Evidence-Based Medicine 20 Years On: A View from the Inside

In this this issue, Seshia & Young report a comprehensive review of the literature[1,2] addressing the evolution of the Evidence-based Medicine (EBM) paradigm.[3] In addition to illustrating the profusion of concepts and terms associated with EBM, the authors have highlighted the current controversies in this area and demonstrated that, for many issues, there are no definitive answers. Although useful, there is a risk that a puzzled reader may be left with nothing but uncertainties. In this editorial we will focus on the current challenges that clinicians face when trying to apply EBM in practice, and how recent developments can help them meet these challenges.

Clinicians must find the current best evidence to answer their questions. In parallel to the widespread uptake of EBM, the volume of research has been dramatically increasing, with now more than 2000 articles published in MEDLINE every day, including 75 randomized controlled trials.[4] Clinicians therefore need resources that filter, appraise, and synthesize the evidence and facilitate access to this processed information at the point of care. Such resources include systematic reviews, synopses of high quality studies or reviews (often published in evidence-based journals, such as ACP Journal Club), online evidence-based textbooks (e.g. UpToDate, Dynamed), and clinical practice guidelines.[5]

As pointed out by Seshia and Young[2], some evidence is more trustworthy than other, and evidence summaries must distinguish between the more and less trustworthy. Whoever makes this assessment must be familiar with many evolving methodological and statistical concepts often far from the expertise of individual clinicians (think for example of non-inferiority trials with composite outcomes, subgroup analysis with tests of interaction, or network meta-analyses). The ambitious clinician can look to publications that provide practical guidance in each of these areas[6], and most EBM resources mentioned provide some guidance regarding the trustworthiness of the evidence.

The state-of-the-art thinking regarding the assessment of trustworthiness of evidence is captured by GRADE (*Grading of Recommendations Assessment, Development and Evaluation*)[7] methodology. Adopted by more than 70 organizations[8], GRADE is a systematic framework for the formulation of treatment recommendations that offers a transparent and systematic assessment of the confidence in estimates of effect (also known as quality of the evidence) for each patient-important outcome (e.g. mortality, major stroke, quality of life).[9]

In the GRADE system, randomized trials begin as high quality evidence (on a four-category rating of high, moderate, low and very low) but can be rated down for one of five limitations in the body of evidence available (Figure).[10] The first of these limitations is risk of bias, resulting from the absence of well-established protective measures (allocation concealment, blinding, intention-to-treat analysis and appropriate follow-up) but also more recently recognized limitations (stopping early for apparent benefit and selective reporting of outcomes). The second component is imprecision in the estimates of effect, assessed by the width of the confidence intervals. The third is indirectness, or the degree in which the results do not apply to the patient population under consideration, are based on inadequate comparisons, or are assessed by surrogate outcomes. The fourth and fifth components relate to inconsistency in

Study Design	Confidence in estimates	Lower if	Higher if
Randomized trial →	High	Risk of bias	Large effect
	Moderate	– 1 Serious	+ 1 Large
Observational study →	Low	– 2 Very Serious	+ 2 Very large
	Very low		
		Imprecision	Dose response
		– 1 Serious	+ 1 Evidence of a gradient
		– 2 Very Serious	
			All plausible confounding
		Indirectness	+1 Would reduce a
		– 1 Serious	demonstrated effect
		– 2 Very Serious	or
			+1 Would suggest a
		Inconsistency	spurious effect when
		– 1 Serious	results show no effect
		– 2 Very Serious	
		Publication bias	
		– 1 Serious	
		– 2 Very Serious	

Figure: Quality assessment criteria in GRADE.

results across studies and publication bias. Observational studies start as low quality, but can be rated up for a number of reasons, including a large magnitude of effect and the presence of a dose-response gradient (Figure).

This systematic and transparent approach to assess quality addresses the common criticism that EBM overly relies on a simplistic hierarchy of evidence. In GRADE, randomized trials (and consequently systematic reviews of such trials) start at high confidence, but are often rated down for one or more of these five limitations. Conversely, observational studies start at low confidence, but can ultimately be judged as providing high quality evidence. Further, GRADE has recently highlighted the importance of methodologically rigorous observational studies in providing trustworthy estimates of baseline risk that will impact on confidence in absolute estimates of treatment effect.[11] Future developments in GRADE also include the evaluation of evidence for questions of diagnosis and prognosis.

