Editorial

Do "Evidence-Based Recommendations" Need to Reveal the Evidence? Minimal Criteria Supporting an "Evidence Claim"

"Evidence-based medicine" (EBM) stresses examining evidence from clinical research[1] as the preferred method of clinical decision making, de-emphasizing intuition, unsystematic clinical experience, and pathophysiologic rationale. Likewise, evidence-based practice (EBP) holds that evidence should be the basis for particular interventions and management plans that are likely applicable to most patients. Logically, this practice would demand that experts develop evidence-based recommendations founded on valid, reliable, and transparent systematic reviews and/or metaanalyses[2,3].

In this issue of *The Journal*, Roubille, *et al* present recommendations for the management of comorbidities, focusing on 8 areas within rheumatoid arthritis (RA), psoriasis (PsO), and psoriatic arthritis (PsA), based on a review of 407 articles[4]. Their report summarizes the results of the Canadian Dermatology-Rheumatology (DR) Comorbidity Initiative's systematic literature searches and consensus-based recommendations from a meeting held in Toronto in 2013, sponsored by the pharmaceutical company AbbVie[4]. The authors report that they did a thorough, systematic review, followed by data extraction and subsequent metaanalyses (including forest plots summarizing the adjusted relative risk estimates, etc.). However, the authors apparently do not want to reveal their explicit findings (yet!), stating that the details and results of the systematic literature review for each topic will be published separately[4]. We have no reason to question the integrity or the content of the recommendations that came out of this work, but we worry about a possible trend that could encourage guideline panels, etc., to publish their evidence-synthesis secondarily to their recommendation while claiming in a peer-reviewed journal (like *The Journal of Rheumatology*) that their work represents evidence-based recommendations.

Although realizing that we might be perceived as having intellectual conflicts of interest (i.e., being editors in the Cochrane Collaboration), we would like to encourage systematic and explicit methods of making judgments because they will likely reduce errors and improve communication[5]. The aim of the Cochrane Collaboration is to help healthcare providers, patients, patient advocates and carers, and policy makers arrive at well-informed decisions on healthcare treatments by preparing, maintaining, and disseminating methodologically strong systematic reviews (http://www.cochrane.org/). Briefly, the Cochrane Collaboration's mission is to provide accessible, credible information to support informed decision-making. We strongly believe that users of clinical practice guidelines and other recommendations should know how much confidence they can place in recommendations. Roubille, *et al*[4] did not report the actual data; instead — "saving the evidence" for later, subsequent papers. We find, therefore, that their claim in a peer-reviewed journal that their recommendations are evidence-based is not supported, as their recommendations are missing any presentation of the actual evidence to the peer reviewer or to users of these guidelines.

We would ask, rhetorically: Do evidence-based recommendations need to reveal their evidence? According to the current thinking in EBM, the answer is probably 'yes.' The Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) Working Group has developed a system of rating quality of evidence that improves reliability in comparison to intuitive judgments about the evidence[6]. One of the key principles behind the GRADE methodology is the aim to make it easier for users to assess the judgments behind recommendations[7,8].

Evidence summaries (e.g., based on forest plots, etc.) should be available to the user of a recommendation, enabling a fair judgment of the quality of evidence supporting the strength of recommendations. Ideally, the summary would consist of full GRADE evidence profiles or a summary-of-findings (SoF) table based on a systematic review[9]. At a minimum, the evidence that was assessed and the methods that were used to identify and appraise it should be clearly described. Making recommendations matter to end users requires specifying *a priori* the relevant setting, population, intervention, comparator, and the relative importance of the outcomes before gathering the evidence and again when evidence summaries are complete[10].

See RA, PsA, and PsO recommendations, *page 1767*

Judging the Quality of the Evidence

Evidence from randomized trials without important limitations constitutes high-quality evidence, whereas observational studies without special strengths or important limitations constitute low-quality evidence. However, while looking at all the evidence after an appropriate "meta-strategy,"[4] evidence-based recommendations should explicitly delineate each of the following criteria that can lower the quality of the evidence (i.e., a consequence of reduced confidence in the estimate of effect)[11]:

- Risk of bias/study limitations[12]
- Risk of publication bias[13]
- Imprecision[14]
- Inconsistency of results[15]
- Indirectness[16].

