EBM Policy - Communication

A Modified Approach to Grading of Evidence

Laxmaiah Manchikanti, MD[1], Frank J.E. Falco, MD[2], Ramsin M. Benyamin, MD[3], Alan D. Kaye, MD, PhD[4], Mark V. Boswell, MD, PhD[5], and Joshua A. Hirsch, MD[6]

Health care organization, delivery, and outcomes of care are the basic requirements of and the focus of health care for research, practice, and policy (1,2). Since the inception of evidence-based medicine in 1992, the emphasis on evidence-based medicine, comparative effectiveness research, evidence synthesis and the development of guidelines continues to grow (1). The emphasis on evidence synthesis and development of guidelines has been the focus of the Institute of Medicine (IOM) leading to the re-engineering of its definition of clinical guidelines in 2011, along with the establishment of standards for systematic reviews (3,4). Clinical practice guidelines are statements that include recommendations intended to optimize patient care that are formed by a systematic review of evidence (3). The IOM also described a systematic review as a scientific investigation that focuses on a specific question and uses explicit, pre-specified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. Meta-analysis, in contrast to a systematic review, is the statistical pooling of data across studies to generate a summaries or pooled estimate of effects (5). However, challenges in implementing the IOM's systematic review standards and deficiencies of IOM's standards in guideline development have ALSO been described. It has been shown that the IOM failed to follow its own standards (6) and that even the Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Centers (EPCs) have found it difficult to implement the IOM systematic review standards (7). West et al (8), in the evidence report describing systems to rate the strength of scientific evidence, conceptualized that a continuum exists from rating the quality of the study to grading the strength of a body of evidence. Grading the strength of a body of evidence incorporates judgments of study quality, but it also includes how confident one is that a finding is true and whether others using different studies or different people have detected the same finding.

Even though the AHRQ has been the federal agency at the forefront of providing research support and policy guidance in health care services research in the United States (8), the National Health and Medical Research Council (NHMRC) of Australia (9) and multiple other organizations have also described similar instruments to assess the level of evidence of clinical studies. All of these organizations consider scientific data to be at the core of evidence-based approaches to clinical or public health issues, emphasizing that evidence needs to be carefully gathered and collated from a systematic literature review of each particular issue in question (8-12).

What constitutes strength of evidence has a range of definitions, all of which take into account the size, credibility, and robustness of the combined studies of a given topic (13-18). Questions remain, however, on the scientific validity of evidence and trustworthiness of clinical guidelines. In contrast, systems for grading the strength of individual articles continue to improve and may be somewhat superior and more consistent than grading the strength of a body of evidence (19-24). Improving the utility of evidence synthesis for public health policy depends on 4 R's (relevance, rigor, readability, and resources) (25). Criteria for rating the overall strength of a body of evidence depends on quality, quantity, and consistency (1,8).

From: [1]Pain Management Center of Paducah, Paducah, KY, and University of Louisville, Louisville, KY; [2]Mid Atlantic Spine & Pain Physicians, Newark, DE, and Temple University Hospital, Philadelphia, PA; [3]Millennium Pain Center, Bloomington, IL, and University of Illinois, Urbana-Champaign, IL; [4]LSU Health Science Center, New Orleans, LA; [5]Department of Anesthesiology and Perioperative Medicine, University of Louisville, Louisville, KY; and [6]Massachusetts General Hospital and Harvard Medical School, Boston. MA.

Author Affiliation and Disclaimer Information on P. E323.

Address Correspondence:
Laxmaiah Manchikanti, MD
2831 Lone Oak Road
Paducah, Kentucky 42003
E-mail: drlm@thepainmd.com

Manuscript received: 03-14-2014
Accepted for publication: 03-25-2014

The literature highlights multiple systems, some of which are extremely cumbersome to use and require substantial resources; whereas others are incomplete and non-comprehensive. West et al (8) reviewed 40 systems that addressed grading of the strength of a body of evidence, 34 were from sources other than AHRQ EPCs and 6 from EPCs. The evaluation criteria involved 3 domains of quality, quantity, and consistency, which are well established variables for characterizing how confidently one can conclude that a body of knowledge provides information.

