EDITORIAL
Updating the Assignment of Levels of Evidence

Orthopaedic surgeons, like all physicians, must make clinical decisions based on the best available evidence. This evidence comes from individual clinical experience and external sources[1]. Although clinical experience is left to the physician, the medical and surgical literature provides the best external evidence. To facilitate the process of determining the best evidence to answer a clinical question, *The Journal of Bone & Joint Surgery* assigns level of evidence (LOE) ratings to all clinical articles.

Since 2003[2], *The Journal* has used a hierarchical rating system based on the recommendations of the Centre for Evidence-Based Medicine (CEBM) in Oxford, United Kingdom, to rank articles according to the study design used to answer the primary research question[3]. In 2011, the CEBM updated its recommendations. After robust internal and external discussion, *The Journal* has decided to keep pace with the CEBM and has updated our LOE table.

The new LOE table emphasizes the clinical applications of research findings and encourages a more holistic assessment of study design and execution. Those familiar with the original table will notice that this update retains many features of its predecessor[4]. Nonclinical articles (such as cadaveric and animal studies) are still excluded from the ranking system, studies are still divided by type (therapeutic, prognostic, diagnostic, or economic analysis), and much of the ranking criteria remains the same.

Although the new table borrows from the original, it also represents an important departure. The most apparent change is structural. The rows and columns have been transposed, and there is an additional column for clinical questions. This new design reflects the order and the types of questions that arise in the process of clinical care. In this way, the table continues to provide a hierarchy of evidence, but it also assumes a new role, guiding busy clinicians to the best available evidence in real time. Whereas interaction with the original table was limited to authors and editors, the new table will engage readers more directly. Readers are encouraged to formulate their clinical question and to consult the table to determine how to conduct their search. For example, if a clinician asks, "does this intervention help?" the table will direct the reader to seek Therapeutic Level-I (randomized controlled trial) studies first, followed sequentially by Levels II (prospective cohort), III (retrospective cohort), and IV (case series). For clinicians who already perform literature searches in this fashion, the table's increased accessibility will provide transparency to *The Journal*'s process of assigning LOEs.

Another important update is the table footnote that allows authors and editors to grade Level-I through IV studies upward due to "dramatic effect" or downward on the basis of "study quality, imprecision, or indirectness or because the absolute effect size is very small." The criteria in the table still guide the process, but this increased flexibility allows for more appropriate LOE assignments when the decision is not obvious. It is also important to note that, although this table is based on CEBM recommendations, we chose not to follow CEBM's policy of reserving the Level-I designation for systematic reviews. Systematic reviews are important, but we believe that high-quality original research merits an equally high LOE[5]. Additionally, *The Journal* recently published guidelines for the submission of systematic reviews and meta-analyses[6].

Lastly, the section on "Economic and Decision Analyses" was eliminated from the CEBM table, but we have elected to include these studies, now referred to as "Economic" in our table, as they are very important in orthopaedic surgery. These research methodologies are performed with use of preexisting data. The quality of these data and the type of analysis affect the LOE[7,8]. In probabilistic sensitivity analysis, each realization of a parameter is drawn from a prespecified distribution. In stochastic sensitivity analysis, the parameter values are selected from plausible ranges, for example, within the 95% confidence interval of the point estimate.

We view the LOE system as a guide to help determine the robustness of research quality but caution that a higher LOE does not necessarily reflect the clinical importance of a given study. The reader is still responsible for examining each article critically and deciding what constitutes the best external evidence for his or her specific clinical question. *The Journal* publishes studies based on quality of evidence and clinical importance and will continue to take both into account. ■

Robert G. Marx, MD
Associate Editor for Evidence-Based
Orthopaedics

Sean M. Wilson, BA

Marc F. Swiontkowski, MD
Editor-in-Chief

Levels of Evidence for Primary Research Question[1,2]