Any of these factors could mandate that we rate down the quality of the evidence. On the other hand, if observational studies are also included in the evidence-synthesis, extra criteria (e.g., a large magnitude of effect) might apply when rating the overall quality of the evidence[17], as their presence might increase our confidence in the estimate of effect[18].

We believe that end users of evidence-based recommendations require succinct, transparent, easily digested evidence summaries. The GRADE process facilitates the creation of summaries in the form of evidence profiles and/or SoF tables[5]. It is important to keep in mind that the judgment available in an SoF table (usually) does not refer to individual studies but rather to a body of evidence. An SoF table presents the credibility of all the available evidence answering a specific clinical question; that is, the body of evidence — such as from a number of trials with good internal validity — may be associated with a low risk of bias, but our confidence in the estimate may be compromised by a number of other factors (as illustrated above)[11].

Although the quality of evidence that goes into an evidence-based recommendation spans a continuum, the approach recommended by GRADE results in either high, moderate, low, or very low confidence in the estimate for any given clinical question[6].

GRADE provides definitions on how to interpret the different levels of evidence[11], which have implications for researchers and funding bodies that are considering whether a research question is valuable or not:

High quality evidence. We are very confident that the true effect lies close to the estimated effect (i.e., no further studies are needed).

Moderate quality evidence. We feel confident in the estimated effect, but it might be different from the true effect (i.e., we still have some research we would like to see).

Low quality evidence. Our confidence in the estimated effect is limited, as the true effect may be substantially different (i.e., we need rigorous prospective studies).

Very low quality evidence. We have very little confidence in the proposed estimated effect, and the true effect is likely substantially different (i.e., at present the evidence is mostly based on opinions and/or unsystematic clinical observations).

Determining the Strength of a Recommendation

We strongly recommend that all organizations adhere to the GRADE process (or a similar transparent approach) when developing recommendations regarding estimation of benefit and harm, as well as when judging the quality of evidence. With the GRADE approach, the strength of a recommendation reflects the extent to which we can be confident that the desirable effects of a management strategy outweigh the undesirable effects. For example, the European League Against Rheumatism (EULAR) states in its standardized operating procedures for the elaboration, evaluation, dissemination, and implementation of recommendations (updated recently)[19] that the GRADE system should be used for guidance throughout the elaboration of clinical recommendations. The American College of Rheumatology (ACR) makes a similar recommendation. Unfortunately, EULAR still endorses a system that is difficult to remember and communicate when writing recommendations[19]. Unlike other previous grading systems, the GRADE approach has only 2 categories for recommendations: either strong or conditional (either for or against). A strong recommendation reflects our confidence that the desirable effects of adherence outweigh the undesirable effects[20]. Conditional recommendations, on the other hand, indicate that although the desirable effects of adherence probably outweigh the undesirable effects, the panel is less confident that health care professionals will recognize that different choices will be appropriate for different patients. Conditional recommendations for a given intervention imply the importance of shared decision making for delivering patient-centered care; healthcare professionals must help each patient to arrive at a management decision consistent with the patient's values and preferences.

In this issue of *The Journal*, Roubille, *et al* present thorough work, based on expert opinions of the Canadian Dermatology-Rheumatology Comorbidity Initiative[4]. As the authors do not report the details and results of their systematic reviews, we do not support the claim that this report represents evidence-based recommendations at this stage[20]. Perhaps once the evidence synthesis is published and made available to the users, it might meet this requirement and earn this label.

