Comparative Effectiveness Review Disposition of Comments Report

Research Review Title: ***Noninvasive Treatments for Low Back Pain***

Draft review available for public comment from April 27, 2015 to May 18, 2015.

Research Review Citation: Chou R, Deyo R, Friedly J, Skelly A, Hashimoto R, Weimer M, Fu R, Dana T, Kraegel P, Griffin J, Grusing S, Brodt E. Noninvasive Treatments for Low Back Pain. Comparative Effectiveness Review No. 169. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2012-00014-I.) AHRQ Publication No. 16-EHC004-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2016. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or E-mail. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Source: https://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=2178
Published Online: February 29, 2016

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #1	The report is very well-done; it is comprehensive, thorough, and methodologically sound. The conclusions appear appropriate. The key questions are appropriate and explicitly stated.	Thank you for the comment.
General	TEP Reviewer #1	Structured Abstract Results section page vi (page 6 on the pdf file): (a) The authors refer to the size of benefits in terms of points on a 0-100 VAS. They might consider adding the number of points on a 0-10 NRS, since currently these are more commonly used to measure pain outcomes in clinical trials.	We added the equivalent for small effects on a numerical rating scale (0.5 to 1.0 points on a numerical rating scale)
General	TEP Reviewer #1	(b) When they say effects on function were “not observed” (line 35 on page vi), it might be good to clarify whether that means there were no effects on function, or whether the study did not measure function.	We revised to be clearer that in some cases there were positive effects on pain but not on function, and that fewer studies measured function than pain.
General	TEP Reviewer #1	(c) It would also be good to clarify the statement “benefits were mostly observed at short-term follow-up” (line 36 on page vi). Does this mean they disappeared by longer-term follow-ups or that the studies did not include longer-term follow-ups?	We revised to state that benefits were mostly *measured* at short-term follow-up.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #1	(d) I am confused by lines 45-48 on page vi in the part beginning "spinal manipulation…" Do the authors mean that spinal manipulation was as effective as other active interventions, defined as ultrasound and TENS, and ultrasound and TENS were no more effective than sham interventions? Or should the left parenthesis preceding ultrasound have been a semi-colon?	We deleted the parenthetical sentence regarding ultrasound and TENS; this sentence was supposed to focus on interventions with some evidence of benefit.
General	TEP Reviewer #1	(e) There seem to be some words missing in the line that ends "associated small effects (SOE: low)" (line 51).	We corrected the typo ("…associated with small effects.")
General	TEP Reviewer #1	(f) Finally, the last sentence in that section (lines 51-54) is unclear. Do the authors mean that among the few trials, there were no clear differences in effects?	We revised to be clearer that the trials generally found no clear differences in effects.
General	TEP Reviewer #1	Executive Summary: Page ES-2 (or page 12 of 923), lines 11-12: I suggest not including interdisciplinary rehabilitation as an example of psychological therapies. Interdisciplinary rehabilitation includes psychological therapies, but also includes other therapies, such as physical and occupational therapy and medication management. This suggestion also applies to the text on page 68 of 923, lines 10-11.	We revised as suggested (we changed the term from interdisciplinary rehabilitation to multidisciplinary rehabilitation to be consistent with the rest of the report).

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #1	Very minor point, but on page 76/923 of the pdf, line 48 (page 10 of the full review) and below in that section, most VAS and NRS scales are 0-100 and thus 101 points rather than 100 possible points. In other places, the authors also refer to a 100-point VAS (e.g., line 33, page 58 of the full review).	We revised to refer to 0 to 100 point VAS and 0 to 10 point NRS when first discussed in the Methods; in subsequent places in the report we referred to as a "100 point VAS" as this is common usage and we think well understood.
General	TEP Reviewer #2	This is a methodologically sound report that covers a very large number of interventions and comparisons. It is an update of the previous AHRQ evidence report also used in part for a clinical practice guideline by the American College of Physicians and Pain Society. Methods, interventions, outcomes and reporting format are similar to the prior report.	Thank you for the comment.
General	TEP Reviewer #2	Authors used existing systematic reviews to update a previous evidence report and then identified and described new information from eligible studies. I believe this is a scientifically sound and programmatically sensible solution to the large amount of information available.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #2	Authors used previously established scale scores to assess their main outcomes (pain and function) and measures of effect size. The majority of studies reported pain (rather than function). They attempted to categorize any effects as “small” “medium” and “large” where significant differences existed and the evidence was of sufficient strength to warrant. They provide numerous references to justify their choices and reporting methods. See my comments below regarding reporting on effect sizes and MID.	Thank you for the comment.
General	TEP Reviewer #2	The authors categorize back pain by duration and “location”: i.e. Acute, Subacute and Chronic; “Isolated Low Back Pain”, Radicular, Spinal Stenosis. They further often “lump” Acute with Subacute and Isolated vs. Radicular/Spinal Stenosis. ACP/APS essentially used a similar categorization in their prior review (though categorized as Acute vs. Subacute or Chronic) and most clinicians approach patients with low back pain in this broad fashion. This is reasonable.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #2	The main messages are that low back pain is very common and can be disabling. Back pain may be clinically conceptualized as acute vs. subacute vs chronic and radicular vs. nonradicular vs. spinal stenosis. Many treatment options exist. None are dramatically better than others or placebo, though some have important benefits and some have important harms and likely costs. There are some notable consistent findings from the prior report and some notable changes.	Thank you for the comment. This is the reviewer's summary of findings and does not require changes to the report.
General	TEP Reviewer #2	Main effectiveness outcomes are "pain" (using a variety of scale scores, with mean changes, percent change from baseline/control, standard effect size) and "function" (similar ways of reporting), short term and long term, then corresponding harms. This is a reasonable way to break out effectiveness outcomes.	Thank you for the comment.
General	TEP Reviewer #2	Harms outcomes are often not reported especially for nonpharmacologic interventions though conceptually most are unlikely to have frequent/serious harms.	Thank you for the comment. Suboptimal harms reporting was noted in the Results for various interventions as well as in the Discussion.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #2	Additional findings: The large majority of patients with acute back pain (whether radicular or not) get better regardless of treatments received (or not). For the large majority of interventions there are few differences versus control (placebo, sham, wait list, “active comparator”) and where they exist effect sizes are generally fairly small. In individuals with chronic back pain the prognosis is worse with a large number of individuals continuing to have pain and dysfunction and with little effective therapy-some of which may have considerable harms and costs.	Thank you for the comment. This is the reviewer’s summary of findings and does not require changes to the report.
General	TEP Reviewer #2	The report is almost entirely “text based”. Tables are very data dense. Few to no figures are provided. The authors attempted to improve clarity of reading by using bulleted summary messages and writing in a clear, concise and standardized fashion. Nonetheless, the amount of data and presentation of results is very difficult to get through the more than 150 comparisons and pull out key messages. This limits clinical usefulness and a point that needs improvement.	We summarized the key findings in bulleted points as well as in the summary of evidence tables, and added additional tables summarizing main findings for acute, chronic, and radicular low back pain.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #2	Summary figures would help and additional organization changes.	We summarized the key findings in bulleted points as well as in the summary of evidence tables, and added additional tables summarizing main findings for acute, chronic, and radicular low back pain.
General	TEP Reviewer #2	Reorganize key comparisons according to back pain characteristics: (given that the full report is written according to intervention I suggest that the reorganization is a “superheader”…and not redo the whole report-rather make some type of organization structure that allows for evaluation by pain categorization in addition to just the list of interventions and their comparators. i. Acute or Subacute vs. Chronic (then subheading radicular vs. nonradicular) 1. Pharmacologic a. Effectiveness i. Harms b. Comparative Effectiveness i. Comparative Harms 2. Nonpharm a. Effectiveness i. Harms b. Comparative Effectiveness i. Comparative Harms	We organized the report according to the Key Questions (pharmacological interventions for Key Question 1 and non-pharmacological for Key Question 2). For each Key Question we presented results for each intervention, organized by comparisons versus placebo/sham/no treatment for acute/subacute low back pain, then chronic, then radicular; we then did the same versus other interventions. We think this is a reasonable way to organize the report; we did add additional tables summarizing results across interventions for acute, chronic, and radicular low back pain.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #2	ii. Display baseline pain/function severity and change from baseline for intervention and control in overview tables (particularly important when intervention is likely to have few harms (patients want something and change from baseline (if harms are low) may be a not unreasonable way for clinicians and patients to look at results.	As stated in the Methods, assessment of outcomes was based on between-group differences at follow-up. We do not think it is appropriate to determine effectiveness based on change from baseline without a comparison group, given that interventions are known to have strong placebo effects in patients with pain and effects of natural history (particularly for acute and subacute pain), regression to the mean, attentional and other non-specific effects.
General	TEP Reviewer #2	Clinicians and patients also think in terms of pain presentation: is it acute or chronic and is it radicular or not…then they consider treatment options. The prior guideline was written in this format.	The large number of interventions, comparisons, and low back pain subtypes (acute, subacute chronic; non-radicular, radicular, spinal stenosis) represent a challenge in presentation of findings. We think there are multiple ways to organize the results but feel that organizing as we did (by intervention and comparison, with subheadings for particular subtypes when data were available) was an appropriate format.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #2	The report is almost entirely text based. While writing is generally clear on an individual item it is very difficult to get main points (even with the key point bullets which are good). I attempted to use this document with a patient having subacute pain. It was very difficult to identify interventions and outcomes rapidly in a way to implement. The summary effect size tables are very laborious to get through but necessary for the highly interested. Additional organization along pain presentation and with “speed sheet” referral graphics (e.g. various shades of grey or arrows showing effect size and SOE would aid in delivering stake holders usable “quick main messages”.	We added summary Tables summarizing the main findings across interventions for acute/subacute, chronic, and radicular low back pain. We used standard text descriptors; it might be appropriate for groups translating the report into practice guidelines to use more graphical representations.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #2	i. Consider adding table/figures (by pain categorization) that provides a better "visual" to display results (perhaps separately for pain and function (though the authors can "play" around with different formats. E.g. use various shades of grey to indicate benefit (or not), or magnitude/strength of evidence of effect with size, width, # arrows in a given direction to display magnitude of benefit, strength of evidence (perhaps include the # studies/patients etc.). The past ACP guideline used a table with headers stratified by pain duration and then interventions and listed SOE and Net Benefits. These authors could at least create a table: 1) Acute or Subacute 2) Chronic…then type of intervention, level of evidence, benefit, harm…with some assessment of magnitude of benefit and harm for each (and SOE for benefit and harm).	We added tables summarizing main findings across interventions for Acute/subacute, chronic, and radicular low back pain.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #2	iii. Greater consistency in magnitude of effect. For some scales a 5-10% change is at least small (e.g. 5-10 point score change on 100 point scale) but the SMD for small was 0.2 to 0.5 (likely a bigger effect size than 5-10%. References are provided but are these really consistent across measurements for small, medium and large-including when incorporating baseline pain/function. 1. This is particularly important depending on patients' baseline measure. E.g. a 5 point change may be very large in an individual with a baseline score of 8 but small in a person with a baseline score of 95.	We relied on pooled results are reported in published systematic reviews. In some cases they were reported as WMD's and in others as SMD's. We described how we categorized effects based on mean differences and standardized mean differences. We agree that the clinical relevance of a similar improvement in pain score might vary depending on baseline symptom severity. We revised the applicability section of the Discussion to be note that magnitude of effects might vary according to baseline severity and that the clinical relevance of similar changes in pain or function scores might vary depending on the baseline score.
General	TEP Reviewer #2	The question is what to do about the group with no benefit versus control but likely very small harms…especially for patients with chronic pain? Both the intervention and control may have "important" benefits from baseline but the difference between intervention and control was either very small or not significant.	Thank you for the comment. The purpose of the report is to summarize the evidence, not provide practice guidelines. Clinicians and policymakers will use information on benefits and harms to make treatment decisions.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #2	a. I request that the authors provide additional information regarding baseline pain and function scores and change from baseline for both intervention and controls-especially for areas where interventions may have few harms or costs. i. This is important to determine the generalizability of findings to others with pain, the potential percent change from baseline and thus magnitude of effect for intervention and control.	As described in the Methods, we focused on between-group differences in outcome measures. We do not think it is advisable to focus on within-intervention changes from baseline given known strong placebo and other non-specific effects. We reported outcomes as reported in the systematic reviews and primary trials not included in systematic reviews; some studies reported the proportion of patients with a 30% or 50% improvement in pain or function (or other dichotomized outcome) but few studies reported the percent absolute improvement in these outcomes.
General	TEP Reviewer #2	b. The authors note the problems with the literature (few studies report % with meaningful change, lack of blinding or placebo and the large number of interventions many with only 1-2 studies or /moderate to high risk of bias thus precluding sufficient evidence. i. For some of these it would be helpful to know pain/dysfunction severity at baseline and the changes from baseline for both intervention and control.	We revised the Applicability section to note that most trials enrolled patients with at least moderate severity symptoms at baseline. We also revised to note that similar absolute effects on an outcome might vary in clinical relevance depending on baseline symptom severity.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #2	c. Much of this report relies on assessing change in scale scores and then determining a “magnitude of effect” and whether the results are “clinically meaningful” and/or achieve a Minimally Important Difference. This is an important when only “mean changes versus control” are reported (rather than change from baseline or “% responding”. A review of the literature regarding “minimally important differences” especially in the pain literature indicates that “MID” are often determined as “MND” minimally noticeable difference as assessed by mean change from BASELINE to end of study rather than change vs. comparator. The authors report on mean effect sizes vs. control and provide no information about change from baseline for intervention and control (it is possible that both the intervention and the control had moderate to large effect sizes but that the differences is small and or not significant).	As the reviewer notes, minimally clinically important differences are typically determined based on within patient changes on an outcome. The thresholds for minimally clinically important differences are then applied to interpret the comparative effectiveness of interventions based on between-group differences. If both groups improve to a moderate or large degree but there is no difference in the degree of improvement, there is no difference in effectiveness. This is the standard method for reporting/assessing minimum clinically important differences.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #2	d. To my knowledge MID are NOT reliably assessed according to "change vs. control though I believe they should. Instead MID and MND are typically derived from change from baseline at a given "disease severity"-often using an anchor method. The authors though (as to many/most (?) reviewers/trialists focus reporting on change vs. control when reporting mean differences (particularly if % achieving MND are not reported).	As the reviewer notes, minimally clinically important differences are typically determined based on within patient changes on an outcome. The thresholds for minimally clinically important differences are then applied to interpret the comparative effectiveness of interventions based on between-group differences. If both groups improve to a moderate or large degree but there is no difference in the degree of improvement, there is no difference in effectiveness. This is the standard method for reporting/assessing minimum clinically important differences.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #2	e. I request that authors provide information regarding change from baseline for both the intervention and the comparator (placebo, sham, "active comparator" especially for treatments where there are likely very few harms or costs (e.g. nonsteroidal anti-inflammatory agents, massage, TENS, ultrasound etc…). This information will be useful for clinicians and patients and can serve assist in clinical practice guidelines. Some guidance is provided in a recently published review: *Johnston et al. Patient-reported outcomes in meta-analyses – Part 2: methods for improving interpretability for decision-makers. Health and Quality of Life Outcomes 2013, 11:211. http://www.hqlo.com/content/11/1/211*	As is standard, we focused on between group differences to assess comparative effectiveness. Baseline scores are reported in the tables for primary trials not included in systematic reviews. It would not be feasible to go back to all the RCTs included in systematic reviews to abstract information about baseline pain and function; however, we revised the Discussion to note that estimates of effectiveness may vary depending on baseline scores and that most trials enrolled patients with at least moderate severity symptoms.
General	TEP Reviewer #3	The "Executive Summary" section is 48 pages in length plus 7 more pages of references cited. This is notably shorter than the full report, which includes 104 pages of text, 103 pages of tables, 36 pages of references, and hundreds more of appendices. However, a 48-page Executive Summary is still lengthy, and if it is truly desired to have this readable in a short period, then I would encourage the authors to further condense the Executive Summary section.	We revised to shorten the Executive Summary substantially.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #3	This is a massive body of literature to review and I generally concur with the findings of the authors.	Thank you for the comment.
General	TEP Reviewer #3	An over-arching question about how this review was conducted is the rationale for not performing new systematic reviews of all the relevant studies. Rather, the approach taken in the current assessment was to include existing systematic reviews that were judged to be of sufficient rigor and then examine clinical studies performed since the prior systematic review. The question is whether the current approach would provide an equivalent result as actually conducting a new systematic review or does the current approach intrinsically have a tendency for some sort of bias. I am not advocating that a new systematic review should or needed to be done, but rather it would be helpful to provide upfront (and not just later in the Discussion section) the rationale & justification for the approach utilized.	The rationale for including systematic reviews (the large number of interventions and literally hundreds of primary studies) is provided in the Methods, in the Study Designs section. We followed AHRQ methods guidance in selecting and incorporating prior systematic reviews.
General	TEP Reviewer #4	This is difficult area due to the large number of trials, many of which are not adequately controlled, and the wide range of clinical presentations, severity and responses.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #5	Yes, the report has clear relevance and value for evaluating the evidence for clinical effectiveness of treatments for back pain. The target population is clearly described and the key questions are explicit and appropriate.	Thank you for the comment.
General	Peer Reviewer #1	This report is very thorough and technically fine, but I think it will be challenging and a bit overwhelming for clinicians to know how to integrate this information into their care. There are so many comparisons - even within therapies - that this is more like a card catalog than a clinically digestible report.	We revised the Executive Summary substantially to shorten the Results and added summary Tables summarizing main findings across interventions for acute, chronic, and radicular low back pain.
General	Peer Reviewer #2	I think the report is correctly targeted and the correct issues are addressed.	Thank you for the comment.
General	Peer Reviewer #2	From an overall usability standpoint I hope ACP/APS will convene a group of multidisciplinary experts to work with the evidence to create clear clinical recommendations. While I have no trouble doing that I work with evidence and guideline development. I think few straight clinicians would wade though this. The final discussion at the end is a great help however.	ACP nominated this topic to AHRQ for a systematic review and intends to use the review to update its guidelines.