The 4 qualitative strength of evidence systems have been commonly utilized in the literature, the first one being the quality of evidence criteria systems as shown in Table 1, used by multiple organizations from AHRQ, which,at the time, was known as AHCPR. This system has simple ratings from A to D with A being strong research-based evidence, based on multiple relevant and high quality scientific studies; B with moderate research-based evidence from one relevant high quality scientific study or multiple adequate scientific studies; C with limited research-based evidence of at least one adequate scientific study; and D with interpretation of information that did not meet inclusion criteria as research-based evidence (24).

The evidence from AHRQ, published in April 2002 (8), provided a description of the strength and consistency of evidence guidelines according to AHCPR. It is described in 2 parts, with Section 1 being type of evidence guidelines and Section 2 the strength and consistency of evidence guidelines (Table 2). Level I evidence consists of a metaanalysis of multiple well-designed controlled studies with consistent findings of type II, III, or IV. Level II evidence consists of at least one well-designed experimental study with evidence of type II, III, or IV, and for which findings are generally consistent. Evidence Level III is based on well-designed quasi-experimental studies such as nonrandomized controlled studies. Level IV is derived from well-designed non-experimental studies, whereas Level V is derived from case reports and clinical examples. Level IV indicates there is essentially little or no evidence or that it is based on case reports and clinical examples. Level V also indicates the existence of case reports or clinical examples. Strength and consistency at Level E describes practice recommended on the basis of the opinion of experts. This evidence criteria from 2002 shows the utilization of nonrandomized studies and differs from the criteria from 1994 (24), which utilized only randomized controlled trials.

Table 1. *Panel ratings of available evidence supporting guideline statements.*

A	Strong research-based evidence (multiple relevant and high-quality scientific studies).
B	Moderate research-based evidence (one relevant high-quality scientific study or multiple adequate scientific studies*).
C	Limited research-based evidence (at least one adequate scientific study* in patients with low back pain).
D	Panel interpretation of information that did not meet inclusion criteria as research-based evidence.

* Met minimal formal criteria for scientific methodology and relevance to population and specific method addressed in guideline statement.

Note: These criteria were derived from Bigos SJ et al. Acute low back problems in adults. Clinical Practice Guideline No.14, AHCPR Publication No. 95-0642. Rockville, Maryland. U.S.A., Agency for Health Care Policy and Research, Public Health Service, U.S., Department of Health and Human Services, December, pp. 1-60, 1994 (24). AHCPR was extinguished by Congress in 1995, changing AHCPR to AHRQ. Acute Low Back Pain Guidelines (24) provide a disclaimer "not for patient care."

Table 2. *Type of Evidence and Strength/Consistency of the Evidence Guidelines According to the AHCPR.*

Type of Evidence Guidelines (section one):
i. Meta-analysis of multiple well-designed controlled studies. ii. At least one well-designed experimental study. iii. Well-designed, quasi-experimental studies such as nonrandomized controlled, single group pre-post, cohorts, time series, or matched case-controlled studies. iv. Well-designed non-experimental studies, e.g., comparative, correlational, descriptive, case control. v. Case reports and clinical examples.
Strength and Consistency of Evidence Guidelines (section two):
There is evidence of type I or consistent findings from multiple studies of type II, III, or IV. There is evidence of type II, III, or IV, and findings are generally consistent. There is evidence of type II, III, or IV, but findings are inconsistent. There is little or no evidence, or there is type V evidence only. Panel consensus: Practice recommended on the basis of opinion of experts.

Agency for Healthcare Research and Quality. Systems to Rate the Strength of Scientific Evidence. Evidence Report/Technology Assessment #47. AHRQ Publication no. 02-EO16, April 2002 (8).

The evidence utilized by Cochrane reviews also is based on randomized trials (26,27). The evidence described falls into 5 levels based on the quality and outcome of studies (26,27), as shown in Table 3. This evidence is rated as strong, moderate, limited, conflicting, and no evidence.

Quality rating by U.S. Preventive Services Task Force (USPSTF) also has been popularized (12). In fact, 2 types of levels of evidence has been described as shown in Tables 4 and 5. Both have been extensively utilized. Both categories utilize evidence from randomized trials and nonrandomized studies (1,12,28,29).

