Comparative Effectiveness Review Disposition of Comments Report

Research Review Title: Treatments for Fibromyalgia in Adult Subgroups

Draft review available for public comment from July 23, 2014 to August 19, 2014.

Research Review Citation: Forte ML, Butler M, Andrade KE, Vincent A, Schousboe JT, Kane RL. Treatments for Fibromyalgia in Adult Subgroups. Comparative Effectiveness Review No. 148. (Prepared by the Minnesota Evidence-based Practice Center under Contract No. 290-2012-00016-I.) AHRQ Publication No. 15-EHC006-EF. Rockville, MD: Agency for Healthcare Research and Quality; January 2015. 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 email. 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: http://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=2040
Published Online: January 26, 2015

Commentator	Section	Comment	Response
Peer Reviewer 4	Structured Abstract: Data sources	I do have some comments on the structured abstract and the executive summary and I'm not sure whether these should be included given the categorization for the review so I am putting them under this first section of comments. Structured Abstract, Data sources: It is not the Cochrane Clinical Trials Registry (CCTR) but the Cochrane Central Register of Controlled Trials (CENTRAL) . This is correct in the main text.	Thank you for noting our error. We changed the abstract Data Sources text to read, *Cochrane Central Register of Controlled Trials (CENTRAL).*
Peer Reviewer 4	Structured Abstract, Results	Instead of nonsubgroup effects perhaps use the term 'overall'.	Edited the Abstract Results text – added the term overall: *Subgroup treatment effects were generally small, beneficial, and similar to overall effects in direction and magnitude.*
Peer Reviewer 4	Structured Abstract, Results	In the next to last sentence of 'Results': 'similar in the MDD subgroup' to what? Do you mean similar in those with and without MDD? Or similar in the MDD subgroup and in the overall population?	Edited the Abstract Results test to read: Adverse effects were reported for subgroups in only one pooled analysis and were similar in adults with and without MDD.
TEP Reviewer 2	ES Background	In the executive summary (p. 8 after line 29), consider adding the following. "Most recently the Wolfe 2010 proposed ACR FM diagnostic criteria have been tested against the 1990 criteria and in a multicenter study. Authors found that the proposed 2010 criteria a robust sensitivity of 83%, but specificity of 67% and correct classification of 74%. Notably the authors propose a shorter questionnaire that had similar sensitivity of 81%, a somewhat better specificity at 80% and 80% correct classification." (reference Bennett, R.M., Criteria for the Diagnosis of Fibromyalgia….Arthritis Care & Research (2014), 66 (9).)	Thank you for the suggestion. We added one statement in the 2nd paragraph of the ES Background (pg. ES-1) and report Background (p. 1): *Alternative diagnostic criteria are under consideration.[Bennett, 2014 et al.]*
TEP Reviewer 2	ES Background	Executive summary p. 8 line 55 and perhaps Appendix E8 & 9 need to mention the Dworkin criteria here. "Minimally clinically significant change for 10 point numerical rating scale for pain or BPI severity is 2 points and BPI interference is 1 point." Reference Dworkin, R.H. Interpreting the clinical importance of treatment outcomes…. (2008) The Journal of Pain 9(2). The back ground for the IMMPACT recommendations was that patients with chronic pain can function with minimal interference if their background pain severity drops by only 10%, though statistical significance may not be reached with such small changes due to variability and small sample sizes.	We did not change the text or citations in this area. The Reviewer's suggestion for the ES-Background text, and Appendix tables E8-E9, do not fit with the existing content in those areas. The IMMPACT recommendations are not specific to fibromyalgia, although the 2-point change in BPI average pain severity from Dworkin et al. 2008 is similar to what we cited for fibromyalgia in the Report, Table 8.

Commentator	Section	Comment	Response
Public Reviewer: American Physical Therapy Association	Executive Summary and Report: Background	Subgroups: Overall, we would suggest the following recommendations to improve the clarity and distinction of the proposed subgroups for the review. First, we are not clear on the rationale for identifying patients as complex. Both clinicians and researchers recognize chronic pain as a complex diagnosis which benefits from multidisciplinary treatments due to the multifaceted nature of pain and its impact on quality of life.	We use the term *complex* to mean FM with coexisting conditions (medical or mental health) or other non-FM factors that impact treatment decisions (such as older age). We edited text as needed throughout the ES and Report to clarify our use of the term *complex*, or replace it.
Public Reviewer: American Physical Therapy Association	Executive Summary and Report: Key Questions	We believe that a number of the subgroup classifications are poorly defined. For example obesity is not clearly defined by BMI parameters; likewise, for older adults, no age criteria was provided. The multiple medical comorbidities subgroup does not indicate a stratification for the number of co-morbidities. Additionally, some co-morbidities, like fibromyalgia, can have delayed diagnosis and can be confounding as a subgroup. The use of depression and anxiety as the primary variables in the categorization of mental health should be considered separately. Lastly, it is important to discuss how subjects who fell into multiple categories were handled and accounted for in the review.	Thank you for your points. Key Questions (ES and Report): The subgroups are not further specified in the Key Questions to enable us to include author-defined subgroups that met our general subgroup classifications. The study Inclusion and Exclusion criteria are listed in the Report Appendices E1-E3, which also list the FM diagnostic criteria used for each study. In general, the literature is much less specific in content and much sparser in number (of articles) than would accommodate what the Reviewer is requesting. Medical comorbidities, if included, were not specified as counts per patient, and counts would not provide any severity information about comorbid conditions. Depression and anxiety were separately reported in several studies (Appendix Tables E4-E7 of the report). We were unable to identify subjects who fell into more than one subgroup per study unless authors identified them as such.
Public Reviewer: American Physical Therapy Association	Executive Summary: Background- Scope and Key Questions	Key Questions: Efficacy and comparative effectiveness are two very different topics. It would seem to be more appropriate to have these as two different questions rather than combined into one.	The Key Questions were developed and refined with expert and other stakeholder input during the Topic Refinement phase of this project that preceded the systematic review. We are unable to modify the key questions (per AHRQ's guidance) after the topic refinement and public posting phase of the project that preceded this review.

Commentator	Section	Comment	Response
Public Reviewer: American Physical Therapy Association	Executive Summary and Report: Background - Key Questions and Analytic Framework	We have questions about the assumptions for the framework of the study: a. Do complex subgroups exist alongside simple subgroups? b. What would be the distinguishing factors between the subgroups? c. Why would treatment be different if pain is complex and treatment is multidisciplinary as a whole? d. What evidence is presented that would support the evaluation of undefined subgroups?	a. We use the term *complex* to mean FM with coexisting conditions (medical or mental health) or other non-FM factors that impact treatment decisions (such as older age). We edited the text throughout the ES and Report as needed to clarify or replace our use of the term *complex*. b. We used author-defined definitions of subgroup membership: See Appendix tables E1-E3 that detail the sample selection criteria used per included study c. The question is unclear. Pain is quantified in most studies (BPI average pain severity, FIQ pain, VAS, other) but patients can be complex due to other non-FM factors that impact their clinical presentation or treatment results. See a. above. d. Undefined means not defined *a priori* by the MN EPC and expert advisors, given the initial literature screen. Other subgroups were included as identified in the literature (by baseline lab values, coping styles, prior antidepressant use, etc.), which were not determined prior to the start of this review.
Public Reviewer: American Physical Therapy Association	Executive Summary and Report: Background	The need for the subgroups should be explained and why the specific variables were chosen to explore. This would greatly improve and clarify the framework of the study	The need for a subgroup study is detailed on ES-2 and page 4 of the report: *We limited this review to subgroup effects because McMaster University in Canada is currently conducting a comprehensive systematic review of randomized clinical trials on interventions for fibromyalgia in adults.*[17] *Our systematic review adds unique information by examining outcomes in fibromyalgia patient subgroups and by including observational literature.*