Section	Commentator & Affiliation	Comment	Response
General	Peer Reviewer #2	IF the table beginning on page 191 is meant to do that I think it probably may need some revision to highlight the most important clinical issues. For instance, the use of steroids for radiculopathy is nearly universal in many places and yet it is not helpful. This is a very important point compared to some of the other medication issues you very appropriately reviewed.	We added Tables summarizing main findings for acute, chronic, and radicular low back pain, including the lack of effect of systemic corticosteroids for radiculopathy. As noted by the reviewer, this finding is already highlighted in the Discussion.
General	Peer Reviewer #2	IN any case my comments here are directed more toward the eventual use of the product rather than the very high quality you have produced.	Thank you for the comment.
General	Peer Reviewer #3	The structured abstract presents a tremendous amount of information in a very small space. I found it distracting that the strength of evidence was cited for some but not all treatments. For the ones were it wasn't cited, did the authors imply that the SOE was strong. This should be clarified.	We revised the abstract so that it is clearer which interventions were assigned which strength of evidence ratings.

Section	Commentator & Affiliation	Comment	Response
General	Peer Reviewer #3	More importantly, I found the mention of studies showing primarily short-term benefits to be inexact. It would be helpful to state what this was defined as in these analyses. In addition, one could consider this differently for those with acute and chronic low back pain. For patients with acute low back pain, a short-term benefit is often one that is assessed immediately after treatment or maybe for a week or two. Longer treatments may be sufficient to manage the acute episode. However, for chronic low back pain, I think of short-term benefits as not being adequate, since the problem is a chronic one. So short-term may mean something quite different. Emphasis on the lack of long-term outcomes for chronic low back pain should be more explicitly stated if this indeed the case.	We revised the Discussion/Applicability section to emphasize the importance of long-term outcomes for understanding effects of treatments for chronic low back pain.
General	Peer Reviewer #3	When referring to the fewer studies examining radicular pain, I think it may help to highlight that most of these studies involve younger individuals with disc herniation or older individuals with spinal stenosis.	We revised the Discussion/Applicability section to note that some studies of radicular pain required imaging confirmation of disc herniation (typically younger persons) or spinal stenosis (typically older persons) or did not require imaging confirmation of symptoms.

Section	Commentator & Affiliation	Comment	Response
General	Peer Reviewer #3	Lastly, though the scope is intended to include benefits and harms, the abstract only mentions benefits. Given the harms that are being identified with the use of opioid agents, this seems like a glaring omission.	We revised the Abstract/Results to note that pharmacological therapies were associated with increased risk of harms versus placebo; there was no increase in risk of serious harms, but trials were not designed to assess such events.
General	Peer Reviewer #3	In summary, this is a very impressive undertaking and the authors deserve credit for doing it with a high methodologic quality. However, it is hard to discern action items that may lead to meaningful change. I read this and say - this is the best we can say about treatments performed millions of times daily at great expense. And yet, there is no sense of this big picture. We have limited evidence supporting much of what we do for low back pain. Reading this should be a call to arms. However, opiate abuse, fueled by prescription opiates for chronic pain - mainly low back pain, is barely mentioned and mainly as a limitation of the study's methods.	We revised the Discussion to be clearer that opioids have been associated with increased risk of serious harms in observational studies, while noting that such studies did not meet inclusion criteria since they did not focus on patients with low back pain. We cited a recent AHRQ-funded review on long-term opioid therapy that we conducted for an NIH Pathways To Prevention Workshop. We added Tables summarizing the main findings across interventions for acute, chronic, and radicular low back pain.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #6	Overall, the authors have done an excellent job of providing a very comprehensive review of the literature on conservative treatments for low back pain. The categorization of studies based on treatment interventions, control group and type of back pain is useful. The same is true of both the individual study quality and overall level of evidence designations within each treatment subcategory.	Thank you for the comment.
General	TEP Reviewer #6	The scope of the review is ambitious, which necessitates a "dense" writing style. In spite of this, for the most part, the report is easy to read. A very few concerns and comments are listed below: Line 15, page ES-34 typo	Thank you for the comment. Typo corrected.
General	TEP Reviewer #6	Lines 42-46 on page ES-43 - not clearly written	We revised to be clearer:

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #6	Lines 35-37 on page ES-44 - not clearly written, typos in lines 13 and 39 on the same page	We revised to make it clearer: "One positive review based findings on a meta-analysis. However, the study in the meta-analysis that reported the largest effect in favor of antidepressants did not report being randomized, it did not include relevant studies that were in the more current review, and it did not report methods for data imputation for two trials included in the meta-analysis."
General	TEP Reviewer #6	Lines 31-34 on page ES-46 are not referenced and appear to be more opinion than fact	We added citations on variability in health insurance coverage of CAM therapies and on lack of reimbursement for multidisciplinary rehabilitation.
General	TEP Reviewer #6	Lines 44-49 on page 15 are not clearly written	We revised the description of sample sizes to be clearer: "The sample size was 456 in one trial and ranged from 40 to 70 in the others."
General	TEP Reviewer #6	Line 33 page 40. What is placebo exercise?	We corrected this to the correct comparator ("no exercise.")

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #6	My primary concern has to do with how some treatments have been categorized, namely Tai Chi, progressive relaxation and foot reflexology. While I understand that there are only a small number of studies on each of these, the way in which they have been lumped into other categories that are not really related creates a precedent that is concerning and could have payment policy implications. You will have Tai Chi therapies arguing that this treatment should be covered under supervised exercise CPT codes, that progressive relaxation should be covered under the psychology codes and that reflexology should be covered under massage CPT codes. Each of these treatments should be given their own category.	We revised the Methods to be clearer that the broad groupings was for organizational purposes only and was not meant to imply that the therapies were equivalent. Results were presented separately for respondent therapies (which includes progressive relaxation) from other behavioral therapies. The massage trials were very heterogeneous, as described in the Results, and it was not possible to define a "standard" massage technique.
General	TEP Reviewer #7	1) Why is Motor Control Exercise split out from other exercise approaches?	We revised the Results so that the motor control exercise evidence is now integrated into the overall results on exercise therapy.
General	TEP Reviewer #7	2) It should be made clear that different exercise approaches aim for different outcomes. Some programs are shooting for symptom relief, others for improvements in flexibility, strength and endurance for specific tasks/work.	We already note that exercise interventions varied and included general strengthening, stretching, or aerobic exercise; motor control and stabilization; physiotherapy, and other techniques.

Section	Commentator & Affiliation	Comment	Response
General	TEP Reviewer #7	3) While pain scales and functional questionnaire scores are the fundamental outcomes in the literature, they don't relate well to vocational disability and social functioning. Since disability from back pain is becoming a major national health problem, the future research section should encourage research testing interventions' effects on disability, beyond symptoms and impairments.	The Research Gaps section states that research is need to understand effects on other outcomes, including return to work.
General	Peer Reviewer #4	The magnitude of low back pain very large and low back pain is a highly clinically meaningful disorder. The authors have clearly made these points.	Thank you for the comment.
General	TEP Reviewer #8	I have rated this report overall "good". The report is certainly meaningful, clinically. If its target population and audience have been explicitly defined, I missed that definition, but I do not see why that should be considered negatively. The contents of the report are very explicitly stated, and anybody with an interest in low back pain can easily tell from reading the table of contents that the report likely does or does not contain information that will be of interest.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Paul Rocktar (APTA)	On behalf of the American Physical Therapy Association (APTA), we would like to thank the Agency for Healthcare Research and Quality (AHRQ) for the opportunity to comment on the "Draft Report: Noninvasive Treatments for Low Back Pain." APTA commends the United States Department of Health and Human Services and AHRQ for these effective healthcare initiatives and the impactful topics which aim to improve the quality of health of all Americans. The American Physical Therapy Association represents more than 90,000 physical therapists, physical therapist assistants, and students of physical therapy nationwide. APTA's goal is to foster advancements in physical therapist practice, research, and education and to further the profession's role in the prevention, diagnosis, and treatment of movement dysfunctions and the enhancement of the physical health and function of members of the public. Physical therapists perform evidenced-based examinations, screenings, evaluations, and interventions for musculoskeletal, neurological, cardiovascular, pulmonary, and integumentary conditions and provide patient centered care that focuses on function and mobility to improve an individual's quality of life.	Thank you for the information about your organization.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Paul Rocktar (APTA)	Role of Physical Therapists in the Noninvasive Treatments for Low Back Pain Physical therapists are essential providers in noninvasive treatments for low back pain (LBP). Physical therapists provide evaluations and evidence-based interventions that may include strengthening, flexibility, manual therapy techniques such as soft tissue mobilization and spinal manipulation, balance, pain management, in addition to various forms of mobility training to support optimal participation at home, at work and in the community throughout the lifespan. Physical therapists participate in collaborative and patient-centered care throughout a number of settings, including hospitals, skilled nursing facilities, inpatient rehabilitation facilities, outpatient practices, and home health agencies.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer (North American Spine Society)	NASS is a multispecialty medical organization dedicated to fostering the highest quality, evidence-based, ethical spine care by promoting education, research and advocacy. NASS is comprised of more than 8,000 physician and non-physician members from several disciplines, including orthopedic surgery, neurosurgery, psychiatry, pain management, neurology, radiology, anesthesiology, research, physical therapy and other spine care professionals. The North American Spine Society (NASS) appreciates the opportunity to comment on the draft Comparative Effectiveness Review on Noninvasive Treatments for Low Back Pain and commends the authors for reviewing and synthesizing a large volume of literature. Members of NASS' Evidence-Based Guideline Development Committee reviewed the review and would like to offer the following comments. Questions may be directed to Karie Rosolowski, NASS Senior Manager of Research & Quality Improvement, at krosolowski@spine.org or 630.230.3692.	Thank you for the information about your organization.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Joseph Shurman (American Board of Anesthesiology, American Academy of Pain Management)	This is my written response to the recent AHRQ report on noninvasive treatments for back pain. This report was not particularly positive on TENS primarily because there was a focus on a limited subset of research and did not evaluate the correct outcome measures in the studies. TENS has been used for many years. It is noninvasive with minimal risk. Hundreds of studies have been done. Most have been very positive. Some have shown an increase in chemical pain mediators. Some have been negative. What has been lacking is a multicenter double-blind randomized study done throughout this country that would cost millions of dollars.	Thank you for the comment. The Results for TENS are based on randomized controlled trials, including trials conducted in the United States.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Joseph Shurman (American Board of Anesthesiology, American Academy of Pain Management)	We know in the pharma world that it can take up to a billion dollars to bring a drug to market. Unfortunately, the companies that produce TENS do not have that kind of financial power to do that kind of study so we are left with relatively small studies. When one looks at everything we do in pain management from interventional procedures to the prescription of opioids, all of which have significant risk, there are no good long-term studies over six months to a year or more to show efficacy in a multicenter way as to improvement of quality of life, improvement of function, and even pain relief.	The available trials on TENS do not suggest a beneficial effect. Therefore, we do not necessarily agree that a large long-term trial is warranted.
General	Public Reviewer Joseph Shurman (American Board of Anesthesiology, American Academy of Pain Management)	Unfortunately, over half the population over 50 suffers from chronic pain. In my own clinical experience, in many patients, the TENS unit has been very beneficial with the patients taking less medications. We have been able to avoid surgeries with minimal to almost no risk. Skin irritation is listed in this report, but we have found it to be minimal. It is easy to deal with it when you can move the patches to different sports on the body. TENS are used in every major pain clinic in this country and are used throughout the world.	Thank you for the comment. The results for TENS were based on the available randomized trials.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Joseph Shurman (American Board of Anesthesiology, American Academy of Pain Management)	I would hope that you would reconsider or possibly have the NIH grant some funding to do this research as has been done with some pharma drugs.	The purpose of the report is to summarize the existing evidence, not to make funding decisions.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	My name is Gert Bronfort and I'm a research professor and clinical research scientist at the University of Minnesota. I have conducted several of the randomized trials included in this draft AHRQ report and I've also been involved extensively in systematic reviews of the literature pertaining especially to spinal manipulation and exercise for low back pain. This brief commentary has been prepared by me with assistance from members of our research unit: Brent Leininger (primary contributor), Corrie Vihstadt and Megan Conlon. We have chosen to focus primarily on spinal manipulation and exercise but we have also briefly assessed the section on acupuncture. We have concentrated mainly on issues that we think, if verified and incorporated by the Report authors, might possibly lead to a change in the overall conclusions regarding the nature and quality of the evidence.	Thank you for the information.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	The key questions highlight radicular low back pain and spinal stenosis as two important conditions. Evidence statements for these two conditions are inconsistently reported within the key findings compared to evidence statements for acute and chronic non-radicular pain. For example, there are no evidence statements for radicular low back pain or spinal stenosis within the spinal manipulation and acupuncture sections and no evidence statements for spinal stenosis within the exercise section.	Given the very large number of conditions, interventions, and comparisons, we did not have separate sections for every situation in which evidence was not available. For example, there were no trials of acupuncture or spinal manipulation for radicular low back pain that met our inclusion criteria. The evidence on exercise and spinal stenosis was limited to a single small trial; in our judgment this did not warrant a separate bullet point. The Discussion highlights the lack of evidence in patients with radicular low back pain.
General	Public Reviewer Gert Bronfort	Additionally, it is not clear why a recent Cochrane review assessing non-operative treatments for spinal stenosis was not referenced within the report (Ammendolia C, Stuber KJ, Rok E, Rampersaud R, Kennedy CA, Pennick V, Steenstra IA, de Bruin LK, Furlan AD. Nonoperative treatment for lumbar spinal stenosis with neurogenic claudication. Cochrane Database Syst Rev. 2013 Aug 30;8:CD010712.	This review was not included b/c it included many different interventions for spinal stenosis, including a number of interventions excluded from our report (including injections, prostaglandins, vitamin B12) or comparisons (vs. surgery or injections) that were excluded from our report. The trials that evaluated relevant interventions and comparisons are reported in their sections (e.g., NSAIDs).