As shown in Table 6, American Society of Interventional Pain Physicians (ASIPP), after convening a panel of experts and exploring (assessing) multiple systems arrived at somewhat different, easily understandable, evidence-based grading of evidence (12,26,27,30-36). This assessment may substitute meta-analysis. Meta-analysis or quantitative synthesis summarizes and helps by highlighting the comparisons that would be made, the outcomes that could be combined, and the study characteristics that should be considered when investigating any variation in effects. This variation in study characteristics is also known as heterogeneity (37). First, it should be determined whether a quantitative synthesis is at all possible and if so, whether it would be appropriate. Meta-analysis is not possible when the necessary data to perform meta-analysis cannot be obtained and may not be appropriate when the data are sparse or when studies are too heterogeneous to be sensibly combined. Essentially, meta-analysis is a tool to increase power, improve precision, and answer questions not posed by individual studies, and metaanalysis may also be used to settle controversies arising from conflicting studies or to generate new hypothesis (38). Meta-analysis, once it has been established that it is possible and appropriate, outcome measures, and a measure of association quantifying the effect of intervention, should be selected to describe the effectiveness. The choice of an effect measure is essential and important. Review should consider what type of outcome measure is being utilized, if the measure is interpretable

Table 3. *Levels of evidence.*

Strong evidence: consistent findings among multiple high quality RCTs.
Moderate evidence: consistent findings among multiple low quality RCTs or 1 high quality RCT.
Limited evidence: 1 low quality RCT.
Conflicting evidence: inconsistent findings among multiple trials.
No evidence: no RCTs.

Source: Staal JB, et al. Injection therapy for subacute and chronic low back pain: An updated Cochrane review. *Spine (Phila Pa 1976)* 2009; 34:49-59 (26) and van Tulder M, et al. Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. *Spine (Phila Pa 1976)* 2003; 28:1290-1299 (27).

Table 4. *Quality of evidence developed by USPSTF.*

I:	Evidence obtained from at least one properly randomized controlled trial
II-1:	Evidence obtained from well-designed controlled trials without randomization
II-2:	Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one center or research group
II-3:	Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled experiments (such as the results of the introduction of penicillin treatment in the 1940s) could also be regarded as this type of evidence
III:	Opinions of respected authorities, based on clinical experience descriptive studies and case reports or reports of expert committees

Adapted from the U.S. Preventive Services Task Force (USPSTF) (12).

Table 5. *Method for grading the overall strength of the evidence for an intervention.*

Grade	Definition
Good	Evidence includes consistent results from well-designed, well-conducted studies in representative populations that directly assess effects on health outcomes (at least 2 consistent, higher-quality RCTs).
Fair	Evidence is sufficient to determine effects on health outcomes, but the strength of the evidence is limited by the number, quality, size, or consistency of included studies; generalizability to routine practice; or indirect nature of the evidence on health outcomes (at least one higher-quality trial or study of diagnostic test accuracy of sufficient sample size; 2 or more higher-quality trials or studies of diagnostic test accuracy with some inconsistency; at least 2 consistent, lower-quality trials.
Limited or poor	Evidence is insufficient to assess effects on health outcomes because of limited number or power of studies, large and unexplained inconsistency between higher-quality trials, important flaws in trial design or conduct, gaps in the chain of evidence, or lack of information on important health outcomes.

Adapted from methods developed by U.S. Preventive Services Task Force (12,28,29).

Table 6. *ASIPP grading of evidence.*

Level I	Evidence obtained from multiple relevant high quality randomized controlled trials or Evidence obtained from multiple high quality diagnostic accuracy studies
Level II	Evidence obtained from at least one relevant high quality randomized controlled trial or multiple relevant moderate or low quality randomized controlled trials or Evidence obtained from at least one high quality diagnostic accuracy study or multiple moderate or low quality diagnostic accuracy studies
Level III	Evidence obtained from at least one relevant moderate or low quality randomized controlled trial study or Evidence obtained from at least one relevant high quality non-randomized trial or observational study with multiple moderate or low quality observational studies or Evidence obtained from at least one moderate quality diagnostic accuracy study in addition to low quality studies
Level IV	Evidence obtained from multiple moderate or low quality relevant observational studies or Evidence obtained from multiple relevant low quality diagnostic accuracy studies
Level V	Opinion or consensus of large group of clinicians and/or scientists.