Commentator	Section	Comment	Response
Public Reviewer: American Physical Therapy Association	Executive Summary and Report: Background - Analytic Framework	The analytic framework for the study key questions includes men in the framework however men are not significantly addressed in the subgroups or in the article. Primary outcomes included pain, symptom improvement, function, participation, fatigue, sleep quality, and health-related quality of-life. However, how symptom improvement is different than the areas of pain, fatigue and function needs to be more clearly differentiated as pain and fatigue are both expressed as symptoms.	Consistent with the fibromyalgia patient population, men accounted for a small minority of patients in studies that included them (see Appendix tables E4-E7: Basic study information tables, and Appendix tables E1-E3: Sample selection criteria for included studies). Outcomes for men (vs. women) are discussed on page 18 of the report under Results by Sex. We added text to the Discussion-Applicability and Limitations of the Evidence Base: *Few men were included in clinical trials*.(ES-14, pg. 39)
Public Reviewer: American Physical Therapy Association	Executive Summary and Report: Background - Analytic Framework	Primary outcomes included pain, symptom improvement, function, participation, fatigue, sleep quality, and health-related quality of-life. However, how symptom improvement is different than the areas of pain, fatigue and function needs to be more clearly differentiated as pain and fatigue are both expressed as symptoms.	Outcomes: See Table 3. PICOTS framework, on page 6 of the report for definitions that accompany the Analytic Framework.
TEP Reviewer 3	Executive Summary: Methods	Page ES-4: Quality (Risk of Bias) Assessment of Individual Studies: The section was difficult to understand. Specific tools that the authors used to assess the bias are listed but how the item responses were consolidated into the ratings of high, moderate, and low is not clear.	ES-4, Methods: Quality (Risk of Bias) Assessment of Individual Studies: we added a statement : (see page 10 of the Report for more detail). A consolidating algorithm was not used.
TEP Reviewer 3	Executive Summary: Methods	Page ES-4: Quality (Risk of Bias) Assessment of Individual Studies: Appendix (C?) that illustrates the risk of bias for each study needs to be mentioned here.	ES-4, Methods: Quality (Risk of Bias) Assessment of Individual Studies: Added reference to (Appendix C) in the Risk of Bias section of ES and Report
TEP Reviewer 3	Executive Summary: Methods	Page ES-5: Strength of the Body of Evidence: It is not clear what this sentence means “Assessing reporting bias was not required’, or why the reporting was not required.	Deleted the statement in ES and Report-it was unnecessary to the section.
TEP Reviewer 3	Executive Summary: Methods	Applicability: The PICOTS framework was used to determine the applicability of the studies. It would be helpful to describe what the PICOTS framework entitles and what kinds of contextual information were specifically identified.	ES-5 Methods-Applicability: We edited the text (and on pg. 12, main Report) to explain the term PICOTS: *(Populations, Interventions, Comparators, Outcomes, Timing, Settings).* We clarified the ES Methods- Applicability text. We did not add the requested information in the Methods because Applicability of findings are covered in the Discussion, pg. 39 of the report.

Commentator	Section	Comment	Response
Public Reviewer: American Physical Therapy Association	Executive Summary and Report: Methods	Strength of Body of Evidence: This section had a very nice outline of the criteria used to evaluate the evidence. It would be more powerful to consistently address each the criterion in all areas reviewed. We realize there is limited space in such a large topic, however in each area it would improve the report to include the criteria consistently.	Methods: ES-5 and Report pg. 12: Strength of the Body of Evidence: The text is consistent with the requirements of AHRQ's CER Methods guidelines. No changes were made. (Risk of Bias elements for individual studies are provided in Appendix tables E10-E12)
Public Reviewer: American Physical Therapy Association	Executive Summary and Report: Methods	Eligibility: It would be helpful to have a better explanation of the phrase "possibly with a third adjudicator." It is more clearly stated later in the article: "by two independent investigators using instruments specific to each study design. The two investigators consulted to reconcile any discrepancies in overall risk of bias assessments and, when needed, a third investigator was consulted to reconcile the summary judgment." Again a clarification with a percentage of the total is needed here.	The Reviewer is referring to two separate sections in the ES: study eligibility and risk of bias assessment. We edited the study eligibility section to read: *Two independent investigators independently determined study eligibility, resolved disagreements through discussions; when needed, a third investigator was consulted until consensus was achieved.* We did not add a percentage to this section.
TEP Reviewer 2	ES Results	Physical treatments p. 17 lines 22-31 and Appendix Table E 10. I'm confused about why so few physical treatments made the inclusion into the ARHQ. There are over 150 exercise trials in fibromyalgia to date. A 2009 article reviews many more: "From 1988-2008, 70 exercise intervention in FM have been published with a total of 4385 subjects completing those studies" (ref Jones KD (2009) Exercise interventions in FM…Rheum Dis Clin N Am 35 (373-391). Similar concerns about lack of mind/body therapies (ref Mist, SD, Complementary and alternative exercise for fibromyalgia: a meta-analysis, 2013, Journal of Pain Research, 6 247-260). Langhorst published a similar meta-anlaysis the previous year.	Thank you for your feedback. See Table 3 PICOTS framework of the report Introduction that details the sample selection criteria for this review of treatment effects in subgroups of adults with fibromyalgia. Articles that did not report subgroup outcomes 3 months or more following initiation of a clinical trial intervention and/or lacked a comparison group were excluded.
TEP Reviewer 2	ES Results	Under psychological therapies, there are 5 mindfulness based stress reduction papers in fibromyalgia. Why were these not included?	Thank you for your feedback. No references were attached for this comment so we cannot provide specifics. See our response to the comment immediately above, and the report Introduction, Table 3- PICOTS framework, which contains the study selection criteria

Commentator	Section	Comment	Response
TEP Reviewer 3	Executive Summary: Discussion/ Conclusions	I think this is the area that further collaborations may be needed. I suggest that following issues to be considered: Page ES-13: Key Findings: Based upon the systematic review, it is concluded that "patient subgroups do not have differential treatment effects". Although the data can technically lead to this resultant comment, further elaborations on this point seem warranted particularly given the limited range of data available. One of the important aspects of the subgrouping may be the nature of the patient divisions, in terms of whether the subgrouping factors can be the target of the treatment. Demographic variables, for example, may be used to subgroup patients but those cannot be the treatment target. On the other hand, there are modifiable variables that can subgroup patients, such as depression and weight status. Other factors could be the presence or extent of symptom severity (pain, fatigue etc.). The question, then, should also include whether the treatment was targeting to modify these subgrouping factors, in addition to delineating whether subgroups responded differentially.	Thank you for your feedback. The Key Findings ES-11 (also Discussion-p. 38, Conclusions) have been edited to better delineate where low and insufficient SoE are for subgroups. Given the limited number of studies that met inclusion criteria, we did not further subgroup the findings to attempt to classify stated or implied treatment targets.
TEP Reviewer 3	Executive Summary: Discussion/ Conclusions	Page ES-14: Applicability and Limitations of the Evidence Base: It is correctly pointed out that the most study participants were middle-aged while female samples. It may be helpful to pint out from the available epidemiological data whether this is relatively representative of the FMS population at large or the other groups (non-white, males) are underrepresented. It is also likely that the study participants were mostly recruited at the urban university based hospital/research institutes, thus over-representing urban FMS patients.	We added the italicized text and references to ES-12 and Report pg. 39: Study patients were largely middle-aged white females with moderate to severe fibromyalgia symptoms at baseline as measured by the FIQ, *which is generally representative of the fibromyalgia patient population seen in clinical practice in the US.[Gore, 2012][Vincent, 2013]. Few men were included in clinical trials.* We did not add text regarding urban vs. rural patients because the distinction was not specifically addressed in most studies.
Peer Reviewer 4	Executive Summary, Background	The background is lengthy but there is not an explicit statement of the rationale for examining subgroup effects. The first sentence under Selection of Patient Subgroups on page 3 would suffice as a rationale.	ES-1, ES-2: Added statement of rationale for examining subgroups to background in Executive Summary as stated in Selection of Patient Subgroups, page 3.
TEP Reviewer 7	ES Introduction	Introduction: ES-1 line 30: Might rephrase as "fibromyalgia has no clearly identified etiology"	Thank you. We edited the statement in ES-1 and the Report Background to: *Fibromyalgia is a chronic, diffuse musculoskeletal pain syndrome that has no clearly identified etiology.*

Commentator	Section	Comment	Response
TEP Reviewer 7	ES Methods	ES Methods: although the main report discusses how subgroup effects were assessed in individual studies, this is missing from the ES methods, but seems critical since this is the focus of the review. I also think that the Synthesis section needs to explicitly discuss issues of particular relevance to assessing subgroup effects e.g. direct (within study) vs. indirect (between study) assessments. I don't really see this in either the ES or full report methods.	See ES-4 and Report pg. 13: We added text to both areas to state that individual RCT subgroup comparisons were within-study (direct).
Peer Reviewer 1	Introduction	While the introduction is generally well written, FM is not a pain syndrome of "unknown etiology". Some of the best mechanistic studies in the pain field have been conducted in FM. We know as much about FM as we do other pain states (which many be suboptimal, but FM is not a special case in this regard).	We edited the statement in the ES and Report and added references: *Fibromyalgia is a chronic, diffuse musculoskeletal pain syndrome that has no clearly identified etiology.*
Peer Reviewer 1	Introduction	The authors used a panel approach to identify patient subgroups. The literature on the other hand, has used empirical approaches (e.g., cluster analytic) approaches for identifying clinically meaningful subgroups of individuals with FM. It is unfortunate that these existing subgroupings were not used in this review.	Thank you for your comment. We clarified the Methods text (Report pg. 3 and ES-2: Scope) that the subgroups were determined *a priori* with expert input based on our preliminary literature scan. It is unclear from the Reviewer's comment how the subgroups we studied differed from those determined exclusively from empirical approaches in the select articles the Reviewer is referring to. From the literature we reviewed, there is likely considerable overlap in subgroup identification. However, it is unclear which articles the Reviewer is referring to because no references were provided.
Peer Reviewer 1	Introduction	To date, most studies have examined FM as a diagnostic class given there has been no compelling rationale for sampling FM in accordance with the subgroupings outlined in this review. It is a bit disingenuous to criticize the FM literature for not subgrouping based upon the categories that were just made up by the paper's "expert" panel.	Thank you for your feedback. The subgroups were determined from the literature and with expert and other stakeholder input. Our Discussion statements regarding the existing literature summarize knowledge gaps and generalizability issues to highlight areas where further investigation and more transparent reporting could advance the field.
Peer Reviewer 1	Introduction	The authors state that there is a strong belief that treatment effects of FM treatment vary by subgroup (ES-15:9). Who believes this? No citation.	Discussion-Research Gaps (ES-11 and Report pg. 38): We added references to support our statement.