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	Outcomes other than pain and function are not consistently reported throughout the report which is problematic given the conclusions that future research on outcomes other than pain and function are needed. For example, the key points for spinal manipulation only include pain and function, but the included systematic reviews and additional trials reported many other secondary outcomes.	Given the vast number of interventions and comparisons, we prioritized results for pain and function, as reported in the Methods. As noted in the Discussion, evidence on effects on other outcomes was very limited. Selected key findings on other outcomes are presented in the Results with additional details in the Evidence Tables.
General	Public Reviewer Gert Bronfort	When looking at the GRADE tables for overall evidence determination, we noticed an example that did not make sense to us. The comparison of Duloxetine vs Placebo for chronic low back pain (Appendix H-3) was deemed to have low study limitations, consistency, directness, precision, and undetected reporting bias; however, the overall strength of evidence was rated as moderate. It's unclear why this comparison didn't receive an overall strength of evidence of high, given none of the individual domains within the GRADE system were downgraded.	Two of the three studies were rated fair-quality so SOE was down-rated for study limitations, we corrected the information in Appendix H-3

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	In the conclusion of the structured abstract there's reference to the magnitude of group comparison differences in effect sizes. We would like to suggest that the report authors include reference to the guidance described in the IMMPACT paper published in 2009. The IMMPACT publication discusses this issue in detail and the summary recommendation is: when considering the clinical importance of group differences, to include responder analyses of the primary outcome measures, the treatment effect size compared to available effective therapies, analysis of secondary time points, the safety and tolerability of treatment, the speed of onset and durability of treatment effect, patient compliance, access, cost and perspective. Reference Dworkin RH, Turk DC, McDermott MP, Peirce-Sandner S, Burke LB, Cowan P, Farrar JT, Hertz S, Raja SN, Rappaport BA, Rauschkolb C, Sampaio C. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. Pain. 2009 Dec;146(3):238-44.	The magnitude of effects classifications were based on suggested thresholds developed through a consensus process using experts and a review of the literature and is described in the Methods. Dichotomized outcomes ("responder analysis") were reported when available, though thresholds for minimum clinically important differences for response rates are not well developed and we did not attempt to formally classify them. All effects were based on between-intervention comparisons. Harms are reported separately and we reported whether outcomes were short- or long-term. Other factors such as patient compliance, access, and cost may be considered when determining when selecting among therapies, and are discussed as factors to consider in the section on Clinical and Policy Decisionmaking, but were not included outcomes for our report.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	SPINAL MANIPULATION It's unclear why two recent spinal manipulation trials (published after the Cochrane reviews) were not included within the report under the spinal manipulation section. • Reference #489 was included under the ultrasound section, but not under the spinal manipulation section of the report: Licciardone JC, Minotti DE, Gatchel RJ, et al. Osteopathic manual treatment and ultrasound therapy for chronic low back pain: a randomized controlled trial. Annals of Family Medicine. 2013 Mar-Apr;11(2):122-9. • Reference #224 was included under the exercise section, but not the spinal manipulation section of the report: Bronfort G, Maiers MJ, Evans RL, et al. Supervised exercise, spinal manipulation, and home exercise for chronic low back pain: a randomized clinical trial. The Spine Journal: 2011 Jul;11(7):585- 98.	The Licciardone trial was not included in the spinal manipulation section because it evaluated multiple osteopathic manual techniques including manipulation, massage, stretching, tender point treatment, and "optional treatments." We added the results of the Bronfort 2011 trial to the spinal manipulation section.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	A recent trial by Goertz et al. (reference #468 within the report) was included within the chronic low back pain comparison of spinal manipulation plus other active treatment versus active treatment alone. The authors' excluded participants with a duration of low back pain greater than 4 weeks resulting in a mean duration of less than 2 weeks. The potential impact of this misclassification should be investigated.	We corrected so the Goertz trial was classified as assessing acute low back pain.
General	Public Reviewer Gert Bronfort	The rationale for downgrading evidence summaries from the Cochrane review of spinal manipulation for chronic low back pain is not clear. Also, the evidence statement about spinal manipulation under chronic low back pain in the results section of the structured abstract appears incomplete with no SOE designation and it is linked to a statement about ultrasound and TENS, which was probably not the intent.	We rated SOE based on study limitations, consistency, directness, and precision, as described in the Methods and in accordance with AHRQ methods. The ratings were based on the totality of evidence (new evidence plus evidence in prior reviews) and in some cases the ratings may have differed from ratings given in the reviews. We reviewed the ratings for spinal manipulation and feel they are accurate based on the available evidence. We corrected the error in the draft (the statement about US and TENS should have been deleted).

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	The Cochrane review reported high quality evidence for spinal manipulation compared to active interventions for pain and function at 1 and 12 months and moderate quality evidence for pain and function at 3 months (downgraded for inconsistency). These results were limited to studies with low risk of bias.	This review dichotomized risk of bias as low or high, but the "low" risk of bias trials still had methodological limitations; e.g. almost all trials used an un-blinded design and trials frequently had high attrition. Therefore, we downgraded for study limitations.
General	Public Reviewer Gert Bronfort	The report describes the strength of evidence for this comparison as "Moderate" with another two trials considered. The evidence was downgraded for moderate study limitations. The justification for this downgrading is suspect given all 6 trials within the Cochrane review and 1 of the 2 additional trials were "good" quality or low risk of bias.	This review dichotomized risk of bias as low or high, but the "low" risk of bias trials still had methodological limitations; e.g. almost all trials used an un-blinded design and trials frequently had high attrition. Therefore, we downgraded for study limitations.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	The Cochrane review also reported high quality evidence for manipulation when added to another intervention for pain and function at 3 and 12 months (2 studies with 994-1078 participants per comparison) and low quality evidence for pain and function at 1 month (2-3 studies with 156-228 participants per comparison; downgraded for risk of bias and imprecision)	This review dichotomized risk of bias as low or high, but the “low” risk of bias trials still had methodological limitations; e.g. almost all trials used an un-blinded design and trials frequently had high attrition. Therefore, we downgraded for study limitations. The SOE was also downgraded for imprecision because of the relatively small samples and because a number of the estimates were imprecise (e.g., mean difference -6 on a 0 to 100 scale with a 95% CI of -11 [moderate effect] to -1 [essentially no effect]).

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	The report describes the strength of evidence for this comparison as “Low” with no additional trials considered. It’s unclear why the authors’ chose to highlight the lower strength evidence finding (which did not include effect estimates from a large low risk of bias study) compared to the high quality evidence findings.	This review dichotomized risk of bias as low or high, but the “low” risk of bias trials still had methodological limitations; e.g. almost all trials used an un-blinded design and trials frequently had high attrition. Therefore, we downgraded for study limitations. The SOE was also downgraded for imprecision because of the relatively small samples and because a number of the estimates were imprecise (e.g., mean difference -6 on a 0 to 100 scale with a 95% CI of -11 [moderate effect] to -1 [essentially no effect]).

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	Regarding spinal manipulation for radicular pain the authors have chosen to include a pilot study by our group. We think this is misleading to bring into evidence since it was specifically designed to be a feasibility study that a priori would not report on between group differences. Instead, we suggest you include the full-scale trial by our group that was published last year: Bronfort G, Hondras MA, Schulz CA, Evans RL, Long CR, Grimm R. Spinal manipulation and home exercise with advice for subacute and chronic back-related leg pain: a trial with adaptive allocation. Ann Intern Med. 2014 Sep 16;161(6):381-91	We added the Bronfort trial, which was published after the original searches had been done.
General	Public Reviewer Gert Bronfort	We realize that this article was published a month after the deadline of your search strategy. However you have included trials in the report published in 2015 (e.g. Hurley DA, Tully MA, Lonsdale C, et al. Supervised walking in comparison with fitness training for chronic back pain in physiotherapy: results of the SWIFT single-blinded randomized controlled trial. Pain. 2015;156(1):131-47) Since it's in an area where there are very few trials it would make sense to include our trial, since is likely to have an influence on the overall evidence for radicular low back pain.	We added the Bronfort trial, which was published after the original searches had been done.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	The evidence statement for motor control exercise is inconsistent with the summary comparing different exercise regimens for acute or chronic low back pain.	The motor control exercise results were incorporated into the general section on exercise; the motor control exercise vs. general exercise findings are reported separately from other comparisons of exercise regimens.
General	Public Reviewer Gert Bronfort	On page 40 of the report, the authors' report "There were no clear differences between different exercise regimens in >20 head to head trials..."	The motor control exercise results were incorporated into the general section on exercise; the motor control exercise vs. general exercise findings are reported separately from other comparisons of exercise regimens.
General	Public Reviewer Gert Bronfort	On page 46 of the report, the authors note that for chronic low back pain, there is low quality evidence motor control exercise is associated with lower pain intensity at short and intermediate term versus general exercise, but effects were smaller and no longer statistically significant at long term. Motor control exercise was also associated with better function in the short and long term.	The motor control exercise results were incorporated into the general section on exercise; the motor control exercise vs. general exercise findings are reported separately from other comparisons of exercise regimens.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	The determination of final evidence statements given the results of the included systematic reviews and additional trials is not clearly documented. For example, the included systematic review by van Middlekoop et al. investigating exercise for chronic low back pain found low quality evidence of no effect for exercise compared to no treatment/waiting list controls and low quality evidence for the effectiveness of exercise compared to usual care. On page 39-40 of the comparative effectiveness report, the authors report moderate quality evidence for exercise compared to usual care. Additional trials are reported on pages 42-44, but it is not clear how the results of these trials influenced the overall evidence and led to the upgrading of the evidence from the systematic review by van Middlekoop et al.	We rated SOE based on study limitations, consistency, directness, and precision, as described in the Methods and in accordance with AHRQ methods. The ratings were based on the totality of evidence (new evidence plus evidence in prior reviews) and in some cases the ratings may have differed from ratings given in the reviews.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Gert Bronfort	ACUPUNCTURE - The exclusion of the recent systematic review of Complementary and Alternative treatments for back pain which was prepared for the Agency for Healthcare Research and Quality is problematic in our view because this review followed a similar methodology as the current report, and is of better quality and more comprehensive than the included reviews. This review included 105 trials of acupuncture for low back pain which is over twice as many trials compared to the included reviews by Lee et al. and Lam et al. Excluded review Furlan A, Yazdi F, Tsertsvadze A, Gross A, Van Tulder, M, Santaguida L, Cherkin D, Gagnier J, Ammendolia C, Ansari M, Ostermann T, Dryden T, Doucette S, Skidmore B, Daniel R, Tsouros S, Weeks L, Galipeau J. Complementary and Alternative Therapies for Back Pain II. Evidence Report/Technology Assessment No. 194. (Prepared by the University of Ottawa Evidence-based Practice Center under Contract No. 290-2007-10059-I (EPCIII). AHRQ Publication No. 10(11)E007. Rockville, MD: Agency for Healthcare Research and Quality. October 2010.	We did not include the Furlan review because it did not stratify results according to duration of symptoms. The reviews by Lee et al and Lam et al are more recent and focused on acute or chronic low back pain; many of the trials in the Furlan review included patients with LBP of mixed or unknown duration. Nonetheless, conclusions of the Lam, Lee, and Furlan reviews were generally consistent, and we added a reference to the Furlan review in the “Findings in Relation To What Is Already Known” section of the report.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Paul Nolet	I have made a few comments for the introduction in this and another submission. The Global Burden of Disease Studies I cited and a couple of papers by Hoy et al might be good to cite in the intro. Hoy D March L Brooks P Woolf A Blyth F et al 2010Measuring the global burden of low back pain. Best Practice Res Clin Rheumatol 24155165Hoy D Bain C Williams G March L Brooks P et al 2012 A systematic review of the global prevalence of low back pain. Arthritis Rheum 64620282037	The Introduction discusses the burden associated with low back pain. Given the size of the review we do not think additional detail is warranted.
General	Public Reviewer David BenEliyahu DC	I don't agree with your conclusion and believe it does not reflect the literature and evidence.	We reviewed suggested citations and added two new trials published since the original searches; they did not affect the SOE ratings.