At least 60% of studies in the direction of the objective being assessed.

by the clinicians using the review, if the measure likely to be considered across the studies is transferrable, and finally, if the measure has mathematic properties required to give a valid answer. Multiple systematic reviews in interventional pain management constantly misunderstand homogeneity and combine multiple heterogeneous studies together. Recently, Pinto et al (39) combined multiple heterogenous studies and utilized active control trials as placebo control, performing a metaanalysis that resulted in inaccurate conclusions (40). Staal et al (26) advised to refrain from performing a formal meta-analysis, if studies were considered clinically heterogeneous and/or studies did not report their results in a way that enabled them to perform statistical pooling. In those cases that were utilized the results were summarized according to a rating system with 5 levels of evidence (best evidence synthesis), based on the quality and the outcome of the studies as described above. Clinical homogeneity is a difficult issue and somewhat rare with interventional techniques. In a systematic review Manchikanti et al (41) assessing the efficacy of randomized controlled trials with 3 types of epidurals, caudal, interlaminar and transforaminal in lumbar disc herniation were unable to find clinical homogeneity among the 23 studies meeting inclusion criteria. However, utilizing best evidence synthesis with 5 levels of evidence they arrived at strong evidence for short-term efficacy and moderate evidence for long-term efficacy for all 3 approaches in managing chronic disc herniation. Thus, it is essential that homogeneity be clinically tested and if the studies meet clinical criteria, that statistical homogeneity be tested as well.

In developing grading of evidence, we have utilized the best evidence systems available (12,26,27,30-36). The evidence developed graded into 5 levels. Level I constitutes the highest level of evidence obtained from multiple relevant high quality, randomized, controlled trials, similar to Cochrane reviews as well as AHCPR grading of evidence.

Since multiple diagnostic accuracy studies are also assessed for evidence we have added Level I evidence for diagnostic accuracy with evidence obtained from multiple high quality diagnostic accuracy studies.

Level II describes the evidence obtained from at least one relevant high quality randomized controlled trial or multiple relevant moderate or low-quality randomized controlled trials. Level II also shows that the evidence was obtained from at least one high quality diagnostic accuracy study or multiple moderate or low quality diagnostic accuracy studies for diagnostic interventions.

Level III incorporates not only the evidence from randomized trials, but also from nonrandomized studies. To be considered Level III evidence there must be evidence obtained from at least one relevant moderate or low quality randomized controlled trial with multiple relevant observational studies or evidence obtained from at least one relevant high quality nonrandomized trial or observational study. Moreover, for diagnostic purposes, Level III describes evidence obtained from at least one moderate quality diagnostic accuracy study in addition to low quality studies. Level III is the interface between randomized and nonrandomized studies.

Level IV is purely nonrandomized studies with evidence obtained from multiple relevant moderate or low quality observational studies. Similarly, it also describes diagnostic accuracy evidence obtained from multiple relevant low quality diagnostic accuracy studies. Finally, Level V evidence is based on the consensus of large group of clinicians and/or scientists.

ASIPP's grading of evidence incorporates quality, quantity, and consistency. Quality is measured by the methodological assessment of relevant studies providing high quality, moderate quality, and low quality evidence which shows a study's design, conduct, and analysis that has minimized selection, measurement, and confounding biases. Quantity is measured by the magnitude of the treatment effect, the number of studies that have evaluated the given topic, and the overall sample size across all included studies. Finally, consistency is provided by various grading levels of evidence based on the quality and quantity of relevant quality studies.

At least 60% of the studies should show the effect in a single direction to provide the qualitative evidence.

For the purposes of grading of evidence, high quality is determined as meeting 8 criteria of 12 from Cochrane review or 32 criteria of 48 from ASIPP assessment grades for randomized trials or observational studies; with moderate quality being defined as scoring 4 to 7 on Cochrane review criteria or 20 to 31 on ASIPP IPM criteria for randomized trials and observational studies. Low quality evidence is described as studies scoring less than 4 on Cochrane review criteria and less than 20 on ASIPP assessment criteria with either randomized controlled trials or observational studies. Consequently, studies scoring less than 4 on Cochrane review criteria and less than 20 on ASIPP criteria are excluded.