Finally, another challenge for clinicians is using evidence at the point of care. As EBM leaders have long pointed out, and as GRADE has emphasized, evidence is not the only ingredient for efficient and patient-centered decision-making. Issuing actionable recommendations not only requires considering the balance between benefits and harms and the quality of this evidence in a specific clinical context, but also patients' values and preferences.[12] Guideline panels using the GRADE approach will issue strong recommendations when they believe that, across the range of patients' values and preferences, virtually all fully informed patients will make the same decision. Conversely, they will issue a weak recommendation when there is a close balance in benefits and harms, when quality of evidence is low, or when there is substantial uncertainty, or large variability, in patients' values and preferences. Thus, weak recommendations, which often account for more than half of the recommendations in GRADE guidelines, identify preference-sensitive alternatives that warrant share-decision making. This approach represents the farthest thing from a "cookbook" approach for decision-making, a common mischaracterization of EBM.

Current guidelines formats will not suffice to make evidence-based share-decision making a reality. New approaches are needed to disseminate guidelines at the point of care. Through several initiatives (http://www.decide-collaboration.eu and http://www.magicproject.org) the GRADE working group continues active research to develop new presentation formats that include interactive decision aids to share appraised evidence with patients in the clinical encounter.[13,14]

In conclusion, we agree with Seshia and Young that EBM is an evolving paradigm[1,2] that integrates new developments in research methods and advances in clinical application of evidence and shared decision making.[15] GRADE, which offers exciting possibilities for clinicians to find, appraise and share current best evidence with patients at the point of care, represents the cutting edge of advances in EBM.

Thomas Agoritsas, Gordon H. Guyatt
Department of Clinical Epidemiology and Biostatistics
McMaster University, Hamilton, Ontario, Canada
Email: thomas.agoritsas@gmail.com

REFERENCES

1. Seshia SS, Young GB. The evidence-based medicine paradigm: where are we 20 years later? Part 1. Can J Neurol Sci. 2013;40 (4):465-74.
2. Seshia SS, Young GB. The evidence-based medicine paradigm: where are we 20 years later? Part 2. Can J Neurol Sci. 2013;40 (4):475-81.
3. Guyatt G. Evidence-based medicine. ACP J Club. 1991;114(suppl 2):A-16.
4. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS Med. 2010;7:e1000326.
5. Haynes RB. Of studies, syntheses, synopses, summaries, and systems: the "5S" evolution of information services for evidence-based healthcare decisions. Evid Based Med. 2006;11: 162-4.
6. Guyatt G, Rennie D, Meade MO, et al: JAMA Evidence users' guide to the medical literature. A manual for evidence-based clinical practice. (ed 2nd). Chicago, IL: McGraw Hill Co; 2008.
7. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336:924-6.
8. Organizations that have endorsed or that are using GRADE - [Accessed 2013 May 14] http://www.gradeworkinggroup.org/society/index.htm.
9. Guyatt GH, Oxman AD, Schunemann HJ, et al. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiol. J Clin Epidemiol. 2011;64:380-2.
10. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011;64: 401-6.
11. Spencer FA, Iorio A, You J, et al Uncertainties in baseline risk estimates and confidence in treatment effects. BMJ. 2012;345: e7401.
12. Andrews J, Guyatt G, Oxman AD, et al. GRADE guidelines: 15. Going from evidence to recommendations: the significance and presentation of recommendations. J Clin Epidemiol. 2013 [Epub ahead of print]
13. Treweek S, Oxman AD, Alderson P, et al. Developing and evaluating communication strategies to support informed decisions and practice based on evidence (DECIDE): protocol and preliminary results. Implement Sci. 2013;8:6.
14. Vandvik PO, Brandt L, Alonso-Coello P, et al. Creating clinical practice guidelines we can trust, use and share: a new era is imminent. Chest. In press 2013.
15. Montori VM, Guyatt GH. Progress in evidence-based medicine. JAMA. 2008;300:1814-6.