Section	Commentator & Affiliation	Comment	Response
General	Public Reviewer Carrie Goettsch	These results directly contradict other quality peer reviewed evidence reports regarding the benefits and cost effectiveness of spinal manipulation and other non-pharmaceutical treatments for low back pain. Please go to www.provenbackcare.com and review the online booklet Evidence Based Spine Care for many quality studies supporting the benefits of SMT for low back pain. Too large of a file to attach I am one of the authors. If the goal of this AHRQ draft evidence report on low back pain is to strengthen the forces behind discrediting or reducing the cultural authority of non-pharmacologic treatments then the goal has been achieved. However if the goal of this report and others to come from AHRQ is to truly tease out the most effective treatments for low back conditions which adds real value to public health then future reports need to refine the searches based on specificity of condition acute or chronic and whether or not they are post surgical and by diagnoses such as disc derangement sacroiliac dysfunction radicular syndromes myofascial as primary pain generator etc. Too broad of a brush creates a blurry picture in art and in science.	We synthesized the evidence on noninvasive therapies for low back pain using a pre-defined protocol developed with the input of Key Informants and a Technical Expert Panel.

Source: https://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=2178
Published Online: February 29, 2016

Section	Commentator & Affiliation	Comment	Response
General	Anonymous Public Reviewer	Well that sucks. And here I thought it was just me.	Thank you for the comment.
Introduction	TEP Reviewer #1	The introduction is appropriate. I have no suggestions for this section.	Thank you for the comment.
Introduction	TEP Reviewer #3	Page vi, Results section: “"spinal manipulation was as effective as other active interventions" - not clear whether this means that SM is an effective therapy compared to placebo, sham, no treatment, usual care, or wait list. Why is it being separated in this fashion from the proceeding sentence that explicitly states a number of interventions that are deemed as effective?	Correct, the results for spinal manipulation are reported for the comparison between other active interventions because the comparison of spinal manipulation vs. sham manipulation suggested no effectiveness. In this situation, we think the evidence for the comparison against active interventions is relevant.
Introduction	TEP Reviewer #4	balanced	Thank you for the comment.
Introduction	TEP Reviewer #5	Very clear, concise, and appropriate.	Thank you for the comment.
Introduction	Peer Reviewer #1	I thought the introduction was quite well written and clear.	Thank you for the comment.
Introduction	Peer Reviewer #2	well done	Thank you for the comment.
Introduction	Peer Reviewer #3	No additional comments beyond those raised in section a.	Thank you for the comment.
Introduction	TEP Reviewer #6	See above.	Thank you for the comment.
Introduction	TEP Reviewer #7	See above.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
Introduction	Peer Reviewer #4	The introduction is well-written and reiterates the major concepts about low back pain epidemiology that have been known for years by healthcare providers and researchers. The straightforward style of writing in the introduction should be advantageous for a lay audience.	Thank you for the comment.
Introduction	TEP Reviewer #8	The introduction is clear and well written.	Thank you for the comment.
Introduction	Public Reviewer Paul Rocktar (APTA)	The Background section notes that "Spinal imaging abnormalities such as degenerative disc disease, facet joint arthropathy, and bulging or herniated intervertebral discs are extremely common in patients with LBP, particularly in older adults, and such findings are poor predictors for the presence or severity of low back pain." It would also be helpful to note that these imaging findings are also common in individuals without LBP, providing additional support for the poor predictive ability of spinal imaging for this population.	We revised to be clearer that such findings are common in persons with or without low back pain.

Section	Commentator & Affiliation	Comment	Response
Introduction	Public Reviewer Paul Nolet	Low back pain can have major adverse impacts on quality of life and function ref.ie. Nolet PS Kristman VL Cote P Carroll LJ Cassidy JD. Is low back pain associated with worse health related quality of life 6 months later European Spine Journal. 2015243458466. Prognosis Many with mild LBP do not seek care and most have a good prognosis 5. A systematic review of 11 studies including those seeking primary care for LBP found that only one third had recovered at the 12 week mark with 65 still reporting LBP 1 year later 65. Hayden JA Dunn KM van der Windt DA Shaw WS 2010 What is the prognosis of back pain Best Practice Res Clin Rheumatol 241671796. Itz CJ Geurts JW van Kleef M Nelemans P 2012 Clinical course of nonspecific low back pain a systematic review of prospective cohort studies set in primary care. Eur J Pain 171515	Thank you for the comment. The Introduction discussed burdens associated with low back pain.

Section	Commentator & Affiliation	Comment	Response
Introduction	Public Reviewer Paul Nolet	I would like to see the global burden of disease study cited in the introduction. Low back pain is the leading cause of years lived with disability 1 and low back pain is in 6th place for disability adjusted life years DALY 2.1. Vos T Flaxman AD Naghavi M Lozano R Michaud C Ezzati M et al 2012 Years lived with disability YLDs for 1160 sequelae of 289 diseases and injuries 19902010 a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380216321962. Murray CJL Vos T Lozano R Naghavi M Flaxman AD et al 2012 Disability adjusted life years DALYs for 291 diseases and injuries in 21 regions 19902010 a systematic analysis for the Global Burden of Disease Study 2010. Lancet 38021972223	The Introduction describes the burden associated with low back pain. Given the size of the report we did not add additional detail.
Introduction	Public Reviewer David BenEliyahu DC	I don't agree with your assessment with respect to evidence for spinal manipulation of neck and back pain. Your assessment is under rating the value of spinal manipulation for neck and back pain	The review does not address interventions for neck pain. The strength of evidence rating for low back pain was based on the available trials, as shown in the Results and the Strength of Evidence Table.

Section	Commentator & Affiliation	Comment	Response
Introduction	Public Reviewer Lisa Culver	The Background section notes that Spinal imaging abnormalities such as degenerative disc disease facet joint arthropathy and bulging or herniated intervertebral discs are extremely common in patients with LBP particularly in older adults and such findings are poor predictors for the presence or severity of low back pain. It would also be helpful to note that these imaging findings are also common in individuals without LBP providing additional support for the poor predictive ability of spinal imaging for this population.	We revised to be clearer that these findings are often present in persons with or without low back pain.
Methods	TEP Reviewer #1	The inclusion and exclusion criteria are justifiable, the search strategies are explicitly stated and logical, and the methods are appropriate.	Thank you for the comment

Section	Commentator & Affiliation	Comment	Response
Methods	TEP Reviewer #1	Assessing Applicability section (page 76/923): I would question whether a difference of 1 point on the Roland-Morris Disability Questionnaire should be considered a small/slight effect; I am not sure the data would support that this is any kind of true effect. Similarly, I am not sure data support calling a 2-point difference on the RDQ a moderate effect. I also note that the authors define a small/slight effect for the Roland as 1-2 points, and a moderate effect as 2-5 points. I suggest redefining these categories to make clear whether 2 points on the Roland is a slight or moderate effect. This might be a place (or perhaps in the Discussion) to mention the difference between defining effect sizes or clinically important differences on outcome measures of pain and function in clinical trials based on within-patient or within-treatment group change scores over time versus based on mean differences between groups at a given point in time (e.g., post-treatment), and the frequent confusion between these.	We added a sentence to this section to be clearer that proposed thresholds for minimum clinically important differences generally fall into our “moderate” classification, such that the clinical relevance of “small/slight” effects is uncertain. This point is also made in the Discussion when discussing Applicability. We also revised to be clearer that estimates of magnitudes of effect are based on between-group differences.

Section	Commentator & Affiliation	Comment	Response
Methods	TEP Reviewer #1	I would also suggest that the authors, when they explain how they defined effect sizes in terms of mean differences (Assessing Applicability section on page 10 of the full review; page 76 of 923), clarify that they mean a mean difference (or adjusted mean difference) between treatment groups at post-treatment or follow-up (if this is indeed what they do mean).	We revised to be clearer that we assessed effects based on between-group differences following treatment.
Methods	TEP Reviewer #3	Page ES-6, Methods, Assessing Applicability, lines 40 – 49: The definitions for small, moderate, and large effects for pain and function are clear. However, they also are somewhat liberal. For example, for pain intensity a change of only 10 points on the VAS (or 1 point on the NRS) would generally not be considered clinically meaningful, even if it was statistically significant. If the authors wish to continue to use these definitions, then it would be helpful to frame them in terms of their clinical meanings and provide a rationale for why use such liberal definitions. This is later addressed in the Discussion section, but it jumps out right away and raises questions.	We revised the Methods and Applicability section of the Discussion to more clearly explain that small/slight effects might be below minimum clinically important thresholds. We explained that we stratified in this way to provide readers with better gradation of results and because the clinical relevance of small/slight effects might vary in individual patients depending on preferences, baseline symptom severity, trade-offs with harms, cost, and other factors.
Methods	TEP Reviewer #4	appropriate	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
Methods	TEP Reviewer #5	The Methods are excellent and appropriate. The authors needed to constrain the scope of their work and made sensible decisions about how to do that.	Thank you for the comment.
Methods	Peer Reviewer #1	The methods are standard and well justified.	Thank you for the comment.
Methods	Peer Reviewer #1	The authors have taken pains to upgrade the quality of "fair" reviews. However, in my experience, primary reviews of non-pharmacologic therapies can be flawed and the authors of this report would have no way of knowing that.	As described in the Methods, we used validated criteria to assess the quality of systematic reviews and focused on good-quality reviews. We only included fair-quality reviews if we could overcome shortcomings (e.g., inadequate quality assessment) through evaluation of the primary studies.
Methods	Peer Reviewer #2	I have issues with the systematic review process as you do not distinguish between reviews and meta-analysis. I believe you should distinguish between systematic reviews without meta-analysis and those with. Many reviews without analysis may merely provide a thorough search of the literature and a grading process, which, as you correctly note, frequently must be repeated by you because the process is not sufficiently strictly performed. (Of course this must be done also for a number of meta-analysis if they include weak articles.)	We included systematic reviews that were qualitative or quantitative (meta-analysis), as long as they met quality criteria using a validated quality assessment instrument.

Section	Commentator & Affiliation	Comment	Response
Methods	Peer Reviewer #2	I was disappointed that you did not go back into previous meta-analysis and add new article to come up with SMDs reflecting all of the current literature. I would strongly urge that this be considered a requirement in the future. We do this and find it enormously helpful.	Given the very large number of interventions and comparisons evaluated it was not feasible to update the meta-analyses.
Methods	Peer Reviewer #2	I personally believe that weak articles should not be discussed at all as evidence when there are stronger articles available because I think it tends to obfuscate the issues. But I understand currently this is the trend and I think you are very clear about the quality of articles you reviewed and the evidence, so that is what is important.	Thank you for the comment.
Methods	Peer Reviewer #3	Here the authors define short-term as <6 mo. and long-term >1 year. As mentioned, this may be reasonable for chronic low back pain, but is a pretty meaningless distinction for acute low back pain.	We revised the Discussion/Applicability section to be clearer that long-term outcomes are of particular importance for chronic low back pain. We also still think it is appropriate to consider long-term effects of treatment for acute low back pain—treatments with sustained benefits would be more clinically relevant than those that provided immediate/short-term relief (though the latter is important too).

Section	Commentator & Affiliation	Comment	Response
Methods	Peer Reviewer #3	There is no mention of the process of evaluating harms from non-RCT studies. It seems that the focus is overwhelmingly on benefits. Harms are given such passing mention, that one should remove mention that this was performed.	In the Methods, we state that cohort studies for harms were included when randomized trials were sparse or unavailable. Unfortunately few observational studies met inclusion criteria because few specifically focused on low back pain (e.g., opioids). For each intervention, evidence on harms is summarized in its own section.
Methods	Peer Reviewer #3	It would be helpful to define what some of the treatments are (e.g. MCE).	We added a description of Motor Control Exercise in the Results section.
Methods	TEP Reviewer #6	See above.	Thank you for the comment.
Methods	TEP Reviewer #7	See above.	Thank you for the comment.
Methods	Peer Reviewer #4	The methods are thorough, sound, and rely on standard systematic review strategies found in high quality reviews.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
Methods	TEP Reviewer #8	The report is very detailed in describing the methodology utilized by the authors. Their work is a systematic review, not a clinical trial. They have analyzed literature that reflects the inclusion and exclusion criteria chosen by others, not themselves. That, of course, is the nature of the systematic review process. They did make clear how the articles that were reviewed were culled from a much larger potential body of literature.	Thank you for the comment.
Methods	TEP Reviewer #8	From my perspective, the statistical methods the authors have utilized are clearly documented, and are very transparent.	Thank you for the comment.
Methods	Public Reviewer Paul Rocktar (APTA)	Focusing on mechanisms of injury may underlie overuse of imaging and procedures. However, more comprehensive information on patient presentation and the impairments underlying LBP could help to explain why particular intervention approaches may be successful or unsuccessful in managing LBP.	Details about patient characteristics were abstracted from the studies and available (for primary studies) in the evidence tables. In general, there was insufficient evidence to determine how patient factors predict responsiveness. As noted in the Discussion, more research is needed to determine factors that predict responsiveness to specific therapies.