This communication provides an analysis of the grading of evidence provides a means of grading evidence using common sense that quantifies the quality, quantity, and consistency.

Acknowledgments

The authors wish to thank Laurie Swick, BS, for manuscript review, and Tonie M. Hatton and Diane E. Neihoff, transcriptionists, for their assistance in preparation of this manuscript. We would like to thank the editorial board of Pain Physician for review and criticism in improving the manuscript.

Author Affiliations

Dr. Manchikanti is Medical Director of the Pain Management Center of Paducah, Paducah, KY, and Clinical Professor, Anesthesiology and Perioperative Medicine, University of Louisville, Louisville, KY.

Dr. Falco is Medical Director of Mid Atlantic Spine & Pain Physicians, Newark, DE; Director, Pain Medicine Fellowship Program, Temple University Hospital, Philadelphia, PA; and Adjunct Associate Professor, Department of PM&R, Temple University Medical School, Philadelphia, PA.

Dr. Benyamin is Medical Director, Millennium Pain Center, Bloomington, IL, and Clinical Assistant Professor of Surgery, College of Medicine, University of Illinois, Urbana-Champaign, IL.

Dr. Alan D. Kaye is Chairman and Professor, Department of Anesthesia, LSU Health Science Center, New Orleans, LA.

Dr. Boswell is Chairman, Department of Anesthesiology and Perioperative Medicine, University of Louisville, Louisville, KY.

Dr. Hirsch is Vice Chief of Interventional Care, Chief of Minimally Invasive Spine Surgery, Service Line Chief of Interventional Radiology, Director of Endovascular Neurosurgery and Neuroendovascular Program, Massachusetts General Hospital; and Associate Professor, Harvard Medical School, Boston, MA.

Conflict of Interest

Dr. Manchikanti certifies that he, or a member of his immediate family, has no commercial association (i.e., consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted manuscript.

Dr. Boswell certifies that he, or a member of his immediate family, has no commercial association (i.e., consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted manuscript.

Dr. Hirsch certifies that he, or a member of his immediate family, has no commercial association (i.e., consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted manuscript.

Dr. Falco is a consultant for St. Jude Medical Inc. and Joimax Inc.

Dr. Benyamin is a consultant and lecturer for Boston Scientific and Kimberly Clark.

Dr. Kaye is a speaker for Depomed, Inc.

