To The Editor:

Manchikanti et al published 2 lengthy critiques (1,2) of guideline efforts sponsored by the American Pain Society (APS) (3-5) that I led. I stand by the work conducted to develop the guideline, which adhered to high methodological standards; utilized an extensive, multidisciplinary external peer review process; and also underwent separate, independent peer review at a high-impact journal prior to final publication. I would like to take this opportunity to respond to the critiques.

In both articles, Manchikanti et al refer to a report by West et al (6) which they state presents Agency for Healthcare Research and Quality (AHRQ) criteria for evaluating the quality of systematic reviews, studies of diagnostic accuracy, and observational studies. This is an inaccurate description. The West et al study was never meant to present new quality ratings criteria. It was a systematic review commissioned by AHRQ to assess the usefulness of currently existing quality rating systems. The criteria presented by Manchikanti et al are simply domains that West et al used to evaluate the usefulness of other quality rating systems, and were not designed to be used to rate study quality themselves. They have never been endorsed by AHRQ or any other organization as quality rating instruments. In fact, West et al themselves recommend a number of existing quality rating systems for use:

Many systems covered most of the domains that are considered generally informative for assessing study quality. From this set, the authors identified 19 generic systems that fully address our key quality domains...In the authors’ judgment, those who plan to incorporate study quality into a systematic review, evidence report, or technology assessment can use one or more of these 19 systems as a starting point. (6)

Manchikanti et al suggest that a weighted system for scoring the criteria included in the Cochrane Back Review Group methods for rating randomized clinical trials is superior to the unweighted system we used. However, their weighted method is not supported by any empiric data or validation studies. In addition, a recent study provides empiric justification for classifying studies as “higher” or “lower” quality based on whether they met more or less than half of the (non-weighted) Cochrane criteria (7). The Cochrane Back Review Group itself does not recommend using a weighted system (8). 

Manchikanti et al state they do not understand how the Oxman and Guyatt instrument (9) for rating systematic reviews is scored. As described in the original Oxman and Guyatt article as well as in subsequent adaptations (10) (and also in the table included in the Manchikanti et al critiques), the final score is not based on simply adding the number of criteria that are met, as they seem to believe. Rather, it is based on an assessment of the type and severity of methodological flaws. For example, if a systematic review combined studies inappropriately, the scoring instructions are that it is likely to have major flaws (i.e. a score of 3 or less on a 1 to 7 scale). In addition, we did not, as suggested by Manchikanti et al, base any of our main conclusions on the efficacy of interventional procedures on previously published systematic reviews of interventional therapies. Rather, as described in our methods (3,4), our sole source of evidence for evaluating efficacy was randomized, placebo- or sham-controlled trials. Previously published systematic reviews were simply described to provide context and to help
identify and explore potential areas of discordance between our review and others’. Manchikanti et al seem to misunderstand the methods used by the US Preventive Services Task Force (USPSTF) to grade a body of evidence. I have extensive experience with USPSTF methods as I have led a number of USPSTF reviews (and am currently the lead investigator on the AHRQ-funded contract to conduct USPSTF evidence reviews). The tables provided in the 2 critiques suggest that the USPSTF grades a body of quality of evidence solely or primarily based on a study design hierarchy. This is incorrect. In fact, as described in the USPSTF methods, study design is only one of many factors used to grade evidence as “good,” “fair,” or “poor” (11). Other factors include the quality of studies, the number and size of studies, consistency between studies, and directness of evidence. Many of these issues—in particular inconsistency and sparse data—are essentially ignored by Manchikanti et al. Such an approach thus falls short of current standards established by the USPSTF, GRADE (12), the Cochrane Collaboration, and others for evaluating a body of evidence, and the growing literature on the importance of such factors (13-16).

Manchikanti et al appear surprised that quality assessments differed when they re-rated the studies included in the APS review. Yet those experienced in conducting systematic reviews know that applying quality criteria inevitably requires some subjective judgment and that some differences in quality ratings always occur (17). In addition, methods for operationalizing quality ratings criteria vary between authors and institutions (e.g., what is required for a study to meet the criteria for adequate allocation concealment, or acceptable loss to follow-up?). We followed standard methods for quality rating by clearly operationalizing each of the criteria we used, conducting dual independent quality ratings, and resolving discrepancies through discussion and consensus. The quality ratings were applied in a standard fashion across all of the studies included in our review (not just the studies of interventional therapies).