Section	Commentator & Affiliation	Comment	Response
Methods	Public Reviewer (North American Spine Society)	Population of Interest NASS questions why spinal stenosis was included in the review population of adults with acute, sub-acute, or chronic non-radicular low back pain, or radicular low back pain. Spinal stenosis is a radiographic finding and a structural diagnosis, not a symptom like low-back pain or radicular pain. In the evidence summaries, all results are clearly delineated by sub-groups, when possible, except for spinal stenosis. Although the authors discuss the limitations of extrapolating evidence findings under the "Applicability" section, no explanation is provided as to why spinal stenosis is not stratified. NASS recommends that all evidence-findings and conclusions be stratified for each patient population searched, including spinal stenosis patients, to assist providers in understanding and applying the evidence to each subgroup of patients. If each subgroup is unable to clearly be delineated, due to inadequate sub-group analysis, then this needs to be discussed in each section. It is also unclear in the inclusion/exclusion criteria if any other structural diagnoses are considered in this review. This should be further clarified.	Radiculopathy is a symptom that is most commonly caused by herniated disc or spinal stenosis. However, many trials of radicular low back pain did not perform imaging correlation, or included populations with mixed underlying conditions. We presented results for Radicular LBP for each intervention, and if there was information about the underlying condition that data was presented. We revised the Methods to be clearer that radicular LBP could be due to a herniated disc or spinal stenosis and revised the Discussion/Applicability section to note that studies of radicular pain evaluated diverse populations (including persons with imaging findings of herniated disc or spinal stenosis or mixed populations/no imaging correlation required).

Section	Commentator & Affiliation	Comment	Response
Methods	Public Reviewer (North American Spine Society)	Not one description of any of the 36 treatments discussed in this report provides information about the patient population with low-back pain who might or might not respond to that intervention. We question if derangement was considered in the population of interest? There is a growing body of research that has successfully identified and validated at least one major LBP subgroup.(Donelson R. Improving spine care using mechanical diagnosis and therapy. SpineLine. 2012; September/October.) That's the "derangement" subgroup in whom "pain centralization" and a "directional preference" are reliably elicited during a standardized baseline mechanical examination. Subgroup-specific RCTs, and systematic reviews then validate this derangement diagnosis, but these studies are routinely overlooked in reviews due to conventional search methods.	The inclusion criteria for trials was reported in evidence tables and in the results. We revised the Exercise section to be clearer that trials of motor control exercise generally enrolled patients with deficits of motor control, though tests to assess motor control and specific inclusion criteria varied. We also revised the Applicability section to note that it is unclear whether effects of MCE vary according to findings of motor control tests. The Research Gaps section notes that research is needed to understand which patients are likely to benefit from which interventions.

Section	Commentator & Affiliation	Comment	Response
Methods	Public Reviewer (North American Spine Society)	In addition, “radicular” pain is not clearly defined in this review. Since this is an ambiguous term, this needs to be clearly defined to ensure that providers/readers fully understand the patient population being addressed.	We revised the Study Selection/Population section to note that Radicular pain was defined as back and leg pain, with or without sensory of motor deficits, in a nerve root distribution; also noting that radicular pain could be based on clinical presentation or require imaging correlation (e.g. a herniated disc or spinal stenosis).
Methods	Public Reviewer (North American Spine Society)	Study Designs NASS is disappointed that observational studies assessing interventional effectiveness were excluded from this review. While we understand that randomized controlled trials and meta-analyses/systematic reviews are considered the gold standard for study design, we noted over 50 topic areas with “Insufficient Evidence.” Excluding these studies provides readers with the false assumption that there is no evidence to address these topics when in fact high quality cohort studies may have provided the evidence necessary to develop conclusions for these topic areas. NASS recommends that observational studies be considered for interventional effectiveness for at least topics areas where there is no or limited evidence.	The review was restricted to randomized controlled trials for effectiveness. Even well-conducted observational studies are susceptible to bias and confounding; such effects have been shown to be more pronounced for subjective outcomes such as pain and function; and observational studies of low back pain interventions have frequently reported results quite inconsistent with subsequent well-done trials.

Section	Commentator & Affiliation	Comment	Response
Methods	Public Reviewer David BenEliyahu DC	Review of recent literature	Noted.
Methods	Public Reviewer Carrie Goettsch	I'm concerned that depending on the selection of research many less specific studies have diluted the benefits of spinal manipulation and conservative treatment in general for low back pain.	We included trials on the interventions specified and stratified results for acute, chronic, and radicular low back pain.
Methods	Public Reviewer Andrew Engel	Page ii states The information in this report is intended to help health care decision makers, patients and clinicians, health system leaders and policymakers among others make well informed decisions and thereby improve the quality of health care services. While this is a noble goal we certainly share your methodology may inadvertently prohibit its actualization by unduly restricting the reports authors from including valuable data. Without the missing data a clinician cannot make evidence based decisions thereby affecting the ability to deliver the safest and highest quality care.	We synthesized the evidence on noninvasive therapies for low back pain using a pre-defined protocol developed with the input of Key Informants and a Technical Expert Panel. The type of evidence included, specifically focusing on randomized trials to evaluate effectiveness, was determined with the input of the Key Informants and the Technical Expert panel.

Section	Commentator & Affiliation	Comment	Response
Methods	Public Reviewer Lisa Culver	Focusing on mechanisms of injury may underlie overuse of imaging and procedures. However more comprehensive information on patient presentation and the impairments underlying LBP could help to explain why particular intervention approaches may be successful or unsuccessful in managing LBP.	Details about patient characteristics were abstracted from the studies and available (for primary studies) in the evidence tables. In general, there was insufficient evidence to determine how patient factors predict responsiveness. As noted in the Discussion, more research is needed to determine factors that predict responsiveness to specific therapies.
Results	TEP Reviewer #1	The amount of detail in the results section is appropriate. The study characteristics are clearly described. The key messages are explicit and applicable. The figures, tables, and appendices are adequate. I am not aware of any studies that were overlooked and should have been included, and I did not identify any studies that should have been excluded.	Thank you for the comment
Results	TEP Reviewer #1	I believe the date of this Cochrane review was 2010, not 2011, and the PMID is 20614428.	We corrected the publication year for the Henschke review.

Section	Commentator & Affiliation	Comment	Response
Results	TEP Reviewer #1	Also, the wording in this Cochrane review is that "Respondent treatment aims to modify the physiological response system to pain, through reduction of muscular tension. The theoretical basis of this approach is the assumption of a pain-tension cycle, where pain is viewed as both a cause and a result of muscular tension. Respondent treatment attempts to interrupt this cycle by using a tension-incompatible reaction, such as relaxation." The authors' re-wording on page 56 of the full review (page 122/923) (last paragraph) refers to a "psychological response to pain" and I think this may be a misinterpretation of the wording in the Cochrane report or a typo. I recommend that the authors revise this sentence to be in accordance with the wording in the Cochrane review.	Changed the word "psychological" to "physiological."
Results	TEP Reviewer #1	The authors might also note that operant treatments include exercise and thus are really combined exercise-psychological treatments.	We revised the description of operant therapies to: "Operant therapies refer to behavioral therapies that encourage healthy behaviors such as exercise and participation in usual activities, and that do not reinforce patient pain behaviors."

Section	Commentator & Affiliation	Comment	Response
Results	TEP Reviewer #1	The sentence on the full review page 56 (or page 122/923), lines 48-50 doesn't make sense: "Operant therapies refer to psychological therapies that encourage healthy behaviors such as exercise and work while discouraging positive reinforcement of pain re-enforcing behaviors." Perhaps reword to something like: "Operant therapies refer to behavioral therapies that encourage healthy patient behaviors such as exercise and participation in customary activities, and that do not reinforce patient pain behaviors." Similarly, perhaps the description of cognitive therapies in the next sentence could be changed to something like: "Cognitive therapies help patients to identify and challenge maladaptive thoughts that contribute to disability and distress." Finally, I would also recommend rewording the following sentence in that paragraph (lines 53-55) that defines respondent therapy, as suggested above.	We revised the descriptions of operant, cognitive, and respondent therapies as suggested.

Section	Commentator & Affiliation	Comment	Response
Results	TEP Reviewer #1	The authors might also want to consider their continued use of the category "respondent therapy;" this is not a term commonly used to describe these therapies (relaxation training, biofeedback).	We used the term "respondent therapy" since it is the term used in the Cochrane review that main results were based on, but we revised the description of respondent therapy to state: "Respondent therapy includes techniques such as relaxation or biofeedback, and is based on the premise that the physiological response to pain…"
Results	TEP Reviewer #1	Page 57 (or page 123/923), lines 46-51: I am curious as to why the authors included a trial of meditation in the category of respondent therapy. Also- Table 15: It is unusual to refer to meditation as a "respondent therapy."	We removed the meditation trials from the section on behavioral therapies as they are not considered a typical behavioral therapy; they are now excluded from the report as meditation is not one of the interventions within the scope.
Results	TEP Reviewer #1	Table page 129 (page 195 of 923): I believe the Henschke (note correct spelling) Cochrane review is 2010, not 2011.	We corrected the date to 2010.
Results	TEP Reviewer #1	In the tables when reporting data on VAS and NRS scales, it would be helpful to include the possible range, as some studies used 0-100 scales and others used 0-10 scales.	We reviewed the Tables and added the pain scales when they were missing.
Results	TEP Reviewer #3	Page ES-8, lines 51 – 52, WMD - this acronym, which is presumably 'weighted mean difference', needs to be spelled out the first time it is used.	We spelled out WMD when it is first presented.

Section	Commentator & Affiliation	Comment	Response
Results	TEP Reviewer #3	Page ES-13, section Motor Control Exercise (MCE): there are two separate bullets on MCE for chronic LBP, both refer to a systematic review, but it is unclear if they are referring to different systematic reviews, which is my guess. In this Executive Summary, references/citations are generally not being included, but it can create ambiguity when there are multiple systematic reviews that are being summarized, as is being done in this section on MCE. I suggest that the text be revised (in the MCE section as well as in other relevant sections) to indicate when different systematic reviews are being summarized.	The results in the main report clearly describe that it is a single systematic review. The results in the Executive Summary were substantially revised so the bullet points are no longer present and the Results no longer refer to specific reviews/studies.
Results	TEP Reviewer #4	Amount of detail is reasonable for someone looking for an in-depth evaluation. Will be difficult for a general practitioner.	Thank you for the comment.
Results	TEP Reviewer #5	The various sections of the report present varying degrees of detail, so readers interested in more or less detail and choose the section best suited for their needs. The studies are clearly described, the key messages are explicit and applicable, the figures, tables, and appendices are satisfactory. I am not aware of any studies that were missed.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
Results	TEP Reviewer #5	ES-1, line 44: this sentence would make more sense if it stated that spinal abnormalities are common in persons WITHOUT back pain (not in persons WITH back pain as written)	We revised to state that spinal abnormalities are common in patients with or without low back pain.
Results	TEP Reviewer #5	E-265: This is my own study and I noticed that this table indicates pain duration as >6 weeks. In fact, the pain duration in the study was >3 months	We corrected this to >12 weeks
Results	Peer Reviewer #1	The authors have done a reasonable job with presenting the results. In some instances where there are multiple reviews that include the same studies, it is difficult to know how much evidence really exists. I think the specific tables may be of interest to researchers who wish to conduct new studies. Like all researchers, I am not expert on every intervention, but they seemed to do a reasonable job.	We added Tables summarizing main findings across interventions for Acute, Chronic, and Radicular low back pain, including the number of trials and strength of evidence ratings.

Section	Commentator & Affiliation	Comment	Response
Results	Peer Reviewer #2	See above studies that might be excluded are those that are very weak and thus provides no real guidance that actually ranks above a multi-disciplinary consensus group.	As described in the Methods, we did not exclude lower-quality studies a priori, given variability in quality assessments, potential misclassification related to what was done versus what was reported, and inconsistency in empiric studies about the effects of study quality assessments on assessments. However, study quality was factored into the evidence syntheses through sensitivity and subgroup analyses and is a core domain for strength of evidence ratings.
Results	Peer Reviewer #2	Sorry I lost the site for this info but it was in interesting article reviewing result changes in the cardiac guidelines.	As described in the Methods, we did not include non-RCTs for evaluating effectiveness, which is the main point of the article cited by the reviewer.
Results	Peer Reviewer #2	Class 1 Recommendations from American College Cardiology & American Heart Association Clinical Practice Guideline of 619 Recommendations 80% or 495 were downgraded.	As described in the Methods, we did not include non-RCTs for evaluating effectiveness.

Section	Commentator & Affiliation	Comment	Response
Results	Peer Reviewer #2	Likelihood of being down graded due to lack of RCT and based on opinion rating only (odds ratio, 3.14, 95% CI 1.69-5.85, P<.001). Downgraded due to being based on observational studies versus RCTs was essentially the same odds ratio 3.49 CI, 1.45-8.41; P=.005). I think this argues against giving lower level studies an evidence grade.	As described in the Methods, we did not include non-RCTs for evaluating effectiveness.
Results	Peer Reviewer #3	In general, it would be helpful to mention when there was no information. For example, that there are no studies of acetaminophen for patients with radicular pain due to a herniated disc or spinal stenosis. In addition, there is inconsistent mention of the follow-up interval where the outcome is reported. I know this is difficult given that this report focuses on the systematic review rather than the individual study, but some mention of the duration of follow-up is important. Some treatments are more likely to have shorter/longer follow-up assessments than other treatments. Acetaminophen: no mention of harms. Opioids: insufficient mention of harms related to long-term risks. More studies have begun to report. They should be included. There is even a systemic review that isn't mentioned.	One of the Key Points for acetaminophen states that "No study evaluated acetaminophen for radicular pain." We don't think it is necessary to add a separate section in the Results to provide the same information. Duration of follow-up is described for each intervention in the beginning of the Detailed Synthesis section and in the Results; we reviewed the Key Points to add this information if it was missing. There is a section in the Results and a Key Point on harms of acetaminophen. We reported harms of opioids from studies that met inclusion criteria and the Discussion discusses other evidence on serious harms that did not meet inclusion criteria.