References

1. Manchikanti L, Falco FJE, Singh V, Benyamin RM, Racz GB, Helm II S, Caraway DL, Calodney AK, Snook LT, Smith HS, Gupta S, Ward SP, Grider JS, Hirsch JA. An update of comprehensive evidence-based guidelines for interventional techniques of chronic spinal pain. Part I: Introduction and general considerations. *Pain Physician* 2013; 16:S1-S48.
2. Gagliardi AR, Brouwers MC. Integrating guideline development and implementation: Analysis of guideline development manual instructions for generating implementation advice. *Implement Sci* 2012; 7:67.
3. Graham R, Mancher M, Wolman DM, Greenfield S, Steinberg E (eds); Committee on Standards for Developing Trustworthy Clinical Practice Guidelines; Institute of Medicine. *Clinical Practice Guidelines We Can Trust*. The National Academies Press, Washington, DC, 2011.
4. Eden J, Levit L, Berg A, Morton S (eds); Committee on Standards for Systematic Reviews of Comparative Effectiveness Research; Institute of Medicine. *Finding What Works in Health Care. Standards for Systematic Reviews*. The National Academies Press, Washington, DC, 2011.
5. Pai M, McCulloch M, Gorman JD, Pai N, Enanoria W, Kennedy G, Tharyan P, Colford JM. Systematic reviews and meta-analyses: An illustrated, step-by-step guide. *Natl Med J India* 2004; 17:86-95.
6. Young BK, Greenberg PB. Are the Institute of Medicine's Trustworthiness Guidelines trustworthy? *R I Med J* 2013; 96:13-14.
7. Chang SM, Bass EB, Berkman N, Carey TS, Kane RL, Lau J, Ratichek S. Challenges in implementing The Institute of Medicine systematic review standards. *Syst Rev* 2013; 2:69.
8. West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, Lux L. *Systems to Rate the Strength of Scientific Evidence*, Evidence Report, Technology Assessment No. 47. AHRQ Publication No. 02-E016. Rockville, MD: Agency for Healthcare Research and Quality, 2002. www.thecre.com/pdf/ahrq-system-strength.pdf
9. How to use the evidence: Assessment and application of scientific evidence. National Health and Medical Research Council, Canberra, Commonwealth of Australia, 2000, pp 1-84.
10. Manchikanti L, Benyamin RM, Helm S, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 3: Systematic reviews and meta-analysis of randomized trials. *Pain Physician* 2009; 12:35-72.
11. Guyatt G, Gutterman D, Baumann MH, Addrizzo-Harris D, Hylek EM, Phillips B, Raskob G, Lewis SZ, Schünemann H. Grading strength of recommendations and quality of evidence in clinical guidelines. Report from an American College of Chest Physicians task force. *Chest* 2006; 129:174-181.
12. Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, Atkins D; Methods Work Group, Third US Preventive Services Task Force. Current methods of the US Preventive Services Task Force. *Am J Prevent Med* 2001; 20:21-35.
13. Ransohoff DF, Pignone M, Sox HC. How to decide whether a clinical practice guideline is trustworthy. *JAMA* 2013; 309:139-140.
14. Nagler EV, Ball S. Quality assessment of clinical practice guidelines: Who watches the watchmen? *Clin Endocrinol (Oxf)* 2013; 78:181-182.
15. Van Biesen W, van der Veer SN, Jager KJ, Fouque D, Wanner C, Vanholder R. What guidelines should or should not be: Implications for guideline production. *Nephrol Dial Transplant* 2013; 28:1980-1984.
16. Aguayo-Albasini JL, Flores-Pastor B, Soria-Aledo V. GRADE system: Classification of quality of evidence and strength of recommendation. *Cir Esp* 2014; 92:82-88.
17. Vandvik PO, Brandt L, Alonso-Coello P, Treweek S, Akl EA, Kristiansen A, Fog-Heen A, Agoritsas T, Montori VM, Guyatt G. Creating clinical practice guidelines we can trust, use, and share: A new era is imminent. *Chest* 2013; 144:381-389.
18. Lenzer J. Why we can't trust clinical guidelines. *BMJ* 2013; 346:f3830.
19. Watine J, Wils J, Augereau C. Clinical practice guidelines: Potential misconceptions of the GRADE approach. *J Clin Epidemiol* 2014; 67:7-9.
20. Furlan AD, Pennick V, Bombardier C, van Tulder M; Editorial Board, Cochrane Back Review Group. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine (Phila Pa 1976)* 2009; 34:1929-1941.
21. Manchikanti L, et al. Development of an interventional pain management specific instrument for methodologic quality assessment of nonrandomized studies of interventional techniques. *Pain Physician* 2014; 17:E291-E318.
22. Manchikanti L, et al. Assessment of methodologic quality of randomized trials of interventional techniques: Development of IPM specific instrument. *Pain Physician* 2014; 17:E263-E290.
23. Marin TJ, Furlan AD, Bombardier C, van Tulder M; Editorial Board of the Cochrane Back Review Group. Fifteen years of the Cochrane Back Review Group. *Spine (Phila Pa 1976)* 2013; 38:2057-2063.
24. Bigos SJ, Boyer OR, Braen GR, BrownK, Deyo R, Haldeman S, Hart JL, Johnson EW, Keller R, Kido D, Liang MH, Nelson RM, Nordin M, Owen BD, Pope MH, Schwartz RK, Stewart DH, Susman J, Triano JJ, Tripp LC, Turk DC, Watts C, Weinstein JN. Acute low back problems in adults. Clinical Practice Guideline No. 14, AHCPR Publication No. 95-0642. Rockville, Maryland. U.S.A., Agency for Health Care Policy and Research, Public Health Service, U.S., Department of Health and Human Services, December, 1994; pp 1-60.
25. Thomson H. Improving utility of evidence synthesis for healthy public policy: The three Rs (relevance, rigor, and readability [and resources]). *Am J Public Health* 2013; 103:e17-e23.
26. Staal JB, de Bie RA, de Vet HC, Hildebrandt J, Nelemans P. Injection therapy for subacute and chronic low back pain: An updated Cochrane review. *Spine (Phila Pa 1976)* 2009; 34:49-59.
27. van Tulder M, Furlan A, Bombardier C, Bouter L; Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. Spine *(Phila Pa 1976)* 2003; 28:1290-1299.
28. Chou R, Huffman L. *Guideline for the Evaluation and Management of Low Back Pain: Evidence Review*. American Pain Society, Glenview, IL, 2009.