Commentator	Section	Comment	Response
Peer Reviewer 2	Introduction	I thought the tone and balance of the introduction were excellent. It well describes the state of the field and how it has changed over the decades while well-acknowledging that much remains unknown.	Thank you for your compliment
TEP Reviewer 1	Introduction	The intro begins on page 27 (preceded by a long executive summary). Like the rest of the document, the intro is well-written and organized, very thorough description of relevant issues	Thank you for your compliment. The ES has been shortened.
TEP Reviewer 2	Introduction	see attached document reviewing entire paper. Mostly concerned that a newer FM diagnostic criteria is gaining traction and should be included (Bennett 2014). It has similar sensitivity but superior specificity to the Wolfe 2010 proposed FM criteria	Added statement in Introduction (ES and report): *Alternative diagnostic criteria are under consideration.(Bennett, RM et al. 2014)*
TEP Reviewer 3	Introduction	Nicely organized.	Thank you for your compliment.
TEP Reviewer 4	Introduction	The Introduction is appropriate, and there are no recommendations for improvement.	Thank you
TEP Reviewer 5	Introduction	Excellent	Thank you for your compliment.
TEP Reviewer 6	Introduction	A good case is made for the importance of finding effective treatment options for patients with fibromyalgia, and that although overall treatment effects have been modest, it remains possible that certain subgroups may achieve more, or less, benefit.	We agree – thank you
Peer Reviewer 4	Introduction	The introduction seems reasonable. If the results had found more information on population subgroups, I would want the introduction to consider how to address subgroups that overlap (e.g., women with depression) as some of these subgroup factors tend to co-exist in fibromyalgia. However, given that there was so little evidence on subgroups, this is probably not important.	Thank you for your feedback.
TEP Reviewer 7	Introduction	Background: Might be useful to note that understanding subgroup effects might help inform treatment decisions--e.g. duloxetine is an antidepressant, so it's possible that it has better effects in patients with fibromyalgia who are also depressed.	Thank you for your suggestion. We added text of the same to the Report pg. 3 and ES-1: *Understanding subgroup effects might help to inform clinical treatment decisions.*
TEP Reviewer 7	Introduction	Background: I'd like a little more detail/data to support how these subgroups were selected. E.g. are women a high prevalence subgroup? Is there data suggesting that FM is more difficult to treat in obese patients or in older patients?	Added text and references to the Report pg. 3-4 under Rationale for Review and Selection of Patient Subgroups

Commentator	Section	Comment	Response
Public Reviewer: Consumer/patient	Introduction	Were all of the patients identified in the RCT's clinically diagnosed, by a physician, with Fibromyalgia? Most studies about Fibromyalgia have the same or similar outcomes, etiology unknown, chronic pain anxiety or depression and Rheumatoid arthritis. The etiology will remain unknown until the clinical studies involve mental Health Specialists. First and foremost I believe that the diagnosis of fibromyalgia should be confirmed by both specialists, the medical Doctor as well as the psychiatrist	Thank you for your feedback. The fibromyalgia diagnostic criteria used per study to determine subject inclusion are listed in Appendix tables E1-E3. Additional sample selection criteria varied per study (see Basic Study tables, Appendices E4-E7)
Peer Reviewer 1	Methods	Overall the methods were standard and justified; but as stated, rigorous methods were applied to a questions for which existing studies are non-existent. Thus an insufficient number of studies were available to answer the varied and diverse questions posed by the study team. This review probably should not have been completed after the first step identified a dearth of data to address the key questions.	Thank you for your comment. One major value of this systematic review is that it highlights the paucity of data available for clinicians on fibromyalgia treatment effects in subgroups, and identifies specific research gaps that future research on fibromyalgia will hopefully address.
Peer Reviewer 1	Methods	The methods would have been appropriate if there had there been sufficient data to analyze. This paper reflects lots of rigor but little data and few meaningful conclusions.	Thank you for your feedback.
Peer Reviewer 1	Methods	The approach to bias assessment should not be universally applied to pharmacological and non-pharmacological studies. For example, one cannot have a patient blinded in a psychological study. Thus it is unfair to rate all psychological studies as having high bias when there are not meaningful alternatives to patient blinding for psychological studies.	Psychological studies were high risk of bias for a number of reasons, not just inability or lack of blinding. We edited the psychological entries in Appendix table E10 to read *blinding not possible,* where applicable, and better stated the methodological shortcomings that account for the Risk of Bias ratings.
Peer Reviewer 2	Methods	I think all were very appropriate. The authors were very thorough and thoughtful in their approach. The inclusion/exclusion criteria were appropriate. The search strategies were explicitly states and logical. The use of all known FMS criteria and standard FMS outcome measures were appropriate. It was not possible to perform a meta-analysis due to the paucity of data.	Thank you for your feedback.
TEP Reviewer 2	Methods	Authors used appropriate search strategies and selection criteria for studies to review, include, and exclude.	Thank you

Commentator	Section	Comment	Response
TEP Reviewer 2	Methods	I am not clear on why several non-pharm studies were exluded. I review these in the attached document. As the paper reads now, third party payers may believe that only drugs should be covered based on the paucity of evidence included for non-pharm therapies.	Thank you for your feedback. We have listed and addressed each comment that was contained in the additional document provided by this reviewer below. Studies that did not meet selection criteria were excluded. See Table 3- PICOTS framework- of the report Introduction that details the sample selection criteria.
TEP Reviewer 3	Methods	In general, the approaches are well delineated and described, including inclusion/exclusion criteria, search strategies, and outcome measures.	Thank you
TEP Reviewer 4	Methods	The Methods are largely appropriate and adequately described. However, due to limitations of the literature, the decision was made not to conduct a formal meta-analysis, and rather they chose to conduct a qualitative synthesis and analysis. This analytic plan is substantially weaker and undermines confidence in the findings and conclusions. Further limiting the value of the report is the fact that, once having determined that a meta-analysis could not be conducted, the authors apparently did not consider relaxing the criteria for inclusion and exclusion of specific studies, potentially further limiting the search for studies that could contribute relatively valuable information for this qualitative review. It is also somewhat disappointing that a common Nursing database, CINAHL, was not included in the search.	Thank you for your comments. Given the number of subgroup-treatment-outcomes combinations and relatively sparse literature base, a meta-analysis was planned but not possible. We agree that qualitative syntheses provide results based on less rigorous methodologies than quantitative summaries. Study selection criteria were decided *a priori* with a literature scan and expert input to best evaluate subgroup treatment effects for FM as a chronic condition. The limited number of eligible studies highlights several important research knowledge gaps that we have identified in the Discussion- that many FM studies evaluate only short (< 12 wk.) outcomes which is of limited utility for this chronic condition, and that subgroup treatment effects have not been a dominant focus of research to date. We added two references to the Methods-Literature Search Strategy (Report pg. 8) that support our decision not to search CINAHL in addition to five other databases for this review of FDA-approved and US-available treatment effects in subgroups.
TEP Reviewer 5	Methods	I believe that the inclusion criteria are justifiable. I was disappointed to see that very few non-pharmacologic studies qualified for the review, but suspect that this is due to the lack of rigor in those studies.	Studies were excluded if they lacked subgroup outcomes reporting at 12 weeks or beyond and did not report subgroup outcomes. See Table 3 PICOTS framework of the Report for study selection criteria