Section	Commentator & Affiliation	Comment	Response
Results	TEP Reviewer #6	See above.	Thank you for the comment.
Results	TEP Reviewer #7	See above.	Thank you for the comment.
Results	Peer Reviewer #4	The authors have largely repeated and updated the findings of the APS/ACP reviews conducted a few years ago by the same lead author.	Thank you for the comment.
Results	Peer Reviewer #4	New findings on a recent acetaminophen trial have been included.	Thank you for the comment.
Results	Peer Reviewer #4	Regarding the exercise section of the executive summary: The rationale for including a separate sub-section for one type of exercise (motor control exercise) is unclear. Numerous back and core specific exercise approaches are available and utilized in healthcare settings (e.g. back strengthening, directional preference), some of which (e.g. back strengthening) have more RCT evidence than motor control exercise. The authors should either include other specific types of exercise in the executive summary as separate sub-sections, or restrict the executive summary to one sub-section on “back and core specific exercise” and clearly identify the specific types of back and core exercises within this sub-section.	We revised the Results so that the motor control exercise evidence is now integrated into the overall results on exercise therapy.

Section	Commentator & Affiliation	Comment	Response
Results	Peer Reviewer #4	Does motor control exercise include core stabilization exercise? The authors should define this in the executive summary and elsewhere.	We added a description of MCE: "MCE (also referred to as specific stabilization exercise) focuses on strengthening of deep muscles of the spine through a specific stabilization protocol, while reducing unwanted overactivity of other muscles."
Results	TEP Reviewer #8	The conclusions that the authors reach are, in each instance, solidly grounded in the evidence they have reviewed. They are very clear about the strengths and weaknesses of the various studies that comprise the systematic review, about the resulting strengths (or lack thereof) of evidence, and about their conclusions based on that evidence.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer Paul Rocktar (APTA)	We note that “exercise" is not defined. In considering the results it would be difficult to come to valid conclusions about comparative effectiveness without a better understanding of the exercise intervention provided. This could be a very broad range of interventions from aerobic conditioning to back specific exercise with or without supervision. Related to this observation, the outcomes appear somewhat better for "motor control exercise" than "exercise." It is unclear if these studies grouped patients to determine if motor control exercise versus exercise was indicated based on the subjects’ signs and symptoms.	We integrated Results for motor control exercise into the general section on exercise. The Exercise section describe the many different types of exercise therapies evaluated in trials; as noted in the Results, head-to-head trials did not find clear differences between exercise regimens; although there was some evidence that MCE might be associated with better outcomes than general exercise, SOE was low and differences were small.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer Paul Rocktar (APTA)	It appears no distinction was made between “acupuncture" and "dry needling." For example, the report states in its summary of acupuncture, “One trial of sham acupuncture using penetrating needles to non-acupuncture points found no effect on pain.” As there appeared to be no distinction made between these two interventions, the report does not distinguish between “acupuncture” as a technique for balancing the flow of energy through meridians and “dry needling” as stimulating myofascial trigger points. It is then a confounding factor that the sham “acupuncture” site might be a trigger point. APTA would also notes that this could be included in “Limitation of the Evidence Base” as we are aware there are studies specifically of dry needling as just described, but we are also aware many studies do not distinguish between the two mechanisms of intervention.	Dry needling was not an included intervention for this report; we considered active acupuncture to be based on performing interventions at defined acupuncture points. Sham acupuncture could involve dry needles into non-acupuncture points or “simulated” placement of needles.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer (North American Spine Society)	Scope NASS questions whether analysis was done regarding the natural history of acute and chronic pain and the context this provides for treatment? Some cases of acute low-back pain may lesson over time. Some chronic low back pain has a directional preference and can be treated with directional exercises while others benefit from long-term treatment management and lifestyle and behavioral changes. Also, although we realize that the comment period for the question protocol has passed for this review, we would like to recommend that smoking cessation and weight reduction strategies be considered for a future review.	Results were stratified according to duration of back pain symptoms. Several validation studies on directional preference/centralization have not been able to demonstrate that it is effective at improving clinical outcomes. We did revise the Exercise section of results to note that most trials of MCE enrolled patients on the basis of tests showing deficits in motor control. Smoking cessation and weight reduction were not included interventions, and we are unaware of any RCT's that have evaluated these interventions for treating LBP.
Results	Public Reviewer (North American Spine Society)	Results Table NASS finds the Summary of Evidence chart very helpful. To make this information even more succinct, NASS suggests incorporating a table (in either appendices or abbreviated report) with symbols to indicate effectiveness of treatment based on literature findings, such as +, ++, +++, etc.	We added Tables that summarize findings across interventions for acute, chronic, and radicular LBP. Magnitude of effects was classified as small, moderate, or substantial, and SOE graded as insufficient, low, moderate, or high.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer (North American Spine Society)	Opioids Because opioids are among the most controversial treatments discussed in this review, NASS would have liked to see more of a discussion on the effectiveness and recommended dosages of different types of opioids. Based on the question and literature search protocol, it appears that these questions could have been addressed. Was there sufficient evidence to address drug superiority and/or dosage?	As described in the Results, there was no clear difference between different long-acting opioids. The doses evaluated in the trials was provided in the Results and Tables; as trials did not compare different doses of the same opioid it was not possible to determine optimal doses. It was outside the scope of this report to make recommendations regarding dose.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer (North American Spine Society)	Antidepressants The use of antidepressants as a non-operative treatment for low-back pain is controversial and not thoroughly investigated in the literature. However, this report provides evidence from 3 systematic reviews and 10 clinical trials evaluating the effectiveness of antidepressants as a non-operative treatment modality for low back pain. In summation, this review identified duloxetine as associated with reduced pain and improved function for chronic pain. Furthermore, the use of antidepressants was identified to engender an increased risk in adverse events compared to placebo. We believe that the combination of systematic reviews and clinical trials in this review was effective in gauging the impact of antidepressants in the use of chronic lower back pain treatment, in integrating various levels of evidence. Additionally, many of the clinical trial studies utilized control arms in the form of placebos to consider specific antidepressants. The large sample sizes of these trials also inform the applicability of this review's conclusions regarding 3 antidepressants.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer (North American Spine Society)	We were similarly appreciative of the breadth of antidepressants considered – including duloxetine, which was studied more frequently in multiple studies, though solely in contrast to placebo. However, some studies included in this review rarely utilized antidepressants, such as clomipramine, which may detract slightly from the applicability of this reviews' findings; nonetheless, the authors do well to highlight this point in their detailed synthesis. We would have liked to see more studies specifically considering the effectiveness of one antidepressant against another, as only 2 trials looked at this, and this would build on the other studies samples, which only consider one antidepressant versus a placebo. Indeed, one of these 2 studies was of lower-quality.	Thank you for the comment. We summarized the available head-to-head evidence on antidepressants and also described which antidepressants were evaluated in the trials.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer (North American Spine Society)	Antiseizure The management of chronic lower back pain is challenging. A wide variety of treatment modalities have been explored for their potential therapeutic effect for this complex pathology. Antiseizure medications are part of the multimodal medication regimen offered to these patients. Medications that have been evaluated in this review include gabapentin, topiramate, and pregabalin. The role of these medications remain controversial based on limited definitive evidence demonstrating their efficacy. The authors conducted a thorough investigation of all the studies on this topic and critically compared the results reported in these studies. We conclude that antiseizure medications is best used as an adjunctive therapy alongside other treatment modalities. Indeed, multimodal pain therapy is a now a well-recognized treatment strategy for back pain, and this review highlights the positive effects of adding antiseizure medications to patients on other medications. Alternatively, antiseizure medications may be particularly attractive in patients who are contraindicated for NSAIDs, opioids, or other drug classes.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer (North American Spine Society)	Corticosteroids This review summarized the benefits and harms of non-invasive treatment modalities for low back pain. Corticosteroids are often administered to patients with back pain, and it is important to analyze their efficacy in relieving pain and improving function. The authors reported on trials that compared the effects of corticosteroids versus placebo on patients with low back pain. We believe that the authors provided a thorough review of the literature on corticosteroid treatment in patients with low back pain. The authors successfully and meticulously described dosages, routes, and duration of corticosteroid usage in each trial as well as the complications reported while also making note of the fact that adverse events were not always well reported. We appreciate the fact that studies were mentioned that were excluded, such as the one that was written in German. Based on this review, we conclude that corticosteroid usage did not provide any more benefit than placebo for both radicular as well as non-radicular low back pain. This information is an important contribution to the literature and should inform future research.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer (North American Spine Society)	Physical Therapy and Multidisciplinary Rehabilitation The referenced studies for Physical Therapy and Multidisciplinary Rehabilitation for this review revolve around exercise/treatment for non-specific LBP (NSLBP). Historically, traditional randomized controlled trials (RCTs) abound with specific and non-specific non-operative treatment for NSLBP, as opposed to specific treatments for specific subgroups of patients. As a result there continues to be no clear physical therapy (PT) treatments, per the literature regarding the care of NSLBP. The articles utilized within this guideline are similar in nature. There are no specific criteria related to physical therapy other than usual 4 recommended care, functional exercise, extension exercises, best care and exercise, myofascial therapy, and typically used modalities. This indicates a wide array of non-specific passive and active treatment, with no clear representation of symptomatic change relative to location but of only 'pain relief'. No mention is made of other mechanical markers that indicate improvement.	Thank you for the comment. To clarify, "physical therapy" was not considered an intervention in this report, though we evaluated a number of therapies (exercise, various physical modalities) that may be performed in physical therapy settings. "Mechanical markers" are considered intermediate outcomes and outside the scope of this report.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer (North American Spine Society)	A body of evidence suggests that there exists a high prevalence of directional preference (DP) for both acute and chronic LBP, with the chronic population varying from 50-71%.(Donelson R, Long A, Spratt K, Fung T. Influence of directional preference on two clinical dichotomies: acute versus chronic pain and axial low back pain versus sciatica. PM R. 2012 Sep;4(9):667-81. doi: 10.1016/j.pmrj.2012.04.013.Epub 2012 Jun 23.) If a DP can be determined and matched with the appropriate directional exercise, the outcome is noted to be very good.(Werneke M, Hart DL, Cook D. A descriptive study of the centralization phenomenon. A prospective analysis. Spine. 1999;24(7):676-83.)(Long A, Donelson R, Fung T. Does it matter which exercise? A randomized controlled trial of exercise for low back pain. Spine. 2004;29(23):2593-602.)	Several studies have attempted to validate classification systems including systems that evaluate preference of directional preference; several trials have been unable to demonstrate a positive effect (see Cleeland JA et al, Phys Ther 2010;90:1239-50; Learman K et al. Physiother Can 2014;66:359-66; Dougherty PE et al. Chiropr Man Therap 2014;18:41). We revised the Results to note that trials of Motor control exercise enrolled patients on the basis of tests showing deficits in motor control and revised the Discussion/Applicability section to note that it is unclear whether effects of MCE are greater in persons classified as having motor control deficits or not. The Research Gaps section notes that additional research is needed to understand which patients will benefit from which therapies.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer (North American Spine Society)	In likewise manner, the lack of a directional preference is associated with poor conservative outcome. The use of this mechanical approach for acute and chronic NSLBP, and even radiculopathy, allows a process to sort out a mechanical origin of symptoms versus a central origin. This allows for proper selection of treatment, and referral of the proper patient to the proper specialty.	This comment appears to be suggesting an approach to low back pain management or guideline recommendations; we did not make any changes.
Results	Public Reviewer (North American Spine Society)	Based on the body of literature that exists regarding this topic, inclusion of DP studies within any LBP guideline seems prudent. Random controlled trials (RCT) comparing 'usual care' with DP, and treatment that does not utilize a symptom-guided approach would be useful to this guideline. This would show the comparative effectiveness of a subgroup matched treatment to the non-matched treatment of a usual care model.	Our report is not a guideline and does not make recommendations. We did not have a key question to evaluate trials comparing management based on directional preference classification vs. usual care, though we agree that more validation studies are needed and that is mentioned in the Research Gaps section.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer (North American Spine Society)	With that in mind, within such studies involving DP, inter-examiner reliability must be taken into account. A study cannot just perform 'extension exercises' or the like, as this is also a nonspecific approach. Clinicians need to be well versed in Mechanical Diagnosis and Treatment (MDT) in order to establish the validity of patient evaluations, and be able to reliably identify directional preference and symptomatic/ROM /functional /neurological patterns within the exam. The literature indicates that only the clinicians with formal MDT training at the level of Certification or Diploma exhibit high reliability (high kappa values), and those with less training only fair-to-poor reliability. Aina S, May S, Clare H. The centralization phenomenon of spinal symptoms - a systematic review. Manual Therapy. 2004;9:134-43.) (Werneke M, Deutscher D, Hart D, et al. McKenzie Lumbar Classification: Inter-rater Agreement by Physical Therapists With Different Levels of Formal McKenzie Postgraduate Training. Spine: 01 February 2014 - Volume 39 - Issue 3 - p E182–E190)	Thank you for this comment. This appears to be focused on what training should be required to administer these therapies in clinical practice, which is beyond the scope of our report.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer (North American Spine Society)	All research up to this point has focused on nonspecific treatment for the non-specific problem of LBP, and has yielded little insight on LBP treatment. Continued use of these methods for NSLBP is unlikely to shed any new light on this issue or result in any meaningful progress in the treatment of LBP. Recommendations for any guideline will continue to suggest the need for more research, just perpetuating the same results. Therefore, in relation to this review, DP studies should be included to encompass all of the available evidence, and for the reasons outlined above.	The Future Research needs states that studies are needed to determine subgroups most likely to benefit from specific treatments.
Results	Public Reviewer David BenEliyahu DC	Your assent with respect to the strength of evidence for spinal manipulation is lower than it should be and should be modified to reflect the latest literature.	The strength of evidence rating was based on the available trials. The reviewer did not provide additional studies for us to evaluate for inclusion, though an update search was performed.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer Andrew Engel	On page ES9 of the report you stated that opioids are associated with greater short term improvement in pain scores... versus placebo. That pain reduction was about 110 on a 10point NRS or VAS. Can a patient truly appreciate a 1 point difference A 1 point reduction is less than the minimal clinically important change. 1 Even if this 1 point reduction is real is it clinically relevant It appears that you have moderate quality data that opioids are identical to placebo for chronic low back pain. Please consider highlighting that a 1 point reduction is not measurable or clinically relevant and even though the quality of the literature is moderate the treatment has no efficacy.	The Methods and the Discussion/Applicability sections describe how the magnitude of effects were classified; we note that the clinical relevance of "small" effects (0.5 to 1 point on a 0-10 scale) may not be clinically relevant.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer Andrew Engel	As your report is written a clinician could misinterpret the results to imply that there is moderate evidence to support the use of opioids for chronic low back pain. The problem with including systematic reviews is that you are working from an assumption that the authors of the review have completed the review in a meticulous and unbiased manner. The authors of the systematic reviews used in this report combined enriched enrollment randomized withdrawal EERW studies with traditional placebo controlled trials which as you stated may also overestimate the efficacy of opioids. In clinical practice clinicians cannot enrich their patient population. The success rate a clinician would see will be lower than what is sold by EERW studies.	We only included high-quality systematic reviews (based on assessment using the AMSTAR instrument). As we noted, Furlan et al performed an analysis that showed that effects in trials of opioids for chronic pain were similar in trials that used an enriched or non-enriched design, though estimates of harms are lower in the enriched designed studies (which we noted). As we noted, 8 of the 23 trials of opioids for low back pain employed an enriched enrollment and withdrawal design.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer Andrew Engel	How did you separate the results of EERW studies and non EERW studies? The review of EERW and non EERW studies by Furlan 2 that you quoted may not be the best example. That review misinterprets the results of Roth et al 3 implying that all doses of OxyContin were superior to placebo. Without reviewing the original data and simply relying on what is reported in the systematic reviews you have inadvertently carried that mistake forward.	The Furlan systematic review was not included since it evaluated opioids for any chronic pain. The Roth trial also was not included because it did not evaluated patients with low back pain.
Results	Public Reviewer Andrew Engel	Your conclusions regarding the efficacy of treatment are substantially limited because you did not review the original data. Since EERW studies were mixed in with non EERW studies the 1point difference you purportedly found in the data would be an overestimate of the outcomes a clinician would expect in his or her own practice.	Furlan et al performed an analysis that showed that effects in trials of opioids for chronic pain were similar in trials that used an enriched or non-enriched design, though estimates of harms are lower in the enriched designed studies (which we noted). As we noted, 8 of the 23 trials of opioids for low back pain employed an enriched enrollment and withdrawal design, but the evidence suggests that this did not impact estimates of efficacy.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer Andrew Engel	The style of study design is not inconsequential EERW designed trials have already been questioned. 4 A careful reading of the Rauck study 5 demonstrates that opioids have a number needed to treat NNT of 33 when all subjects who attempted the opioid in study are included in the analysis. This NNT is what a provider could expect in clinical practice. Therefore a clinician following your report who prescribed opioids to all low back pain patients would need to un-treat 33 patients before pain relief could be attributed to the provider.	It is unclear how the reviewer determined a NNT of 33. The Rauck trial reports that 68% of patients randomized to opioids had a >30% improvement in pain vs. 31% in the placebo group, for an absolute difference of 36% and a NNT of 2.8. For >50%, the rates were 48% vs. 23%, for a NNT of 4.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer Andrew Engel	Please comment on how you can have moderate quality of data that a treatment is superior to placebo if the NNT is negative and we are not discussing number needed to harm yet. The long-term harms of opioids are extremely serious. For example CDC states: In 2010 nearly 60 percent of the drug overdose deaths (22,134) involved pharmaceutical drugs. Opioid analgesics such as oxycodone hydrocodone and methadone were involved in about 3 of every 4 pharmaceutical overdose deaths 16651 confirming the predominant role opioid analgesics play in drug overdose deaths. 6I understand that the CDCs posting of opioid deaths in 2010 an observational study does not have the same academic rigor as a double blind placebo controlled trial but in this case the information from the observational study is as valuable as data from a randomized controlled trial. Since the long-term negative effects of opioids include death it would seem important to highlight this point. Death is not an acceptable risk for treatment of a subjective symptom.	It is not clear why the reviewer is saying the NNT Is “negative” as he seems to be referring to the Rauck trial based on a previous comment. The Rauck trial reports that 68% of patients randomized to opioids had a >30% improvement in pain vs. 31% in the placebo group, for an absolute difference of 36% and a NNT of 2.8. For >50% improvement, the rates were 48% vs. 23%, for a NNT of 4. The Results and Discussion are clear that the trials were not designed to assess risk of serious harms and refers to another AHRQ-funded review on opioid therapies.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer Andrew Engel	I am concerned that the failure to include observational data especially related to complications and risk benefit analysis has flawed your assessment of opioids by underestimating the substantial risk associated with this treatment.	As described in the methods, cohort studies of harms were included. However, all cohort studies on serious harms of opioids did not meet inclusion criteria because they were not focused on patients with low back pain. This is discussed in the Discussion and a reference to an AHRQ-funded report on benefits and harms of opioids for chronic pain in general is provided.
Results	Public Reviewer Andrew Engel	Please consider revising your conclusion. There is moderate quality evidence that opioids are not clinically superior to placebo since many non-responders were prematurely removed from the data set. For clinicians who elect this treatment even though there is no data to support its use they can expect their patients to experience results on par with placebo while increasing the patients risk of iatrogenic death.	Furlan et al performed an analysis that showed that effects in trials of opioids for chronic pain were similar in trials that used an enriched or non-enriched design, though estimates of harms are lower in the enriched designed studies (which we noted). As we noted, 8 of the 23 trials of opioids for low back pain employed an enriched enrollment and withdrawal design, but the evidence suggests that this did not impact estimates of efficacy.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer Andrew Engel	Reference 1. Mannion A Junge A Grob D Dvorak J Fairbank J. Development of a German version of the Oswestry Disability Index. Part 2 sensitivity to change after spinal surgery. Eur Spine J 200615 6673. 2. Furlan A Chapparro LE Irvin E et al. A comparison between enriched and non-enriched enrollment randomized withdrawal trials of opioids for chronic noncancer pain. Pain Research Management. 201116337351. 3.Roth S Fleischmann R Burch F et al. Around the clock Controlled release oxycodone therapy for osteoarthritis related pain. Arch Intern Med 2000160853860. 4. Engel A. All failures count. Pain Med 201516404. 5. Rauck R Nalamachu S Wild J et al. Single entity hydrocodone extended release capsules in opioid tolerant subjects with moderate to severe chronic low back pain A randomized Double blind placebo controlled study. Pain Med 2014159755.	We reviewed these references and found no new references meeting inclusion criteria.
Results	Public Reviewer Andrew Engel	6. httpwww.cdc.govmediareleases2013p0220drugoverdosedeaths.html	This study does not meet inclusion criteria.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer Lisa Culver	We note that exercise is not defined. In considering the results it would be difficult to come to valid conclusions about comparative effectiveness without a better understanding of the exercise intervention provided. This could be a very broad range of interventions from aerobic conditioning to back specific exercise with or without supervision. Related to this observation the outcomes appear somewhat better for motor control exercise than exercise. It is unclear if these studies grouped patients to determine if motor control exercise versus exercise was indicated based on the subjects signs and symptoms. It appears no distinction was made between acupuncture and dry needling. For example the report states in its summary of acupuncture One trial of sham acupuncture using penetrating needles to non-acupuncture points found no effect on pain. As there appeared t be no distinction made between these two interventions the report does not distinguish between acupuncture as a technique for balancing the flow of energy through meridians and dry needling as stimulating myofascial trigger points.	We integrated Results for motor control exercise into the general section on exercise. The Exercise section describe the many different types of exercise therapies evaluated in trials; as noted in the Results, head-to-head trials did not find clear differences between exercise regimens; although there was some evidence that MCE might be associated with better outcomes than general exercise, SOE was low and differences were small. Dry needling was not an included intervention for this report; we considered active acupuncture to be based on performing interventions at defined acupuncture points. Sham acupuncture could involve dry needles into non-acupuncture points or “simulated” placement of needles.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer Lisa Culver	Options for this population could include movement coordination impairments mobility deficits in acute LBP and patients with acute LBP with referred pain post lumbar microdiscectomy Interventions could include and should be granular enough to distinguish intervention such as mobilization manipulation including thrust trunk coordination strengthening exercises and progressive endurance exercises to name a few.	We state in the Research Gaps section that additional research is needed to identify patient subgroups that may benefit from specific therapies.
Results	Public Reviewer Lisa Culver	1 1. Delitto A George S Van Dillen L et al. Low Back Pain Clinical Practice Guidelines Linked to the International Classification of Functioning Disability and Health from the Orthopaedic Section of the American Physical Therapy Association of Functioning Disability and Health from the Orthopaedic Section of the American Physical Therapy Association. Orthop Sports Phys Ther. 2012424A1A57. doi10.2519jospt.2012.0301	This is a guideline and does not meet inclusion criteria.