The above helps to explain minor discrepancies in quality ratings, but marked discrepancies are still troubling. Yet the numerous statements by Manchikanti et al that our quality ratings are “incorrect” do not hold up to scrutiny. As an example, take the very first randomized trial (by Mathews et al) from the Manchikanti et al therapeutic interventions critique in which there was a substantial difference between quality ratings from Manchikanti et al (8/11) and the APS (4/11) review. Manchikanti et al rates the randomization criterion as “yes” even though the trial never describes the method used to generate the random sequence (the criterion requires description of an appropriate randomization method, such as computer generated randomized numbers or a random numbers table). They rate the drop-out criterion as “yes” even though 21% of patients randomized to epidural injections and 41% randomized to control dropped out (the criterion requires less than 15% drop-outs overall and for drop-outs to be roughly equal between groups). They rate timing of outcome assessment criterion as “yes” even though the trial states that “assessments were made at least 4 times in the first 2 weeks” without a more precise description, and no results were reported for the first 2 weeks. Finally, they rate the intention-to-treat criterion as “yes” even though 9% (5/57) of the persons randomized to epidural steroid injections or control were not included in the analysis (the criterion requires no more than 5% of randomized patients to have been excluded). In other instances, Manchikanti et al seem to confuse issues related to external validity with quality (internal validity). For example, in the section on caudal epidural injections, they describe studies with short duration of follow-up, lack of placebo-control, or use of high volume injections as poor-quality even though none of these issues are associated with bias per se.

There are other inaccurate or misleading critiques. Again looking at caudal epidural interventions, the first intervention discussed in the therapeutic interventions critique, Manchikanti et al states that we should have excluded a trial by Manchikanti et al (18) since it only addressed adhesiolysis. In fact, this trial had 3 arms, one of which evaluated “catheterization without adhesiolysis, followed by injection of local anesthetic, normal saline, and steroid.” This certainly sounds like a trial comparing epidural injection to 2 different types of adhesiolysis. Manchikanti et al questions inclusion of a trial with follow-up of 24 weeks (19), because the duration of follow-up was too short, even though many of the other trials they included only followed patients for 3 months or less; in addition the APS exclusion criteria did not require longer follow-up. Manchikanti et al also critique us for excluding a foreign-language article (20) as well as 2 active-controlled trials (21,22), even though our methods clearly state that we excluded foreign-language articles and did not independently abstract or rate ac-
tive-controlled trials already included in other higher-quality systematic reviews (since these were not the focus of our review). A basic principle of conducting systematic reviews is to follow the methods laid out for including and excluding studies, in order to remove subjective bias in selection of studies.

It is not feasible to respond to every critique in the articles by Manchikanti et al. However, radiofrequency denervation provides a good example of how the American Society of Interventional Pain Physicians (ASIPP) approach falls short. Of the 6 trials of radiofrequency denervation versus sham denervation included by Manchikanti et al, they eliminated 4 trials (23-26) due to alleged technical or methodological flaws (though we disagree with some of these exclusions, we will accept them for the purposes of this discussion). This leaves 2 trials with a total of 100 patients. Of the 2 trials, one by Nath et al (27) was rated highly by both Manchikanti et al and APS, but it had a serious methodological problem. Specifically, there was a statistically significant difference in important baseline characteristics, including generalized, back, and leg pain scores. This is a serious flaw because it indicates failure of randomization (due to chance, a flawed randomization method, or intentional subversion), and because the sham control group (which had higher baseline scores) had greater potential to experience improvement from baseline (the main outcome analyzed in the trial). Manchikanti et al seem to downplay this issue and then go on to contest our statement that final scores in the active and sham radiofrequency denervation groups were identical. However, the results speak for themselves: generalized pain 4.1 vs. 4.0, back pain 3.9 vs. 3.7, and leg pain 2.7 vs. 2.6. It is difficult to see how the Nath et al trial could be taken as reliable evidence that radiofrequency denervation is effective. This leaves one (n=60) trial by Tekin et al (28) of borderline methodological quality (rated 6/11 by Manchikanti et al and 5/11 by APS), with small effects on pain (0.2 to 0.8 points on a 0 to 10 scale at 6 months) and function (4 points on the 0 to 100 Oswestry Disability Index at 6 months). In other words, a single, small trial with some methodological flaws and marginally clinically meaningful benefits. Such sparse and underwhelming evidence is obviously a poor basis with which to guide clinical decision-making.