Commentator	Section	Comment	Response
TEP Reviewer 6	Methods	The eligibility criteria were reasonable, but it would have been ideal to include non-English studies. We have found that approximately 11% of interventional trials for fibromyalgia are published in non-English languages (1). *Busse JW, Bruno P, Malik K, et al. An efficient strategy allowed English-speaking reviewers to identify foreign-language articles that met eligibility criteria for a systematic review of management for fibromyalgia. Journal of Clinical Epidemiology. 2014; 67(5): 547-53.*	Thank you for your suggestion. Our search strings were not limited in language. Under funding from AHRQ, our review is limited to treatments that have FDA-approval or are available in the US, which differs from the article selection criteria for the ongoing systematic review of controlled trials at McMaster University, the topic of the referenced screening article. Given that limitation, we expect that the overwhelming majority of treatments used in the US for FM are examined in English-language articles. We added citations (Report Methods, pg. 8 and ES-3) that support using an English-language restriction, especially for conventional medical therapies that are most commonly employed for FM.
TEP Reviewer 6	Methods	The authors, correctly, completed study selection and data abstraction in duplicate; however, it would have been helpful to report measures of agreement (e.g. phi or kappa) at least for selection of eligible studies.	Thank you for your suggestion. The review was conducted in accordance with AHRQ's CER methods guidance.
TEP Reviewer 6	Methods	The authors declare their subgroups of interest a priori, which is appropriate; however, they do not appear to pre-specify the anticipated direction of effect for each factor which is also recommended by the paper by Sun et al. that they reference.	Thank you. The Sun et al.2010 article suggests that RCT investigators who plan to conduct subgroup analyses should pre-specify the direction of anticipated effect for their subgroups. We did not pre-specify a direction of effect for this systematic review.
TEP Reviewer 6	Methods	The authors were not able to statistically analyze their data.	Added to the Data Synthesis section: *However, a quantitative analysis with pooling of data was not possible due to the relatively few studies and large number of subgroup-treatment-outcome combinations.*

Commentator	Section	Comment	Response
Peer Reviewer 4	Methods	It seems that the review intends to address both 1) the comparative effectiveness of particular treatments for fibromyalgia in subgroups (e.g., patients with fibromyalgia and MDD), and 2) the comparative effectiveness of particular treatments for fibromyalgia in these subgroups, compared to the comparative effectiveness of particular treatments for fibromyalgia in patients not in these subgroups (e.g., without MDD). These two questions are comingled throughout the report, and it is sometimes confusing to follow which question is being addressed. I am not sure that the second question directly addresses the review question of interest, but in some situations a statement comparing treatment effects for persons within the subgroup and persons outside the subgroup, or comparing treatment effects for persons within the subgroup and treatment effects for the overall set of study participants, may be the only available piece of information upon treatment effectiveness. Perhaps this could be clarified in the methods, and reported more clearly in the results?	We added text to the Methods-Data Synthesis section (pg. 11 Report; ES-4) *Wherever possible, we report interaction results that assessed whether treatment effects varied in subgroups. If interaction results were not reported, we included within stratum (stratum-specific) results (such as change from baseline pain score in those with MDD (treated vs. control) and in those without MDD (treated vs. control) when data were presented.* We edited the text throughout the Report and ES to be more consistent in this distinction.
Peer Reviewer 4	Methods	Data Synthesis, did the authors consider calculating SMD for outcomes that measured the same underlying concept but used different instruments? This wasn't clear.	Our Methods text focus is on what was done for this review. Standardized mean differences were not relevant with few studies and limited available subgroup comparisons for this review.
TEP Reviewer 7	Methods-Applicability	Methods-Applicability: The methods don't really discuss important issues that might affect applicability. I also don't really understand the sentence "...this would not limit applicability but rather limit the number of studies with adequate subgroup inclusion and reporting." I think including higher functioning/less impaired patients impacts applicability (e.g. in analyses of older patients or women) as well as the ability to look at lower functioning/more impaired subgroups)	Thank you for your suggestion. We revised the Applicability text (ES-5 and Report pg. 12).
TEP Reviewer 7	Methods	The "pure subgroup" studies really only provide data for between study comparisons which I would consider "indirect" evidence; the within-study comparisons are what I would consider "direct" evidence. Might be good to state this explicitly in the methods.	Thank you for your suggestion. We added statements of such in the Methods under Inclusion criteria (Report pg. 9, 13)

Commentator	Section	Comment	Response
TEP Reviewer 7	Methods	Would like a little more discussion of how the IPD studies are handled; many such studies are funded by pharmaceutical companies and are often quite selective in which studies they include, how they analyze the data (sometimes they just lump data across studies versus basing analyses on pooled within-study analyses), and which outcomes are reported. Also, how did you handle situations where a study was included in an IPD but also had results reported separately? Wouldn't want to double count these studies.	The pooled IPD analyses were pooled within-study comparisons. We edited the Methods pg. 10 and added Appendix Table E17. We avoided double counting by examining non-overlapping information from 4 RCTS that were used in 3 pooled analyses included in this review. Appendix Table E17 shows differing RCT versus pooled analysis features that provide rationale for our inclusion of 4 RCTs and three pooled analyses of those RCTs in this review. These features include differences in outcome measures, outcomes timing, drug dosages (single dose vs. aggregated several doses for pooled analysis into one group), and omitted treatment groups in pooled analyses. The differences were not readily determined from nontransparent Methods reporting in pooled analyses, but are more apparent in Appendix Table E17. The actual number of patients included in the 3 pooled analyses was not directly determinable; only patients *randomized* in the input RCTs were reported in the pooled analyses. We estimated the actual number of patients that were available for these 3 pooled analyses from the attrition in the input RCTs (Appendix E4), which was approximately 56.5% of the reported number of patients in 3 pooled analyses (see footnote, Appendix table E17)
Peer Reviewer 1	Results	Little could be said for many of the questions. Looking only at the psychological treatments, only 4 studies could be discussed. Yet there were 8 subgroups of interest x two key questions. The authors just didn't have a large enough literature available to address all of the questions that they posed.	Thank you for your comment.

Commentator	Section	Comment	Response
Peer Reviewer 1	Results	If these key questions are truely important, a case for answering the questions should be made and the lack of data discussed. Trying to review the few studies that partially examined some aspects of the key questions does not come across as a credible review and certainly the conclusions about the equivicality of need for more intensive care for subgroups cannot be reliably discussed.	Thank you for your comment.
Peer Reviewer 1	Results	At times, the authors state conclusions not about differences between subgroups but about the efficacy of specific treatments based upon their review sample. This is wholey inappropriate given the rather small and unrepresentativeness of the literature that met the selection criteria for this review.	Thank you for your comment.
Peer Reviewer 2	Results	I thought that the results section was very thorough and detailed. Figures and tables were more than adequate. I did not note any studies that were missed or believe there are any that warrant inclusion.	Thank you
TEP Reviewer 2	Results	The detail presented is extensive, with the justification for all decisions clearly spelled out. Figures/tables/appendices are appropriate, perhaps more than necessary.	Thank you
TEP Reviewer 2	Results	several key studies, especially r/t tramadol and sodium oxybate were excluded. I reference these in my attached document.	Thank you for your feedback. Studies without subgroup outcomes reporting were excluded. See Table 3- PICOTS framework- of the report Introduction that details the sample selection criteria.
TEP Reviewer 2	Results	Table 5 p 50. Please consider including tramadol and sodium oxybate in this table. There were 2 phase III clinical trials in tramadol with and without acetaminophen (Refererences Bennett RM and Russell IJ). The company at the time (Ortho) decided for political reasons to try for an indication for pain rather than fibromyalgia, but the FDA approved the drug on 3 studies, 2 of which were in fibromyalgia.	Thank you for your feedback. Studies that lacked subgroup outcomes reporting, or otherwise did not meet selection criteria were excluded. See Table 3- PICOTS framework- of the report Introduction that details the sample selection criteria.
TEP Reviewer 3	Results	The results are presented in the well organized fashion addressing each Key Points. Tables are busy but it is the nature of the systematic reviews such as this one. It is not clear why some subgrouping categories such as anxiety and weight, are not included in Table B.	Thank you for your suggestion. We added a footnote to Table 12 (Report) and Table B (ES) of the same to clarify: **Table 12 shows strength of evidence for subgroup-treatment-outcomes combinations with at least two relevant studies. Other comparisons that had insufficient evidence (addressed by single studies that had high risk of bias and small sample sizes) are not shown.*