Study Type	Question	Level I	Level II	Level III	Level IV	Level V
Diagnostic—Investigating a diagnostic test	Is this (early detection) test worthwhile?	• Randomized controlled trial	• Prospective[3] cohort[4] study	• Retrospective[5] cohort[4] study • Case-control[6] study	• Case series	• Mechanism-based reasoning
	Is this diagnostic or monitoring test accurate?	• Testing of previously developed diagnostic criteria (consecutive patients with consistently applied reference standard and blinding)	• Development of diagnostic criteria (consecutive patients with consistently applied reference standard and blinding)	• Nonconsecutive patients • No consistently applied reference standard	• Poor or nonindependent reference standard	• Mechanism-based reasoning
Prognostic—Investigating the effect of a patient characteristic on the outcome of a disease	What is the natural history of the condition?	• Inception[3] cohort study (all patients enrolled at an early, uniform point in the course of their disease)	• Prospective[3] cohort[4] study (patients enrolled at different points in their disease) • Control arm of randomized trial	• Retrospective[5] cohort[4] study • Case-control[6] study	• Case series	• Mechanism-based reasoning
Therapeutic—Investigating the results of a treatment	Does this treatment help? What are the harms?[7]	• Randomized controlled trial	• Prospective[3] cohort[4] study • Observational study with dramatic effect	• Retrospective[5] cohort[4] study • Case-control[6] study	• Case series • Historically controlled study	• Mechanism-based reasoning
Economic	Does the intervention offer good value for dollars spent?	Computer simulation model (Monte Carlo simulation, Markov model) with inputs derived from Level-I studies, lifetime time duration, outcomes expressed in dollars per quality-adjusted life years (QALYs) and uncertainty examined using probabilistic sensitivity analyses	Computer simulation model (Monte Carlo simulation, Markov model) with inputs derived from Level-II studies, lifetime time duration, outcomes expressed in dollars per QALYs and uncertainty examined using probabilistic sensitivity analyses	Computer simulation model (Markov model) with inputs derived from Level-II studies, relevant time horizon, less than lifetime, outcomes expressed in dollars per QALYs and stochastic multilevel sensitivity analyses	Decision tree over the short time horizon with input data from original Level-II and III studies and uncertainty is examined by univariate sensitivity analyses	Decision tree over the short time horizon with input data informed by prior economic evaluation and uncertainty is examined by univariate sensitivity analyses

1. This chart was adapted from OCEBM Levels of Evidence Working Group, "The Oxford 2011 Levels of Evidence," Oxford Centre for Evidence-Based Medicine, http://www.cebm.net/ocebm-levels-of-evidence/. A glossary of terms can be found here: http://www.cebm.net/glossary/.

2. Level-I through IV studies may be graded downward on the basis of study quality, imprecision, indirectness, or inconsistency between studies or because the effect size is very small; these studies may be graded upward if there is a dramatic effect size. For example, a high-quality randomized controlled trial (RCT) should have ≥80% follow-up, blinding, and proper randomization. The Level of Evidence assigned to systematic reviews reflects the ranking of studies included in the review (i.e., a systematic review of Level-II studies is Level II). A complete assessment of the quality of individual studies requires critical appraisal of all aspects of study design.

3. Investigators formulated the study question before the first patient was enrolled.

4. In these studies, "cohort" refers to a nonrandomized comparative study. For therapeutic studies, patients treated one way (e.g., cemented hip prosthesis) are compared with those treated differently (e.g., cementless hip prosthesis).

5. Investigators formulated the study question after the first patient was enrolled.

6. Patients identified for the study on the basis of their outcome (e.g., failed total hip arthroplasty), called "cases," are compared with those who did not have the outcome (e.g., successful total hip arthroplasty), called "controls."

7. Sufficient numbers are required to rule out a common harm (affects >20% of participants). For long-term harms, follow-up duration must be sufficient.

References

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996 Jan 13;312(7023):71-2.

2. Wright JG, Swiontkowski MF, Heckman JD. Introducing levels of evidence to the journal. J Bone Joint Surg Am. 2003 Jan;85(1):1-3.

3. OCEBM Working Group. The Oxford 2011 Levels of Evidence. Oxford Centre for Evidence-Based Medicine. 2011. http://www.cebm.net/ocebm-levels-of-evidence/. Accessed 2014 Oct 23.

4. Howick J, Chalmers I, Glasziou P, Greenalgh T, Heneghan C, Liberati A, Moschetti I, Phillips B, Thornton H. Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence (Background Document). Oxford Centre for Evidence-Based Medicine. http://www.cebm.net/wp-content/uploads/2014/06/CEBM-Levels-of-Evidence-Background-Document-2.1.pdf. Accessed 2014 Oct 23.

5. Edwards TB. What is the value of a systematic review? J Shoulder Elbow Surg. 2014 Jan;23(1):1-2. Epub 2013 Nov 4.

6. Wright JG, Swiontkowski MF, Tolo VT. Meta-Analyses and Systematic Reviews: New Guidelines for JBJS. J Bone Joint Surg Am. 2012 Sep 5;94(17):1537.

7. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. JAMA. 1996 Oct 16;276(15):1253-8.

8. Caro JJ, Briggs AH, Siebert U, Kuntz KM; ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling good research practices—overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force—1. Value Health. 2012 Sep-Oct;15(6):796-803.