ROBIN CHRISTENSEN, BSc, MSc, PhD,
Musculoskeletal Statistics Unit,
The Parker Institute,
Department of Rheumatology,
Bispebjerg and Frederiksberg Hospital,
Copenhagen, Denmark;
JASVINDER A. SINGH, MBBS, MPH,
Medicine Service, Birmingham VA Medical Center,
Department of Medicine at School of Medicine,

and Division of Epidemiology at School of Public Health,
University of Alabama, Birmingham, AL;
and Department of Orthopedic Surgery,
Mayo Clinic College of Medicine,
Rochester, MN, USA;
GEORGE A. WELLS, PhD,
University of Ottawa Heart Institute,
Ottawa, Canada;
PETER S. TUGWELL, MD,
School of Epidemiology, Public Health and Preventive Medicine,
University of Ottawa,
Ottawa, Canada.
The Musculoskeletal Statistics Unit at the Parker Institute (RC) is supported by grants from the Oak Foundation. JAS is supported by grants from the Agency for Health Quality and Research Center for Education and Research on Therapeutics (U19 HS021110), US National Institute of Arthritis, Musculoskeletal and Skin Diseases (P50 AR060772 and U34 AR062891), National Institute of Aging (U01 AG018947), National Cancer Institute (U10 CA149950), and the Patient Centered Outcomes Research Institute (research contract CE-1304-6631).
Address correspondence to Prof. Christensen, Copenhagen University Hospital at Frederiksberg, Musculoskeletal Statistics Unit, The Parker Institute, Nordre Fasanvej 57, Copenhagen F, DK-2000, Denmark.
E-mail: robin.christensen@regionh.dk

REFERENCES

1. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. JAMA 1992;268:2420-5.
2. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009;62:e1-34.
3. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol 2009;62:1013-20.
4. Roubille C, Richer V, Starnino T, McCourt C, McFarlane A, Flemming P, et al. Evidence-based recommendations for the management of comorbidities in rheumatoid arthritis, psoriasis, and psoriatic arthritis: expert opinion of the Canadian Dermatology-Rheumatology Comorbidity Initiative. J Rheumatol 2015;42:1767-80.
5. Ghogomu EA, Maxwell LJ, Buchbinder R, Rader T, Pardo Pardo J, Johnston RV, et al; Editorial Board of the Cochrane Musculoskeletal Group. Updated method guidelines for Cochrane Musculoskeletal Group systematic reviews and metaanalyses. J Rheumatol 2014;41:194-205.
6. Mustafa RA, Santesso N, Brozek J, Akl EA, Walter SD, Norman G, et al. The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. J Clin Epidemiol 2013;66:736-42.
7. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al; GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ 2004;328:1490.
8. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:924-6.
9. 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.
10. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol 2011;64:395-400.
11. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64:401-6.
12. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence — study limitations (risk of bias). J Clin Epidemiol 2011;64:407-15.
13. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias. J Clin Epidemiol 2011;64:1277-82.
14. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence—imprecision. J Clin Epidemiol 2011;64:1283-93.
15. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al; GRADE Working Group. GRADE guidelines: 7. Rating the quality of evidence—inconsistency. J Clin Epidemiol 2011;64:1294-302.
16. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al; GRADE Working Group. GRADE guidelines: 8. Rating the quality of evidence—indirectness. J Clin Epidemiol 2011;64:1303-10.
17. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al; GRADE Working Group. GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol 2011;64:1311-6.
18. Glasziou P, Chalmers I, Rawlins M, McCulloch P. When are randomised trials unnecessary? Picking signal from noise. BMJ 2007;334:349-51.
19. van der Heijde D, Aletaha D, Carmona L, Edwards CJ, Kvien TK, Kouloumas M, et al. 2014 Update of the EULAR standardised operating procedures for EULAR-endorsed recommendations. Ann Rheum Dis 2015;74:8-13.
20. Schünemann HJ, Wiercioch W, Etxeandia I, Falavigna M, Santesso N, Mustafa R, et al. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. CMAJ 2014;186:E123-42.

J Rheumatol 2015;42:1737–9; doi:10.3899/jrheum.150846