Commentator	Section	Comment	Response
TEP Reviewer 4	Results	The Results are appropriately presented, including level of details of the studies, their characteristics, and so forth. Tables, figures, and appendices are appropriate. Although I may be biased, and although I haven't conducted my own review, I wondered if studies of psychological interventions were adequately identified. Finally, I wonder if including CINAHL would have identified additional studies.	We appreciate your concern. Most studies were excluded due to lack of subgroup outcomes reporting or due to outcomes of less than 12 weeks from treatment initiation. We added two references to the Methods-Literature Search Strategy (Report pg. 8) that support our decision not to search CINAHL in addition to five other databases used for this review of FDA-approved, US-available subgroup treatment effects.
TEP Reviewer 5	Results	The amount of detail is adequate. A recent review by Daniel Clauw suggested that there is strong evidence for education, cognitive behavioral therapy and exercise (Clauw, DJ (2014) Fibromyalgia: A Clinical Review. JAMA, 311 (5), pp. 1547-55) but his methods were not as rigorous as those used for this report.	Thank you for your feedback. We read the mentioned article prior to the posting of this report draft.
TEP Reviewer 6	Results	The results are presented with considerable details in a number of tables.	Thank you for your feedback.
TEP Reviewer 6	Results	The authors attempt to contextualize the magnitude of reported treatment effects by comparing them to established minimally important difference (MID) thresholds. Although it is tempting to conclude that mean differences less than the MID are not worthwhile, and mean difference exceeding the MID suggest that most or all patients will benefit from treatment, this conclusion is misguided. Consider an example where the MID is 0.50 and patients mean improvement vs. control is 0.25. This could mean that 75% had no improvement and 25% experienced a mean change of 1.0, which would result in a NNT of 4, a clearly important benefit.	Thank you for your suggestion. We corrected the Results text pg. 16. We deleted one column from Table 8 that had erroneously reported MCID for a difference across groups. We edited the text of the Results (pg. 17) and Discussion accordingly

Commentator	Section	Comment	Response
TEP Reviewer 6	Results	Authors may wish to knowledge English-language bias as a limitation, as well as challenges associated with assessing the credibility of subgroup analyses without access to study protocols. We have recently found that a third of trials that claimed to have pre-specified their subgroup analyses were not supported by the corresponding protocol (2). *Kasenda B, Sun X, von Elm E, Schandelmaier S, et al. Subgroup analyses in randomised controlled trials: cohort study on trial protocols and journal publications. BMJ. 2014; 349: g4539.*	We added text and references to address both points under Limitations of the Comparative Effectiveness Review Process, pg. 40 (and ES-13): *In assessing subgroup pre-specification for included studies, we relied on information in each article, which may overstate the actual number of subgroups that were determined a priori in randomized clinical trials. (Kasenda, 2014).* *This review was limited to English-language publications. The possibility of missing clinical trials with subgroup reporting for treatments that were FDA-approved and/or available in the US with this restriction is remote, especially for conventional medical therapies. (Moher, 2000) (Morrison, 2012)(Pham, 2005)*

Commentator	Section	Comment	Response
Peer Reviewer 4	Results	I found the Results section to be very difficult to follow, and it took more effort to understand and mentally organize the results than I would have expected from such a limited number of studies. Some of this may be because of the multiple types of studies, the multiple treatments considered, and the range of subgroups examined, but there may still be room for improved clarity and organization. In terms of clarity, sometimes it was not clear what outcome(s), subgroup(s), or comparators were being referred to in the text. Some minor examples: The heading of the second column of Table 12 is 'Intervention/Comparator' but no comparators are listed in the column. 'Other Subgroup Outcomes' on page 19 does not specify the subgroup or subgroups for which within-subgroup changes were reported by Bhadra et al. and although there is a reference to Table 6, a statement that within-group changes 'for multiple other subgroups' were assessed would be helpful for the reader. On page 21 the description of the observational biofeedback study does not specify the comparator.	Thank you for your feedback. We did not change the main structure of the Results (by main categories of treatments, with subgroups listed within those headings). While we considered using subgroups as headings, there were too many "empty cells" (no comparisons) that would have required additional listing of such. WE rearranged subgroups in the Results under each class of treatment to present higher SoE studies first. We rearranged Results information within treatment subheadings so that comparisons with higher strength of evidence are presented first. Table 12 (and ES Table B): We added the comparator to each row of the table. We added text to pg. 19 re: Bhadra et al. as per your suggestions under Other Subgroup Outcomes: We edited the text for the observational biofeedback study (pg. 20): *One high risk of bias observational biofeedback study examined the benefits of using EMG-reduction training of visual and auditory feedback to teach two groups of subjects the same muscle relaxation techniques. (Drexler 2002)*
Peer Reviewer 4	Results	I found Tables 6-11 to be very long, relative to the amount of useful information in them, and would be fine with putting them in the Appendices instead of including them in the Results.	These are the main Results tables; we left them in the Report
Peer Reviewer 4	Results	In terms of organization, I would have preferred the organization of Results for KQ1 to be by subgroup rather than by type of therapy, as it seems more clinically intuitive to begin with considering the patient rather than the therapy, however I realize that everything is organized this way, including tables, and this is probably an unrealistic change to suggest.	Thank you for your suggestion. We did not change the report format.

Commentator	Section	Comment	Response
Peer Reviewer 4	Results	For KQ2, it makes more sense to consider reporting by types of treatment, and I am not sure how useful it was to distinguish between information from pure-subgroup, mixed sample, pooled IPD and observational studies.	Thank you for your suggestion. We did not change the format, given the brevity of this section (4 studies). The format clearly shows the limited AE information, especially from drug RCTs that examined subgroup interaction effects.
TEP Reviewer 7	Results	The Key Points are a mixture of key results as well as bullet points with descriptions of methodological shortcomings etc. SOE ratings are not always clearly presented when relevant (for key results) and it seems that many of the descriptions of methodological shortcomings could be incorporated into the bullet points with key results (i.e., noting that conclusions are based on studies with methodological shortcomings etc.). It also looks like blanket ratings of "low SOE" were given even though the evidence seems to be insufficient for some subgroups and low for others. It may also be appropriate to point out as a key finding that evidence is largely insufficient to determine subgroup effects for interventions basically other than duloxetine.	Thank you for your feedback. The Results-Key Points, have been edited in Report pg. 15 and ES. We reviewed our information and edited two items in Table 12 (Table B ES) with inconsistent findings, and changed the ratings from low to *insufficient*: both for duloxetine effects on BPI average pain (sex and race subgroups). We added a footnote to Table 12 (Table B, ES): **Table 12 shows strength of evidence for subgroup-treatment-outcomes combinations with at least two relevant studies. Other comparisons that had insufficient evidence (addressed by single studies that had high risk of bias and small sample sizes) are not shown.* We added the following to the Key Points and Conclusion: *The fibromyalgia evidence is largely insufficient to determine subgroup effects for interventions basically other than duloxetine.*
TEP Reviewer 7	Results	There should be more detail about the pooled IPD analyses--it is unclear if the estimates/results from IPD studies were based on pooled within-study subgroup comparisons or if the IPD analyses pooled subgroup effects across studies (e.g., took all patients with depression from all the studies and pooled those results, and separately pooled results for all patients without depression). The former would be considered more reliable for interpreting subgroup effects.	We clarified the text in the Methods-Type and Labeling of Included Studies (pg. 14 and ES-6): *…, which were pooled within-study comparisons.*
Public Reviewer: Consumer/patient	Results	Although the FDA has approved Lyrica and Cymbalta for the treatment of Fibromyalgia, they do not work, nor does either medication relieve pain, these medications do cause fog brain and do nothing to stimulate the patient into self help.	Thank you for your opinion.

Commentator	Section	Comment	Response
Public Reviewer: American Physical Therapy Association	Results	Review of Trials Overall: It would be helpful to discuss the number of subjects that the trial contained which lends power to the conclusions drawn by the authors. For example, a trial with 250 subjects compared to a trial of 16 subjects would increase the power of the conclusions and the effectiveness of the outcome of the study.	The number of adults who were randomized (or enrolled at baseline) per included study is provided in Appendix Tables E4-E7, (overall and for each subgroup). Sample size and related factors (such as attrition), are taken into consideration during different stages of the review process, including risk of bias assessment (Appendix Tables E10-E12). The summary information from the entire review process is provided in Table 12- Strength of Evidence, which gives a measure of confidence in the review findings for various subgroup-treatment-outcome combinations. Attrition was high and generally poorly reported, especially for subgroups, which impacts power.
Peer Reviewer 1	Discussion/ Conclusions	The discussion is a bit repetitious perhaps due to the fact that there is very little that can be said about these data.	Thank you for your feedback. The Discussion has been edited.
Peer Reviewer 1	Discussion/ Conclusions	The actual FM literature review alluded to as being conducted at McMaster University would seem to be the more relevant review of the FM treatment literature.	Thank you for your opinion.
Peer Reviewer 1	Discussion/ Conclusions	The conclusions are of insufficient quality to inform policy or practice.	Thank you for your opinion.
Peer Reviewer 2	Discussion	I think the discussion is well done. The lack of good data makes it impossible to come to any strong conclusions. I think the report makes it clear that future research needs to consider subgroups - as it has really not done so to date.	Thank you for the compliment
TEP Reviewer 2	Discussion/ Conclusions	Conclusions are clearly and comprehensively presented and well defended. Research gaps are clear, based on findings of this review.	Thank you for the compliment
TEP Reviewer 2	Discussion/ Conclusions	Future research needs to include 2+ measures for outcomes (Dworkin criteria) and BPI pain interference over BPI pain severity	Thank you for your suggestion. We did not add this reference on how best to measure outcomes changes in chronic pain since that is not the topic of this review.
TEP Reviewer 4	Discussion/ Conclusions	I think that the conclusions are reasonable, although an alternative would be to acknowledge the limitations of the available empirical literature as well as the limitations of the qualitative methods of analysis, leading to a conclusion that the key questions could not be answered.	Thank you for your feedback.

