGREE) [10] and ADAPTE [11] (<http://www.g-i-n.net/activities/adaptation>), that have attempted to enhance guideline development and use. However, substantial gaps in guideline methodology remain, reflecting that most guidelines focus on a single target condition, process, or diagnostic strategy rather than multiple target conditions, their associated comorbidities, and complex care pathways. It is not obvious how a GRADE evidence profile for one outcome should be integrated into this complex clinical context. These are issues health care providers and policy makers grapple with as they attempt to implement guidelines. Research by the Guideline Implementability Appraisal group [12] provides some guidance on identifying barriers to implementation, and work by the Supporting Policy Relevant Reviews and Trials (SUPPORT) group addresses issues around the application of evidence from systematic reviews [13]. However, there is scope for additional research on how to improve the “implementability” [14] of guideline recommendations. Ideally, the implementation of high-quality evidence improves organization and quality of care, but many studies show variable success in achieving these goals [15]. This challenge highlights that the field of implementation science or knowledge translation is still in its infancy and that there are vast gaps in our knowledge of how best to implement evidence from guidelines and systematic reviews.

We invite readers to identify key challenges in guideline methodology and implementation science and propose issues that you think this journal should address in this area. We would also like to hear from people who have used GRADE so that others and we may benefit from your experience. If you would like to offer a suggestion, send correspondence to the *Journal of Clinical Epidemiology* at [ltugwell@uottawa.ca](mailto:ltugwell@uottawa.ca).

\* Corresponding author. Department of Medicine, St Michael's Hospital, 30 Bond Street, Toronto, Ontario, Canada M5B 1W8. Tel.: 416-864-3068.  
E-mail address: [sharon.straus@utoronto.ca](mailto:sharon.straus@utoronto.ca) (S.E. Straus).

## References

- [1] National Guidelines Clearinghouse. Available at [www.guidelines.gov](http://www.guidelines.gov). Accessed December 2, 2010.
- [2] West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, et al. Systems to rate the strength of scientific evidence. Evidence Report/Technology Assessment No. 47 (prepared by the Research Triangle Institute-University of North Carolina Evidence-based Practice Center under Contract No. 290-97-0011). AHRQ Publication No 02-E016. pp. 64–88. Rockville, MD: Agency for Healthcare Research and Quality; 2002.
- [3] Ferreira PH, Ferreira ML, Maher CG, Refshauge K, Herbert RD, Latimer J. Effect of applying different “levels of evidence” criteria on conclusions of Cochrane reviews of interventions for low back pain. *J Clin Epidemiol* 2002;55:1126–9.
- [4] Fahey TP, Peters TJ. What constitutes controlled hypertension? Patient based comparison of hypertension guidelines. *BMJ* 1996;313:93–6.
- [5] Manuel DG, Kwong K, Tanuseputro P, Lim J, Mustard CA, Anderson GM, et al. Effectiveness and efficiency of different guidelines on statin treatment for preventing deaths from coronary heart disease: modeling study. *BMJ* 2006;33:1419.
- [6] GRADE guidelines 1. Introduction—GRADE evidence profiles and summary of findings tables
- [7] Canadian Task Force on Preventive Health Care. Available at <http://www.canadiantaskforce.ca/>. Accessed December 2, 2010.
- [8] Bedford M, Pettersen K, Minhas R. Strength of evidence and handling uncertainty: practical considerations and general observations. *J Clin Epidemiol*, in press.
- [9] English M, Opiyo N. Getting to grips with GRADE: perspectives from a low income setting. *J Clin Epidemiol*, in press.
- [10] Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *J Clin Epidemiol* 2010;63:1308–11.
- [11] Harrison M, Graham I, Fervers B; on behalf of the ADAPTE group. Adapting knowledge to a local context. In: Straus SE, Tetroe J, Graham I, editors. *Knowledge translation in health care*. Oxford, UK: Wiley Blackwell; 2009; pp 116–25.
- [12] Shiffman RN, Dixon J, Brandt C, Essaihi A, Hsiao A, Michel G, et al. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. *BMC Med Inform Decis Mak* 2005;5:23.
- [13] Lavis JN, Oxman AD, Souza NM, Lewin S, Gruen RL, Freheim A. SUPPORT Tools for evidence-informed health Policymaking (STP) 9: assessing the applicability of the findings of a systematic review. *Health Res Policy Syst* 2009;16(7 Suppl 1):S9.
- [14] O. Bhattacharyya O, Estey E, Kastner M, Straus S, Grimshaw J, Zwarenstein M, et al. Adopting a realist review approach to conceptualizing the relationship between the perceived characteristics of clinical practice guidelines and their uptake. Chicago, IL: Guidelines International Network Conference; August 25-28, 2010, Chicago, IL.
- [15] Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2004;8:iii–iv. 1–72.