Section	Commentator & Affiliation	Comment	Response
Results	Public Reviewer Lisa Culver	It is then a confounding factor that the sham acupuncture site might be a trigger point. APTA would also notes that this could be included in Limitation of the Evidence Base as we are aware there are studies specifically of dry needling as just described but we are also aware many studies do not distinguish between the two mechanisms of intervention.	Dry needling was not an included intervention for this report; we considered active acupuncture to be based on performing interventions at defined acupuncture points. Sham acupuncture could involve dry needles into non-acupuncture points or "simulated" placement of needles.
Discussion/ Conclusion	TEP Reviewer #1	The implications of the major findings are clearly stated. The limitations of the review and included studies are described adequately. I am not aware of any important literature that was omitted in the discussion. The future research section is clear and translatable into new research.	Thank you for the comment.
Discussion/ Conclusion	TEP Reviewer #1	I wonder if the authors want to mention that many trials of psychological approaches not included in the review may be relevant because, although the sample did not exclusively consist of patients with back pain, the majority had back pain. For example, in the trial of Wetherell et al. (2011) of ACT and CBT, 79% of participants had back pain. It seems quite plausible, if not highly likely, that results would be generalizable to a sample of patients who exclusively had back pain.	We revised the Discussion/Limitation of the Review to note that we excluded pain treatment trials that weren't restricted to patients with LBP, and that the applicability of such studies would depend on the proportion of patients with LBP and the other conditions present.

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	TEP Reviewer #1	It might be noted that for many trials of psychological therapies, scores on depression measures were very low at baseline, beneath thresholds for clinically meaningful depression, thus leaving little possible room for improvement with treatment.	We revised the Applicability section to note that most trials of antidepressants excluded patients with depression or only enrolled a small proportion with depression, so that it is not clear if there would be additional effects on mood in more depressed patients.
Discussion/ Conclusion	TEP Reviewer #1	There is a word left out in the first paragraph of the Key Findings Section of the Discussion (page 182 of the report, or page 248/923, lines 26-27) – a word such as “range” should be added before “for pain”. The next sentence could also be revised for clarity.	Corrected the typo and revised the sentence on function to be clearer that effects were small, in some cases there were positive effects on pain but not function, and fewer studies measured pain than function.
Discussion/ Conclusion	TEP Reviewer #1	In the Implications for Clinical and Policy Decisionmaking, the authors might mention in parentheses what nonpharmacological therapies were found to be effective (page 188, or page 254/923, lines 45-46).	We added a sentence noting that a number of pharmacological and nonpharmacological therapies are supported by some evidence of effectiveness for pain or function, and added references to tables summarizing main findings.

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	TEP Reviewer #1	I was curious about the sentence on page 190 (256/923), lines 11-13: "Because responses to pain treatments tend to be bimodal, (reference 620) or marked benefit, assessment of outcomes based on continuous outcomes could obscure treatment effects." I had not previously seen articles showing that responses to pain treatments tend to be bimodal in this way, so I looked up the reference cited to support this statement (reference 620). I read reference 620, which is about analgesic drug trials. In this article, the authors make this statement and provide two citations, both of which were articles authored by the author of reference 620. The first of those two articles cited simply repeats that statement with no supporting evidence. The second article cited summarized data on patients with fibromyalgia treated with medication; the article reported that one-third of patients withdrew before trials ended, one third had pain that was worse or improved less than 30%, and one third had pain reduction of at least 30% from baseline (considered the minimally clinically meaningful level of improvement.) (continues)	We deleted the reference to a bimodal distribution and revised to simply note that assessing continuous as well as dichotomous outcomes would provide a more complete assessment of treatment effects.

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	TEP Reviewer #1	(continued) I would recommend that the authors therefore delete the statement that responses to pain treatments tend to be bimodal, with patients either experiencing no benefit or marked benefit. I think that such a statement cannot be supported by the evidence, and there is a risk that this statement could be picked up and used by others in future articles, without supporting data.	We deleted the reference to a bimodal distribution and revised to simply note that assessing continuous as well as dichotomous outcomes would provide a more complete assessment of treatment effects.
Discussion/ Conclusion	TEP Reviewer #3	n/a	Noted
Discussion/ Conclusion	TEP Reviewer #4	well done	Thank you for the comment.
Discussion/ Conclusion	TEP Reviewer #5	The implications of the major findings are complex but presented in as clear a way as possible. The limitations and Applicability sections should mention that very few of the studies reviewed included any or many older adults, so it is not clear how well the findings apply to older adults. This should also be highlighted as an important area for future research. No important literature was omitted. Future research section is good, except omission of need for research on older adults.	We revised the Applicability and Future Research sections to note that few trials enrolled older adults and the need for research in this area.