Commentator	Section	Comment	Response
TEP Reviewer 4	Discussion/ Conclusions	The section on research gaps is inadequate. The call for well powered RCTs and other prospective studies to examine subgroups and specific fibromyalgia treatments is unrealistic given multiple restrictions of funding and other feasibility issues. The authors should consider a call for authors of RCTs to conduct post-hoc analyses even with limited power. The use of observational methods to examine existing EHR data (e.g., Kaiser, VA, other integrated healthcare systems) or emerging registry data should be encouraged. Finally, consistent with an acknowledgement of high rates of concurrent pain conditions and a growing interest in central mechanisms that may underlie fibromyalgia and other pain conditions (e.g., headache, TMD, IBS), the authors may suggest that future research consider examination of persons with multiple conditions rather focus on fibromyalgia in isolation.	Thank you for your suggestions. Our initial statement on clinical trials was not restricted to RCTs; therefore, we did not change the text in this area (p.41): *Transparently-reported, sufficiently powered clinical studies with a priori subgroup and hypothesis specifications were lacking.* We disagree with the addition of a statement to call for more underpowered studies; such investigations already dominate this literature set. We added text on multimorbid FM patients to the Research gaps on pg. 41. We appreciate this suggestion from you and several reviewers. We added Discussion text on the use of EHR data on pg. 41.
TEP Reviewer 5	Discussion/ Conclusions	This section is very good. my only comment would be to clarify what is meant by functional outcomes, does that include cognitive function, participation, as well as physical function? The work being done by the Omeract group might help to inform an expansion of this part of the discussion. http://www.omeract.org/ Additionally, it might be important to mention the challenges of measuring physical function in older groups--I have used the Late Life Function and Disability Index to capture a more accurate measure of physical function in my research of older adults living with FM for this reason.	Thank you for your suggestions. We added text to the Discussion about function and better defined it. We did not add a discussion of functional measures which is beyond the scope of this project.
TEP Reviewer 6	Discussion/ Conclusions	Discussion/ Conclusion: The findings are clearly reported. The authors do point out the limitations of the current literature, but there are no declarative statements regarding specific recommendations for future research. Such recommendations would be helpful.	Thank you for your suggestion. We edited the Discussion and added some areas for future research.

Commentator	Section	Comment	Response
Peer Reviewer 4	Discussion/ Conclusions	Is there evidence for no differential treatment effects or is there no evidence for differential treatment effects? Even though it is stated that the evidence is limited and low in strength, I think that stating there is evidence for no differential treatment effects may be overstating the conclusions. Also, to the extent that there is evidence, isn't it primarily for patients with MDD, not for subgroups generally, and not for complex patient subgroups (as stated in the Conclusions of the Abstract)? It should also be noted that evidence is only on the short term (e.g., not 1-2 years of treatment). While this is mentioned under Applicability and Limitations of the Evidence Base, it should be included in the overall conclusion as well. Finally, it should be noted that the greatest amount of evidence is upon SNRIs versus placebo, and this is a question of efficacy more than effectiveness, as placebo is not a plausible real-world treatment.	We edited the Discussion and Key Findings pg. 15 text for clarity. We emphasized that evidence is overwhelmingly short term (Report pg. 15). The Discussion already stated that evidence is largely short term. (pg. 39: *Outcomes are overwhelmingly reported for short term not long term outcomes, the latter of which is of greatest interest in the management of chronic fibromyalgia syndrome*. Also, short term evidence was mentioned on pg. 41 of Research Gaps.
Peer Reviewer 4	Discussion/ Conclusions	The Key Findings and Strength of Evidence is easier to follow than the Results, because it organizes the findings by subgroup when summarizing the findings. It might be worth mentioning that meta-analysis was not possible (and why) because the listing of findings from multiple studies (e.g., three studies showed no differences...) seems to allow a meta-analytic approach.	Added text to the Methods pg. 11 and Discussion pg. 40: *A meta-analysis was not possible due to relatively few studies and the large number of subgroup-treatment-outcome combinations.*
Peer Reviewer 4	Discussion/ Conclusions	Minor point: The paragraph beginning "For pain," (line 6, p. 39) should begin "For depression,".	Thank you. We edited the statement in the Discussion to, *For depression,….*
Peer Reviewer 4	Discussion/ Conclusions	The Research Gaps do list all the gaps in the research base, however there is no suggestion as to which subgroups might be most important to investigate. Perhaps this is beyond the scope of the report, but bringing in information about the most common comorbidities or most difficult to treat types of patients into the discussion might be helpful in suggesting the most useful next steps.	We added a paragraph to the Discussion, Research Gaps (p. 41) regarding the need for evidence for multimorbid FM patients.
TEP Reviewer 7	Discussion/ Conclusions	I think the findings are generally sound but the paucity of evidence makes it difficult to translate clinically, e.g. should clinicians preferentially use SNRI's in patients with FM plus depression? There clearly needs to be more evidence for interventions other than duloxetine and for many of the subgroups.	Added to the Discussion – Research Gaps: pg. 41 and ES: *There is a clear need for more evidence for interventions other than duloxetine, and for many of the fibromyalgia subgroups, to better inform clinical decisionmaking and improve patient outcomes and management.*

Commentator	Section	Comment	Response
Public Comment: Consumer/patient	Discussion/ Conclusions	A comparative non-biased RCT study will result in these findings. There is another concept regarding the treatment of Fibromyalgia. In my opinion, there is a very strong link between clinical depression and the diagnosis of fibromyalgia. Why treat the symptoms such as pain, anxiety or fog brain and not treat the underlying cause related to behavioral health. Clinical trials should be considered to determine if the treatment of fibromyalgia would best be approached using a serotonin uptake inhibitor, an antianxiety medication, and an over the counter medication for pain. This combination of treatments for fibromyalgia has worked for me since 2002. The healing process in terms of feeling sorry for yourself and time for the serotonin uptake inhibitor to take effect, as well as the self help components, is a very slow process. In addition to my initial concept about the causes of fibromyalgia, it is my firm belief and opinion that the cause is also linked to stress in any form. This methodology has worked for me since 2002. I do not take narcotics only across the counter pain relievers and that is not everyday. I do not have fog brain as experienced with both Lyrica and Cymbalta. I have never seen a Rheumatologist, however that is my next step to better recovery and improved self help. As I age the symptoms of stiffness, decreased mobility of one arm does cause some concern. Exercise does help. I'm currently working on my doctorate in nursing, as I plan to return to work. I hope this information will assist you in future studies.	Thank you for your specific feedback and suggestions. Some of the articles in this review (pooled analyses, few RCTs) of duloxetine discuss the same issue: trying to determine how much of the effects of duloxetine on fibromyalgia outcomes are due to its effects on depression vs. fibromyalgia – or both.
Peer Reviewer 1	Clarity and usability	The structure is fine and organized.	Thank you
Peer Reviewer 2	Clarity and Usability	This report is well structured and organized. The content will be informative to policy and make it clear that there is no data to support some clinical therapeutic beliefs.	Thank you
TEP Reviewer 2	Clarity and Usability	Structure and organization are excellent. Length may prevent most clinicians from using clinically; however, the executive summary is more than sufficient. It provides clear data; however, the data does not exist to fully inform clinical decision making. The full report is well structured to guide future research.	Thank you