Regarding diagnostic therapies, Manchikanti et al do not appear to understand the major issue. That is, it is impossible to calculate the diagnostic accuracy of invasive diagnostic techniques because there is no accepted reference standard for discogenic or facetogenic low back pain. Imaging studies are an inappropriate reference standard because they do not correlate with the presence of or severity of back symptoms, and no other reliable reference standard is available. Therefore, studies that attempt to characterize diagnostic accuracy are not interpretable, and it is necessary to move up on the evidence hierarchy to studies that evaluate effects of invasive diagnostic (versus no invasive testing) on clinical outcomes (29,30). In fact, Manchikanti et al do not seem to dispute that there is little evidence showing that provocative discography or facet joint blocks improve clinical outcomes. Furthermore, a recent trial found that patients who underwent radiofrequency denervation based on clinical findings alone experienced a higher rate of clinical success than those who underwent radiofrequency denervation only if they had one or 2 positive facet joint blocks (31).

Manchikanti et al also call into question the integrity of the APS review and guideline, which is unjustified and uncalled for. I would like to clarify several points. First, AHRQ is not my employer. Although I have received funding from AHRQ to conduct research, my employer is Oregon Health & Science University, as stated in the guideline documents. Suggesting that my employer is AHRQ would be like suggesting that any researcher who has received funding from the National Institutes of Health (NIH) is an employee of the NIH. Furthermore, there is no evidence that conducting publicly-funded work is associated with bias. In fact, bias related to commercial or other financial interests (such as from performing a procedure) is much more likely to be a problem (32-34). Second, it is incorrect to assert, as Manchikanti et al do, that any experts invited to participate on the APS low back pain guideline panel withdrew their support. Although ASIPP requested that one expert whom they had nominated for the panel not list himself as an author on the guideline article, he did not withdraw from the panel and in fact fully agreed to be listed as a member. Finally, Manchikanti et al suggest that professional society sponsorship was not fully disclosed. In fact, as stated in the guideline and accompanying articles, the APS was the sole sponsor of the surgery and interventional therapies guideline (3-5). Although the American College of Physicians were a co-sponsor of an earlier guideline (35) on primary care evaluation and management of low back pain (thus the original press release), they were not.
References
1. Manchikanti L, Datta S, Derby R, Wolf-er LR, Benyamin RM, Hirsch JA. A criti-
cal review of the American Pain Society
clinical practice guidelines for interven-
tional techniques: Part 1. Diagnostic in-
terventions. Pain Physician 2010; 13:
E141-E174.
2. Manchikanti L, Datta S, Gupta S, Mung-
lani R, Bryce DA, Ward SP, Benyamin
RM, Sharma ML, Helm II S, Fellows B,
Hirsch JA. A critical review of the Ameri-
can Pain Society clinical practice guide-
lines for interventional techniques: Part
2. Therapeutic interventions. Pain Phy-
R. Nonsurgical interventional therapies
for low back pain: A review of the evi-
dence for an American Pain Society cli-
nical practice guideline. Spine (Phila
4. Chou R, Huffman LH. Guideline for the
evaluation and management of low back
pain: Evidence review. American Pain
Society. www.ampainsoc.org/pub/pdf/
5. Chou R, Loeser JD, Owens DK, Rosen-
quist RW, Atlas SJ, Baiden J, Carragee
EJ, Grabois M, Murphy DR, Resnick DK,
Stanos SP, Shaffer WW, Wall EM; Ameri-
can Pain Society Low Back Pain Guide-
line Panel. Interventional therapies,
surgery and interdisciplinary rehabili-