    www.americanpainsociety.org/uploads/pdfs/LBPEvidRev.pdf
29. Manchikanti L, Abdi S, Atluri S, Benyamin RM, Boswell MV, Buenaventura RM, Bryce DA, Burks PA, Caraway DL, Calodney AK, Cash KA, Christo PJ, Cohen SP, Colson J, Conn A, Cordner HJ, Coubarous S, Datta S, Deer TR, Diwan SA, Falco FJE, Fellows B, Geffert SC,

Grider JS, Gupta S, Hameed H, Hameed M, Hansen H, Helm II S, Janata JW, Justiz R, Kaye AD, Lee M, Manchikanti KN, McManus CD, Onyewu O, Parr AT, Patel VB, Racz GB, Sehgal N, Sharma M, Simopoulos TT, Singh V, Smith HS, Snook LT, Swicegood J, Vallejo R, Ward SP, Wargo BW, Zhu J, Hirsch JA. An update of comprehensive evidence-based guidelines for interventional techniques of chronic spinal pain: Part II: Guidance and recommendations. *Pain Physician* 2013; 16:S49-S283.

30. Lohr KN, Carey TS. Assessing 'best evidence': Issues in grading the quality of studies for systematic reviews. *Joint Commission J Qual Improvement* 1999; 25:470-479.
31. Clarke M., Oxman AD. Cochrane Reviewer's Handbook 4.0. The Cochrane Collaboration, 1999.
32. National Health and Medical Research Council (NHMRC). How to Use the Evidence: Assessment and Application of Scientific Evidence. Canberra, Australia: NHMRC; 2000.
33. van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions. *Spine (Phila Pa 1976)* 1997; 22:2128-2156.
34. Guyatt GH, Cook DJ, Sackett DL, Eckman M, Pauker S. Grades of recommendation for antithrombotic agents. *Chest* 1998; 114:441S-444S.
35. Gross PA, Barrett TL, Dellinger EP, Krause PJ, Martone WJ, McGowan JE Jr, Sweet RL, Wenzel RP. Purpose of quality standards for infectious diseases. Infectious Diseases Society of America. *Clin Infect Dis* 1994; 18:421.
36. Gray JAM. *Evidence-Based Healthcare.* Churchill Livingstone, London, 1997.
37. Deeks J, Khan KS, Song F, Popay J, Nixon J, Kleijnen J. *Stage II – Conducting the review, Phase 7 – Data synthesis. In: Undertaking Systematic Reviews of Research on Effectiveness.* CRDs guidance for carrying out or commissioning reviews. CRD Report Number 4 (2nd), CRD Centre for Reviews and Dissemination, University of York, York, UK, March 2001. www.york.ac.uk/inst/crd/report4.htm
38. Deeks JJ, Higgins JPT, Altman DG (eds). Analysing and presenting results. Cochrane Handbook for Systematic Reviews of Interventions 4.2.6 [updated September 2006]; Section 8. In: *The Cochrane Library*, Issue 4, 2006. John Wiley & Sons, Ltd., Chichester, UK.
39. Pinto RZ, Maher CG, Ferreira ML, Hancock M, Oliveira VC, McLachlan AJ, Koes B, Ferreira PH. Epidural corticosteroid injections in the management of sciatica: A systematic review and meta-analysis. *Ann Intern Med* 2012; 157:865-877.
40. Manchikanti L, Falco FJE, Hirsch JA. Epidural corticosteroid injections in the management of sciatica. *Ann Intern Med* 2012; 157:865-877; online comment posted March 29, 2013.
41. Manchikanti L, Benyamin RM, Falco FJ, Kaye AD, Hirsch JA. Do epidural injections provide short- and long-term relief for lumbar disc herniation? A systematic review. *Clin Orthop Relat Res* 2014 Feb 11. [Epub ahead of print].