Commentator	Section	Comment	Response
TEP Reviewer 2	Clarity and Usability	Efficacy is not defined, yet policy makers use “efficacy” as a key driver in their formulas on what is covered for reimbursement. I suggest a definition of efficacy that includes minimally clinically significant change (Dworkin).	Thank you for your feedback. We do not think that such a definition is necessary for this review. We used MCIDs that were specific to fibromyalgia, rather than chronic pain in general.
TEP Reviewer 3	Clarity and Usability	The paper is well written and organized. The key points are clearly delineated. It is also clear that the available data on this topic is quite limited thereby restricting its impact on the clinical and policy domains. However, the topic is of great interest. The paper can serve as a starting point to facilitate further research to delineate clinically meaningful subgroups that will lead to the development of tailored treatment protocols specifically addressing subgroups’ needs.	Thank you for you compliment.
TEP Reviewer 4	Clarity and Usability	The Executive Summary is nearly as long as the report. I recommend a much more concise ES, especially given the weak nature of the study and the limited value of the conclusions. The authors should be more clear that the state-of-the literature precludes its use in informing policy and practice recommendations.	Thank you for your suggestions. We cut the length of the ES and added the following clarification to the Conclusion (Report pg. 42) and Abstract: *The limitations of the primary literature preclude any change of policy or practice based on these findings.*
TEP Reviewer 5	Clarity and Usability	Yes to all of the above. I believe that this report will stimulate increased interest in these subgroups, which is needed. I am not sure that the lack of evidence shown in this review will inform policy or practice decisions other than stimulating more study.	Thank you for your feedback
TEP Reviewer 6	Clarity and Usability	The report is well-structured and organized; however, the limitations of the primary literature preclude any change of policy or practice based on the findings.	Thank you- Added to Conclusion (pg. 42, ES-18), Abstract: *The limitations of the primary literature preclude any change of policy or practice based on the findings.*
Peer Reviewer 4	Clarity and Usability	The report is a complete recounting of findings, but might be improved by some of the clarifications suggested above.	Thank you – see specific responses above

Commentator	Section	Comment	Response
TEP Reviewer 7	Clarity and Usability	The main points are pretty clearly presented, though I think it would be useful to more clearly describe variability in evidence for different interventions and subgroups (right now it's all lumped together as SOE: low but it seems to me it really ranges from insufficient to low).	We edited the initial paragraphs of the Discussion to better report the evidence base for a range of subgroups-treatments-outcomes. We also edited the Key Findings in the Results. We added a footnote to the SoE table (Table B ES, Table 12 Report) that single high risk of bias for any treatment-=subgroup-outcome combination were determined to be insufficient evidence and are not included in the SoE tables. Table 12 has been updated.
Peer Reviewer 1	General	The two key questions are potentially interesting; but the initial review step revealed that very few studies had been designed to address the questions or satisfied the selection criteria. At that point, it would have been sufficient to note that studies had not been powered or designed to address the various questions and that further research was needed.	Thank you for your opinion – we appreciate your feedback.
Peer Reviewer 1	General	A statement that further research is needed is really all that can be reliably concluded from this paper. Instead, the authors attempt at reviewing a sparse literature has lead to a bizarre sub-sampling of the existing FM clinical trials literature. The conclusion that treatment need not be different based upon the various subgroupings is not convincing given the very small number of studies upon which they base their conclusions. In some cases, (e.g., psychological treatments) there were more potential subgroupings than studies offering data.	Thank you for your feedback.
Peer Reviewer 1	General	At best, this paper reveals that the authors took a great deal of time and effort to conclude very little. At worse, policy makers will assume that more intensive interventions for more challenging special populations with FM don't need more intensive care (a conclusion no clinician who sees individuals with FM will believe). Lack of data is just that, a lack of data, it does not warrant making any conclusions about any of the key questions.	Thank you for your comment. This review highlights the paucity of data available for clinicians on fibromyalgia treatment effects in subgroups, and identifies specific research gaps that future research on fibromyalgia will hopefully address.

Commentator	Section	Comment	Response
Peer Reviewer 2	General	Given the paucity of data about the topic and the nature of the therapeutic literature of fibromyalgia, I thought the report was exceptionally clear in its methods and message. The clinical meaning of the report is somewhat difficult to assess - mostly because the data they reviewed is of low quality and at high risk for bias. The key questions are appropriate (and important) and well stated.	Thank you for your comments. We edited the Discussion (Report pg. 38) to provide clinical context for the state of the literature on fibromyalgia subgroups: *Clinicians and patients are thus left with little to guide their treatment decisions……*
TEP Reviewer 2	General	The report represents an extraordinary amount of high quality work. Unfortunately, the studies necessary to provide good answers to the primary questions (benefits and harms of various fibromyalgia treatment in specific subgroups of patients) have not been performed. Yet the authors have done an exceptional job of reviewing the studies that are available, and were able to provide conclusions, albeit tentative due to paucity of data.	Thank you for your compliments and comments.
TEP Reviewer 2	General	yes, if Dworkin criteria for minimally clinically effective change are added	We used MIDs that were specific to fibromyalgia in this review. The IMMPACT consensus recommendations are for chronic pain.
TEP Reviewer 3	General	This is a systematic and comprehensive review of the literature on the topic of how subgrouping of adult patients with fibromyalgia syndrome (FMS) impacts treatment response. FMS is a prevalent, chronic, and often debilitating pain disorder. Despite extensive research in the past 4 decades, no single modality has been found to be universally effective to ameliorate FMS related symptoms and disability. Many scientists and clinicians have raised a question of the patient heterogeneity within the FMS population that identification of FMS subgroups may enhance our understanding of the condition and help develop specific ally tailored and efficacious treatments based upon the needs of the subgroups. The present paper addresses whether various treatments would differentially treat FMS subgroups. The topic is significant and quite timely. The review approach is systematic and well documented. There are however, some areas that would benefit from further elaboration and clarification. Most of my concerns are minor and overall, the paper presents significant merit.	Thank you for your summary of the main issues facing the field for the treatment of adults with fibromyalgia. We believe that we have included this information in the Background and Discussion of the report.

Commentator	Section	Comment	Response
Peer Reviewer 3	General	I don't have much to say about this report. Given the nature of such reports, the authors went about getting and examining the data correctly, and writing the correct type of report. I have no disagreement with what they did, nor with their conclusions.	Thank you for your comments.
Peer Reviewer 3	General	I wonder, however, why AHRQ would ever have thought that a study like this would have yielded useful results. Given the limited efficacy of treatments overall, it is hard to see why sub-group treatments would have been effective. Where did AHRQ get the idea that such a study, doubtless an expensive study, would be needed?	Thank you for your comments. AHRQ received a nomination to review Treatments for Fibromyalgia. During the process of topic refinement, it became apparent that such a review of the topic was already underway. Discussions with the key informants and technical experts suggested that there would be a merit to a review focusing on subgroups of patients with common comorbidities that often present challenges for clinical decision making. Despite paucity of information, a review can still provide service to the field by highlighting research gaps that can help advance the field and help clinicians and patients.
Peer Reviewer 3	General	I am also puzzled about those "experts" who recommended a study of subgroups. With similar logic, one could study the depressed and the elderly who suffer with acne. There are always subgroups. While observational studies cannot tell one much about efficacy, they can tell us about whether these subgroups have worse outcomes in the community. Maybe they don't require special treatment. Maybe men with fibromyalgia are the group at risk.	Thank you for your excellent points.
Peer Reviewer 3	General	In addition, there is confusion about the nature of fibromyalgia and the direction of causality. Perhaps it is mental issues, obesity, smoking, education and lifestyle issues that contribute to fibromyalgia symptoms.	Thank you for your comments.
Peer Reviewer 3	General	The authors point out the limitations of short term studies. For chronic problems like fibromyalgia, short-term studies tell us very little about how to treat fibromyalgia patients, whether they are in subgroups or not. The authors understand these issues and discuss them fully in their report. Why didn't AHRQ know this to start with?	Thank you for your feedback. We added a statement to the Discussion, pg. 42: *For clinicians, short-term studies provide very little information about how best to treat adults with fibromyalgia over time.*

Commentator	Section	Comment	Response
Peer Reviewer 3	General	Overall, I found the report to be thorough and thoughtful. We all knew that the limitations were caused by inadequate data. Such information has been in the literature for years.	Thank you for your comments.
TEP Reviewer 4	General	This report will have limited clinical utility due largely to the limitations of the empirical literature on the topic of this review. This limitation may be a function of the narrow specification of the criteria for selecting studies to be included in the review and/or the narrow specification of the questions. The clinical value is limited to a conclusion that there is only low strength of evidence to either reject or confirm common clinical assumptions or beliefs about the moderating role of important individual difference variables on patient responses to common fibromyalgia treatments. Given all of the limitations of the evidence and the qualitative nature of the methods employed, it may be more appropriate to conclude that the questions posed for this review could not be adequately answered, rather than concluding that there is low strength of evidence that there is efficacy, comparative effectiveness or differential risk of harms from treatments for fibromyalgia for the subgroups examined.	Thank you for your helpful feedback. We edited the first paragraphs of the Discussion (Report pg. 38 and ES**)** to better specify the limitations of the literature base for FM subgroups, and how this lack of information likely impedes clinical decision-making. In doing so, we included some of your suggested text at the start of the Discussion (pg. 38**)**: *There is low-strength or insufficient evidence to reject or confirm common clinical assumptions or beliefs about the moderating role of important individual difference variables on patient responses to common fibromyalgia treatments.* We edited the first segments of the Discussion (Report, ES) to clarify areas where there is low vs. insufficient SoE. The Conclusions (pg. 44 Report, ES, Abstract) have been edited: *The fibromyalgia evidence is largely insufficient to determine subgroup effects for interventions basically other than duloxetine. The limitations of the primary literature preclude any change of policy or practice based on these findings.*
TEP Reviewer 5	General	This report is excellent and reveals significant gaps in research guiding treatment of Fibromyalgia in adult subgroups. Hopefully this report with stimulate greater interest and inclusion of adult subgroups in FM research. The subgroups that are the target population for this review are clearly defined.	Thank you for your compliments and comments.
TEP Reviewer 6	General	A carefully done and meticulously reported review exploring the evidence for subgroup effects among adult patients diagnosed with fibromyalgia. Findings are limited by sparse data that precluded meta-analyses, and by the low quality of much of the literature, including limited credibility of reported subgroup analyses. The conclusions of the report are that the limited evidence has not established any important subgroups among fibromyalgia patients.	Thank you for your feedback.