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	TEP Reviewer #5	It would be interesting to note if there are any treatment categories for which studies have consistently found that functional outcomes were as good or better than pain reduction. This is important since most clinicians and researchers in the field now consider functional outcomes to be more important.	Unfortunately, even for treatments aimed at improving function such as exercise, behavioral therapies, or multidisciplinary rehabilitation, effects on function were generally small and frequently smaller than effects on pain.
Discussion/ Conclusion	Peer Reviewer #1	I thought the discussion was quite good in terms of talking about the clinical relevance of the evidence for general populations.	Thank you for the comment.
Discussion/ Conclusion	Peer Reviewer #1	Missing is a discussion about older adults, wherein evidence is largely missing and of more seriously injured populations, such as veterans.	We revised the Applicability and Future Research sections to note the paucity of evidence in older adults and the need for studies in this population.
Discussion/ Conclusion	Peer Reviewer #2	I think in the research gap section you did not mention a point you made elsewhere. Using means as the effect doesn't really address the number of successful treatments versus those not successful and that really is the question. We frequently have difficulty with the current studies because it seems there might be a subgroup of patients who would have a very good response but we can't identify that using the traditional statistics. I very much appreciated the need for long term follow up you highlighted.	The issue with reporting of continuous and dichotomous outcomes is addressed in the “Limitations of the Evidence Base” section.

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	Peer Reviewer #3	Very bland reciting. Good for a government report. Unlikely to shake up much. As noted in section f, I think highlighting key gaps identified in bullets or tables would be helpful.	The Discussion summarizes key findings and gaps and has separate sections on Limitations of the Evidence base and on Research needs.
Discussion/ Conclusion	TEP Reviewer #6	See above.	Thank you for the comment.
Discussion/ Conclusion	TEP Reviewer #7	See above.	Thank you for the comment.
Discussion/ Conclusion	Peer Reviewer #4	Discussion/ Conclusion: This discussion/conclusion is successful in addressing the clinical implications of the findings and makes adequate recommendations for clinicians and researchers.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	TEP Reviewer #8	I think that a major shortcoming of the report is that it did not address in the detail I think that it should have, the harms of opioid therapy. I understand that they have produced a systematic review, but I don't think that the comment "Trials were not designed to assess risks of overdose, abuse and addiction, or long-term harms" is at all adequate for a report that purports to analyze both benefits and harms of treatments for low back pain. If the literature they reviewed was limited by their study design to studies that did not assess long-term harms, in my opinion they should have searched elsewhere for quality literature that does in fact examine such issues.	We applied inclusion and exclusion criteria as outlined in the Methods and PICOTS. We revised the Discussion to more clearly highlight risks of opioids reported in observational studies that did not meet inclusion criteria. However, it would not be appropriate to selectively include studies that did not meet inclusion criteria.
Discussion/ Conclusion	Public Reviewer Paul Rocktar (APTA)	APTA would like to thank AHRQ for the opportunity to comment on the draft report on a critically important topic, “Noninvasive Treatments for Low Back Pain.” We look forward to working with AHRQ in the future to ensure that this process is comprehensive and reflects best practices. If you have any questions, please feel free to contact Heather Smith, PT, MPH, Program Director of Quality, at 703-706-3140 or heathersmith@apta.org; or Lisa Culver, PT, DPT, MBA Senior Specialist, Clinical Practice at 703-706-3172 or lisaculver@apta.org.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	Public Reviewer Paul Rocktar (APTA)	Important parameters of exercise and manipulation are related to dosage. In considering comparative effectiveness the results are limited by the absence of dosage information. APTA believes the lack of granularity of the description of these interventions is a "Limitation of the Evidence Base" and should be highlighted.	This is discussed in the "Applicability" section: "For nonpharmacological treatments, the applicability of our findings is affected by the variability between trials in the interventions evaluated. For example, trials of acupuncture varied in the sites in which needles were applied, the length of acupuncture sessions, the number of sessions, and the time period over which the sessions were performed."
Discussion/ Conclusion	Public Reviewer Paul Rocktar (APTA)	The issue of multi-modal treatments should not be considered a "Limitation of the Evidence Base" as this suggests research should not include multi-modal treatments. APTA would suggest the issue of multi-modal treatments be considered an issue of "Applicability" not a "Limitation of the Evidence base"	Evaluation of multi-modal treatments is not described as a Limitation of the Evidence Base, so it is unclear what this comment is referring to. The Research Gaps section notes that research is needed on which combinations and sequences of therapy are most effective.

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	Public Reviewer Paul Rocktar (APTA)	We would recommendation further research based on stratifying the population and matching interventions based on patient characteristics with nonspecific LBP. Options for this population could include movement coordination impairments, mobility deficits in acute LBP, and patients with acute LBP with referred pain post lumbar microdiscectomy, Interventions could include, and should be granular enough to distinguish intervention such as mobilization, manipulation including thrust, trunk coordination, strengthening exercises, and progressive endurance exercises to name a few. (Delitto A, George, S, Van Dillen L, et al. Low Back Pain: Clinical Practice Guidelines Linked to the International Classification of Functioning, Disability, and Health from the Orthopaedic Section of the American Physical Therapy Association of Functioning, Disability, and Health from the Orthopaedic Section of the American Physical Therapy Association. Orthop Sports Phys Ther. 2012;42(4):A1-A57. doi:10.2519/jospt.2012.0301)	The Research Gaps section notes that studies are needed to understand which patients are most likely to benefit from specific therapies. There are many potential classification systems and we do not think any have been sufficiently validated to warrant highlighting at this time (see Cleeland JA et al, Phys Ther 2010;90:1239-50; Learman K et al. Physiother Can 2014;66:359-66; Dougherty PE et al. Chiropr Man Therap 2014;18:41).

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	Public Reviewer (North American Spine Society)	Research Gaps We would like to highlight an area of limited evidence and suggest the consideration of this in the Research Gaps discussion. Two surveys of international low-back pain researchers both reported that the #1 research priority was “Can different varieties or subgroups of LBP be identified, and if they can, what criteria can be used to differentiate them?” (Borkan JM, Koes B, Reis S, Cherkin DC. A report from the Second International Forum for Primary Care Research on Low Back Pain. Reexamining priorities. Spine (Phila Pa 1976). 1998 Sep 15;23(18):1992-6. Review) (Costa Lda C, Koes BW, Pransky G, Borkan J, Maher CG, Smeets RJ. Primary care research priorities in low back pain: an update. Spine (Phila Pa 1976). 2013 Jan 15;38 (2):148) To further this research, reliability studies are needed to demonstrate how to easily identify these subgroups.(Spratt K. Statistical relevance. In: Fardon DF, editor. Orthopaedic Knowledge Update: Spine. 2nd ed. Rosemont, IL: American Academy of Orthopaedic Surgeons; 2002. P 497) Reliable subgroups then feed subgroup specific cohorts to identify potentially effective treatments for the subgroup.	We revised the Research Gaps section to note: “More research is needed to help understand whether nonradicular low back pain can be reliably classified into clinically meaningful subgroups, and which patients or subgroups are most likely to benefit from specific therapies."

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	Public Reviewer David BenEliyahu DC	Comment on Draft Reports and White PapersI do not agree with your assessment with respect to evidence for spinal manipulation of neck and back pain. Your assessment is under rating the value of spinal manipulation for neck and back pain. Your assessment with respect to the strength of evidence for spinal manipulation is lower than it should be and should be modified to reflect the latest literature. There is strong evidence for subacute chronic back leg pain. Studies you did not cite include RCTs comparing spinal manipulation to diclofenac microdoscectomy usual medical care and exercise. See references below. Recent References 1. Dose Response and efficacy for SMT. Haas et.al. Spine Journal 142104 11062.	We rated SOE based on study limitations, consistency, directness, and precision, as described in the Methods and in accordance with AHRQ methods. The ratings were based on the totality of evidence (new evidence plus evidence in prior reviews) and in some cases the ratings may have differed from ratings given in the reviews. Comparisons involving invasive therapies were outside the scope of this report.
Discussion/ Conclusion	Public Reviewer David BenEliyahu DC	SMT and exercise for seniors with chronic neck pain. Maiers et.al. Spine J. 149 21043	Neck pain was outside this report's scope.
Discussion/ Conclusion	Public Reviewer David BenEliyahu DC	Spinal Manipulation and home exercise with advice for subacute and chronic related Leg pain. Branford et.al. Annals of Int Med 2104 1614	We believe this is the Bronfort trial, which was published after the original searches and has been added.

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	Public Reviewer David BenEliyahu DC	Comparison of Spinal manipulation Methods and usual Medical care for acute and subacute Low back Pain. Schneider et.al. Spine Vol 40 No. 4 pg. 2095	This trial was published after the original searches and has been added.
Discussion/ Conclusion	Public Reviewer David BenEliyahu DC	Spinal manipulation in acute nonspecific low back pain a double blinded RCT comparing SMT to diclofenac and placebo Von Heyman et.al. Spine 2013 11387 pg. 540548 Conclusion I don't agree with your conclusion that there is only moderate evidence and believe it does not reflect the latest literature and mischaracterizes the utility validity of Spinal Manipulation. Respectfully submitted David J Ben Eliyahu DC DABCSP DAAPM Selden NY 11784	The Von Heymann trial was included. It was rated poor quality and showed small effects.
Discussion/ Conclusion	Public Reviewer David BenEliyahu DC	There is strong evidence for subacute chronic back leg pain. Studies you did not cite include RCTs comparing spinal manipulation to diclofenac microdoscectomy usual medical care exercise. See references below	Comparisons involving invasive therapies were outside the scope of this report.

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	Public Reviewer Lisa Culver	Important parameters of exercise and manipulation are related to dosage. In considering comparative effectiveness the results are limited by the absence of dosage information.	This is discussed in the "Applicability" section: "For nonpharmacological treatments, the applicability of our findings is affected by the variability between trials in the interventions evaluated. For example, trials of acupuncture varied in the sites in which needles were applied, the length of acupuncture sessions, the number of sessions, and the time period over which the sessions were performed."

Section	Commentator & Affiliation	Comment	Response
Discussion/ Conclusion	Public Reviewer Lisa Culver	APTA believes the lack of granularity of the description of these interventions is a Limitation of the Evidence Base and should be highlighted. The issue of multimodal treatments should not be considered a Limitation of the Evidence Base as this suggests research should not include multimodal treatments. APTA would suggest the issue of multimodal treatments be considered an issue of Applicability not a Limitation of the Evidence base We would recommendation further research based on stratifying the population and matching interventions based on patient characteristics with nonspecific LBP.	This is discussed in the "Applicability" section: "For nonpharmacological treatments, the applicability of our findings is affected by the variability between trials in the interventions evaluated. For example, trials of acupuncture varied in the sites in which needles were applied, the length of acupuncture sessions, the number of sessions, and the time period over which the sessions were performed." Evaluation of multi-modal treatments is not described as a Limitation of the Evidence Base, so it is unclear what this comment is referring to. The Research Gaps section notes that research is needed on which combinations and sequences of therapy are most effective.
Clarity and Usability	TEP Reviewer #1	The report is well-structured and organized. The main points are clearly presented. The conclusions can be sued to inform policy and practice decisions.	Thank you for the comment.
Clarity and Usability	TEP Reviewer #3	n/a	Noted
Clarity and Usability	TEP Reviewer #4	Suitable for policy discussions. The ability of the data to provide concise, consistent guidelines for the practitioner is limited.	The purpose of this report is to summarize the available evidence, not to provide clinical recommendations.

Section	Commentator & Affiliation	Comment	Response
Clarity and Usability	TEP Reviewer #5	Yes to all of this. An excellent and clearly written report!	Thank you for the comment.
Clarity and Usability	Peer Reviewer #1	The report is well structured and the main points are clear. There is so much information because of the different types of back pain, etc. that informing policy and practice will be tough in many instances. Even though the SOE is often 'low' the voluminous nature of the comparisons makes it unlikely that this situation will really improve.	Thank you for the comment. We added tables summarizing main findings across interventions for acute, chronic, and radicular low back pain.
Clarity and Usability	Peer Reviewer #2	Please see my comments under the intro re usability. I liked the structure and found it was easy to locate specific information that I think a reader might want to pursue and know more details about, while the key points where succinct and on target.	Thank you for the comment.
Clarity and Usability	Peer Reviewer #2	Consider reviewing the table at the end a little more perhaps for readability for clinicians. However it is already complete.	We added summary Tables summarizing main findings for Acute, Chronic, and Radicular low back pain.

Section	Commentator & Affiliation	Comment	Response
Clarity and Usability	Peer Reviewer #3	I think the authors could do a better job highlighting by means of bullet points, tables, etc.: 1. Recommendations that have changed as a result of this review 2. Treatments where there is strong evidence of moderate benefit or greater 3. How this study was not designed to adequately address risks 4. Where the gaps are greatest in terms of methods (such as need to define standard criteria for outcome assessment), treatments not supported by evidence that are expensive or potentially risky. In general, given how poor the evidence base - where should we start?	The review summarizes the current evidence and does not make recommendations. The Discussion highlights areas in which findings differ from the prior APS/ACP review. There were no interventions supported by high SOE; we added Tables summarizing findings across interventions for acute, chronic, and radicular low back pain, and interventions supported by evidence are highlighted in the Discussion.
Clarity and Usability	TEP Reviewer #6	See above.	Thank you for the comment.
Clarity and Usability	TEP Reviewer #7	See above.	Thank you for the comment.
Clarity and Usability	Peer Reviewer #4	The report is very well-written and is the most up-to-date and thorough report available on low back pain. However, the transfer of knowledge and dissemination of findings from reports, reviews, and guidelines such as this are usually major shortcomings that inhibit change in policy and clinical practice patterns.	This report will be used by the nominator (ACP) to update its clinical practice guidelines.
Clarity and Usability	TEP Reviewer #8	The structure and organization of at the report are excellent. The report is very clear, and well written.	Thank you for the comment.

Section	Commentator & Affiliation	Comment	Response
References	Public Reviewer David BenEliyahu DC	1. Dose Response and efficacy for SMT. Haas eval Spine Journal 142104 11062.SMT and exercise for seniors with chronic neck pain. Makers eval. Spine J. 149 21043. Spinal Manipulation and home exercise with advice for subacute and chronic related Leg pain. Branford eval. Annals of Int Med 2104 1614. Comparison of Spinal manipulation Methods and usual Medical care for acute and subacute Low back Pain. Schneider eval. Spine Vol 40 No. 4 pg 2095.Spinal manipulation in acute nonspecific low back pain a double blinded RCT comparing SMT to diclofenac and placebo Von Heyman eval. Spine 2013 11387 pg. 540548	Neck pain was outside the scope of this review. The Bronfort and Schneider trials were published subsequent to the original searches and have been added. The Von Heymann trial is already included.