tation for low back pain: An evidence-
based clinical practice guideline from
the American Pain Society. Spine (Phila
6. West S, King V, Carey TS, Lohr KN, McK-
oy N, Sutton SF, Lux L. Systems to Rate
the Strength of Scientific Evidence.
Rockville, MD: Evidence Report/Techn-
ology Assessment No. 47. The Agen-
cy for Healthcare Research and Qual-
ity; Rockville: MD; AHRQ Publication
No. 02-E016; April 2002.
7. van Tulder MW, Suttorp MJ, Morton S,
Bouter LM, Shekelle P. Empirical evi-
dence of an association between inter-
val validity and effect size in random-
ized controlled trials of low-back pain.
Spine (Phila Pa 1976) 2009; 34:1685-
1692.
8. Furlan AD, Pennick V, Bombardier C, van
Tulder M. 2009 updated method guide-
dines for systematic reviews in the Co-
chrane Back Review Group. Spine (Ph-
9. Oxman AD, Guyatt GH. Validation of an
index of the quality of review articles. J
10. Furlan AD, Clarke J, Emslie R, Sinclair S,
Irvin E, Bombardier C. A critical review
of reviews on the treatment of chronic
2001; 26:E155-E162.
11. Harris RP, Helfand M, Woolf SH, Lohr
KN, Mulrow CD, Teutsch SM, Atkins D;
Methods Work Group, Third US Preven-
tive Services Task Force. Current meth-
ods of the U.S. Preventive Services Task
12. Guyatt GH, Oxman AD, Vist GE, Kunz R,
Falck-Ytter Y, Alonso-Cuello P, Schüne-
mann HJ; GRADE Working Group. Rating
quality of evidence and strength of rec-
ommendations: GRADE: an emerging
consensus on rating quality of evidence
and strength of recommendations BMJ
2008; 336:924-926.
13. Ioannidis JPA. Contradicted and initial-
ly stronger effects in highly cited clini-
14. Ioannidis JPA. Why most published re-
search findings are false. PLoS Med
15. Ioannidis JPA. Why most discovered true
associations are inflated. Epidemiology
16. Rerkasem K, Rothwell PM. Meta-analy-
sis of small randomized controlled trials
in surgery may be unreliable. Br J Surg
2010; 97:466-469.
hazards of scoring the quality of clinical
trials for meta-analysis. JAMA 1999;
282:1054-1060.
18. Manchikanti L, Rivera JJ, Pampati V,
Damron KS, McManus CD, Brandon DF,
Wilson SR. One day lumbar epidural ad-
hesiolysis and hypertonic saline neu-
rolysis in treatment of chronic low back
pain: A randomized, double-blind trial.
19. Ackerman WE, 3rd, Ahmad M. The effi-
cacy of lumbar epidural steroid injec-
tions in patients with lumbar disc herni-
ations. Anesth Analg 2007; 104:1217-
1222.
20. Hesla PE, Brievik H. Epidural analgesia
and epidural steroid injection for treat-
ment of chronic low back pain and sci-
atica. Tidsskr Nor Laegeforen 1979;
99:936-939.
21. McGregor AH, Anjarwalla NK, Stambach
T. Does the method of injection alter the
outcome of epidural injections? J Spinal
Disord 2001; 14:507-510.
22. Revel M, Auleley G, Alaoui S. Forceful
epidural injections for the treatment of
lumbosacral pain with post-operative
lumbar spinal fibrosis. Rev Rheum Eng
23. J Gallagher, PL Petriccione di Vadi, JR
Wedley, W Hamann, P Ryan, I Chikanza,
B Kirkham, R Price, MS Watson, R Gra-
ham, and S Wood. Radiofrequency fac-
et joint denervation in the treatment of
low back pain: A prospective controlled
double-blind study to assess its efficac-
Y, Rossignol M. Radiofrequency fac-
et joint denervation in the treatment of
low back pain: A placebo-controlled clini-
cal trial to assess efficacy. Spine (Phila
25. van Kleef M, Barendse G, Kessels A,
Voets H, Weber W, de Lange S. Random-
ized trial of radiofrequency lumbar facet
denervation for chronic low back pain.
Spine (Phila Pa 1976) 1999; 24:1937-
1942.
26. van Wijk R, Geurts J, Wynne H, Hammink
E, Buskens E, Lousberg R, Knape JT,
Groen GJ. Radiofrequency denervation
of lumbar facet joints in the treatmen-

Roger Chou, MD
Oregon Health & Science University
3181 S.W. Sam Jackson Park Rd.
Portland, Oregon 97239-3098
E-mail: chour@ohsu.edu

Pain Physician: January/February 2011; 14:E69-E86
Letters to the Editor