Commentator	Section	Comment	Response
Peer Reviewer 4	General	The report addresses the comparative effectiveness of treatments for fibromyalgia in subgroups of patient populations with a higher prevalence of fibromyalgia, higher severity or duration of symptoms of fibromyalgia, or particular comorbidities. The rationale for investigating treatments in these subgroups, and the choice of specific subgroups, both appear reasonable. However, the major findings are that there is little evidence. I am not sure how much this review adds to the McMaster work other than by establishing that there is little information within RCTs upon subgroups and there is also little observational literature available on the topic. The findings of the review are somewhat predictable and the major benefit of doing this review is the opportunity to ask 'what next?'	Thank you for your comments. We edited the Conclusions (pg. 42 and ES) as well as the Key Findings of the Discussion (Report, ES) to better delineate where there is low vs. insufficient evidence. Conclusion: *The fibromyalgia evidence is largely insufficient to determine subgroup effects for interventions basically other than duloxetine. The limitations of the primary literature preclude any change of policy or practice based on these findings.*
TEP Reviewer 7	General	This CER focuses on how effects of treatments for FM might vary in different subgroups. While the methods and conclusions generally appear to be sound I think there may be some areas that could improve the report, particularly with regard to assessment of methodological issues related to subgroup effects and presentation of findings. I also think there are some potential issues with including pooled IPD analyses of selective studies that might be highlighted better.	We edited the Methods text of the report and ES per your specific comments that are addressed individually above.
Public Reviewer: American Physical Therapy Association	General	The authors have been very thorough in their preparation for review of the literature for treatments for fibromyalgia. It takes a tremendous amount of work to prepare and complete such an extensive review.	Thank you
Public Reviewer: Physical Therapy Association	General and Abstract	Throughout the text, we would have preferred to see actual data rather than generalized statements. For example in the abstract "over half were drug trials" it would be clearer to provide a percentage to represent the number of trials and to indicate if these were all randomized controlled trails (RCT's) or all studies combined.	General: Thank you for your feedback. In presenting information, it is our preference to make summary statements first, followed by detail in text, when indicated, or with reference to a table that contains the supporting detail for our comment. Abstract: After updating our literature review, we added the specific proportion of included studies that were drug trials.

Commentator	Section	Comment	Response
Public comment: Pfizer Inc. Representative	General	Having reviewed the your draft comparative effectiveness document, we wish to note that the following information we provided in our write up to you is not acknowledged, included or referred to: In our submitted document, on page 1, under the section entitled, "Treatment for Fibromyalgia in Adult Subgroups and Individuals with coexisting mental health conditions", we provided a section pertaining to our study (Study 1275) by Arnold et al, on Fibromyalgia Patients with Comorbid Depression. Our write up provided key aspects of this study including patient inclusion criteria, study design, patient demographics, treatment emergent adverse events and study results which concluded that pregabalin significantly improved FM pain in patients with FM and comorbid depression receiving a concurrent antidepressant medication compared with placebo. In addition, we also stated that the safety profile of pregabalin in this population was consistent with previous studies and current product labelling. (Arnold LM, Sarzi-Puttini P, Arsenault P, et al. Efficacy and Safety of Pregabalin in Patients with Fibromyalgia and Co-Morbid Depression Receiving Concurrent Antidepressant Therapy: A Randomized, 2-Way Crossover, Double-Blind, Placebo-Controlled Study [abstract]. Presented at American College of Rheumatology; 2013 Oct 25-30; San Diego, CA.)	Thank you for the additional information. Our systematic review included published studies, and the referenced information based on a professional meeting abstract from October 2013. Although we were able to locate the presentation slides, we did not find a published article for this information.

Commentator	Section	Comment	Response
Public comment: Pfizer Inc. Representative	General	In our submitted document, page2, under the same section, “Individuals with coexisting mental health conditions”, we provided a write up from Arnold et al on a pooled analysis using data from 3 pregabalin studies to investigate the relationship between the effect of pregabalin on pain in relation to baseline anxiety and depression in 2,013 patients with fibromyalgia. We provided information on mean baseline pain scores, HADS-A and HADS-D across treatment group. Pregabalin 300, 450, 600 mg/day showed significant improvement in pain compared with placebo ($p<0.0001$). Pregabalin 450 mg/day was associated with improvement in both HADS-A and HADS-D compared with placebo ($p<0.05$ for both), while the 600 mg/day group was associated with improvement in HADS-A compared with placebo ($p=0.009$). An examination of linear models of the association between baseline HADS–A and HADS–D for all treatment groups pooled indicated the improvement in pain was not related to baseline levels of anxiety or depression and correlation coefficients between changes in pain and changes in HADS-A and HADS-D indicated low-to-moderate association between changes in pain and mood. Finally we provided information on path-analysis which showed that most of the pain relief observed with pregabalin treatment was from a direct analgesic effect and was not explained by improvement in mood. Similar results were seen in another study. (Arnold LM, Crofford LJ, Martin SA, et al. The effect of anxiety and depression on improvements in pain in a randomized, controlled trial of pregabalin for treatment of fibromyalgia. Pain Med. 2007;8(8):633-638.).	The article by Arnold et al. 2010 in *Psychosomatics* is included in this report-see Appendix Table E6 for pooled studies (*Arnold LM, Leon T, Whalen E, et al. Relationships among pain and depressive and anxiety symptoms in clinical trials of pregabalin in fibromyalgia. Psychosomatics. 2010 Nov-Dec;51(6):489-97. PMID 21051680.)* The second study mentioned by the Reviewer (Arnold et al., Pain Med 2007) was excluded due to the short duration of follow-up. It is a secondary analysis of data from one RCT with 8-week outcomes (Crofford et al. 2005, reference #33, page E-47 of our report).

Commentator	Section	Comment	Response
Public comment: Pfizer Inc. Representative	General	In our submitted document, page 3, under "Older Adults", we refer to Emir et al and a pooled, post hoc analysis of 4 randomized, placebo-controlled, multicenter, 8- to 14-week clinical studies in 2,061 patients with FM to evaluate the effect of patients' characteristics at baseline on the magnitude of pain response to pregabalin 300 mg/day and 450 mg/day. Significant treatment by baseline interactions were observed for baseline mean pain (p=0.037), treatment by baseline sleep score (p=0.071), and treatment by age (p=0.051) and a trend for a treatment by duration of FM interaction was observed but did not meet statistical significance. We noted that the results of these analyses suggest that pain improvement may not depend on baseline levels of depressive and anxiety symptoms or duration of FM but that older age, more impaired sleep, and higher levels of pain at baseline were associated with greater improvements in pain at endpoint. (Emir B, Murphy TK, Petersel DL, et al. Treatment response to pregabalin in fibromyalgia pain: effect of patient baseline characteristics. Expert Opin Pharmacother 2010; 11(14):2275-80.).	We excluded articles that focused on predictors of treatment response. This is a *post hoc* analysis of predictors of treatment response to pregabalin (article p. 2278) in patients who were required to stop all other FM-impacting drugs prior to the RCTs.
TEP Reviewer 2	Appendix	In the table of off label drugs for FM, it is important to note that NIH funded a study of Neurontin in fibromyalgia and it had similar effect sizes as the larger multisite pregabalin trials, though the study tested Neurontin vs placebo not Neurontin vs pregabalin. Arthritis Rheum. 2007 Apr;56(4):1336-44. Gabapentin in the treatment of fibromyalgia: a randomized, doubleblind, placebo-controlled, multicenter trial. Arnold LM1, Goldenberg DL, Stanford SB, Lalonde JK, Sandhu HS, Keck PE Jr, Welge JA, Bishop F, Stanford KE, Hess EV, Hudson JI. Supported by NIH grant N01-AR-2-2264 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (Dr. Arnold, Principal Investigator).	Thank you for the information. Appendix table A is a list of medications used off-label for FM; it does not provide study details.

Key:
ES: Executive Summary
FM: Fibromyalgia
pg.: page
RCT: randomized clinical trial
SoE: strength of evidence