Comparative Effectiveness Research Review Disposition of Comments Report

Research Review Title: *Pharmacotherapy for Adults With Alcohol-Use Disorders in Outpatient Settings*

Draft review available for public comment from October 24, 2013 to November 20, 2013.

Research Review Citation: Jonas DE, Amick HR, Feltner C, Bobashev G, Thomas K, Wines R, Kim MM, Shanahan E, Gass CE, Rowe CJ, Garbutt JC. Pharmacotherapy for Adults With Alcohol-Use Disorders in Outpatient Settings. Comparative Effectiveness Review No. 134. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center under Contract No. 290-2012-00008-I.) AHRQ Publication No. 14-EHC029-EF. Rockville, MD: Agency for Healthcare Research and Quality; May 2014. 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 comparative effectiveness 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 comparative effectiveness 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: www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1908
Published Online: May 13, 2014

Comment #	Section	Comment	Response
TEP Reviewer #1	Quality of the Report	Superior	Thank you.
Peer Reviewer #1	Quality of the Report	Good	Thank you.
TEP Reviewer #2	Quality of the Report	Superior	Thank you.
Peer Reviewer #2	Quality of the Report	Superior	Thank you.
TEP Reviewer #3	Quality of the Report	Superior	Thank you.
Peer Reviewer #3	Quality of the Report	Superior	Thank you.
Peer Reviewer #4	Quality of the Report	Superior	Thank you.
Peer Reviewer #5	Quality of the Report	Superior	Thank you.
Peer Reviewer #6	Quality of the Report	Poor	From the specific comments, the reviewer generally did not like the scope of the report and would like to have seen some different questions addressed.
TEP Reviewer #1	General/ overall	This report is highly clinically meaningful. Its length, however, may discourage busy clinicians to review it and it will therefore be important to present the ES as both a separate and integrated document.	Thank you, we agree. We hope that a journal article publication will serve as a useful source for busy clinicians too.
TEP Reviewer #1	General/ overall	The target population is well defined. The audience is defined in the preface but it could be more explicit if also included in the ES and main body of the review.	Thank you. To minimize redundancy, we prefer not to repeat this information in multiple places.
TEP Reviewer #1	General/ overall	The key questions are appropriate and explicitly stated.	Thank you.

Comment #	Section	Comment	Response
TEP Reviewer #1	General/ overall	The term "addiction" is used on few occasions; "substance use disorder" or "use disorder" might be considered instead.	We have removed some uses of the term “addiction” as suggested. We now only retain its use when describing guidelines (and we are using their words), or describing a study (and using the authors description of something), and in parentheses when providing definitions for alcohol dependence in the Intro (to indicated that it is generally considered a synonym).
TEP Reviewer #1	General/ overall	The abstract does not mention disulfuram; since it is one of the FDA-approved medications, it seems important to include it even if findings are NS. Differentiating po vs. IM naltrexone might also be considered.	For disulfiram, we have added the following to the abstract: “Evidence from well-controlled trials does not support efficacy of disulfiram, except possibly for patients with excellent adherence.” We have also differentiated the findings for oral and IM naltrexone as suggested by this reviewer and others.
Peer Reviewer #1	General	Well conducted, not quite up to date. New data on nalmefene are to be included, according to the authors' statement.	Thank you, we have updated our literature searches and added new studies on nalmefene meeting our inclusion criteria.
TEP Reviewer #2	General/ overall	In general I think the reader gets lost in the trees and perhaps cannot see the forest. It almost appears that some of the lesser-studied compounds are given as much space as the main ones. Since many of these don’t have many trials describing specific trials is weighted almost as heavily in writing space as larger meta-analytic evaluations. This detracts from the overall impact in my opinion. It would much rather see statement like there is not enough evidence to comment on a particular outcome or medication then try to summarize what is know (even if it might be spurious).	In the full technical report, we only include evaluations for drugs with multiple studies. We have placed the assessment of drugs with just 1 available study to an Appendix. We have followed the suggestion of this reviewer in the executive summary.
TEP Reviewer #2	General/ overall	NNT contain bias since some placebo groups received CBT and CBI etc. This is a big problem/concern since primary care health care providers and others will get a biased estimate of NNT which is likely too high if one used trials where CBT and other effective therapies are also used to inflate placebo response relative to active medication. These NNT should be calculated separately for trials with and without ancillary therapy.	We disagree. Studies typically included psychosocial co-interventions; effect sizes reflect the added benefits of medications (i.e., the benefits beyond those of psychosocial interventions, rather than the benefits of medications when used alone). We have added this point to the abstract, to try to make this more clear very early in the report. We also added more explicit statements about this issue in the Results of the Executive Summary and full report.

Comment #	Section	Comment	Response
TEP Reviewer #2	General/ overall	In a similar vein, combining studies where medications are evaluated in dual-diagnosis populations with those where these were excluded reduced the validity of the data and does not do the review justice. This really detracts from the overall interpretation of the results and could mislead particularly uninformed readers like policy makers and primary care health providers.	This is an empiric question that we can explore with data. It is not clear whether findings would differ if combining studies of people with dual diagnosis and those without. We think that it might be somewhat artificial to separate studies that describe a dual diagnosis population from those that do not when we know that many people with alcohol dependence have additional psychiatric diagnoses (even if studies didn't report it), and that many studies don't report information on co-occurring diagnoses. To explore this issue, we have added sensitivity analyses to the report that stratify studies that reported dual diagnoses and those that did not. We conducted these for acamprosate and naltrexone (the 2 medications with enough studies to conduct a stratified analysis). Ultimately, the analyses are not very revealing and do not significantly impact our findings because there are so few studies that specify enrolling a population with dual diagnoses (just 1 trial for acamprosate and 6 for naltrexone that contributed data to any of our meta-analyses). We also added a paragraph to the Discussion section of the full report under Limitations of the Comparative Effectiveness Review Process to describe this: "We combined studies that described including populations with a dual diagnosis (e.g., alcohol dependence and depression) and those that did not in our meta-analyses. To determine whether this potential population heterogeneity would have a significant impact on our conclusions, we conducted sensitivity analyses for acamprosate and naltrexone (the medications with enough studies to conduct a stratified analysis by presence of dual diagnosis). Ultimately, the analyses were not very revealing and do not significantly impact our findings because there are so few studies that specify enrolling a population with dual diagnoses (just 1 trial for acamprosate and 6 for naltrexone that contributed data to any of our meta-analyses, Appendix F).

Comment #	Section	Comment	Response
TEP Reviewer #2 (continued)	General/ overall	In a similar vein, combining studies where medications are evaluated in dual-diagnosis populations with those where these were excluded reduced the validity of the data and does not do the review justice. This really detracts from the overall interpretation of the results and could mislead particularly uninformed readers like policy makers and primary care health providers.	Effect sizes did not change significantly and were sometimes identical. The one possible exception, that might be considered a significant change in the effect size, was for heavy drinking days and naltrexone. When including all studies, subjects treated with naltrexone had 3.8 percent fewer heavy drinking days than those treated with placebo (WMD, -3.8; 95% CI, -5.8 to -1.8; 11 trials); when excluding 3 studies that enrolled patients with dual diagnoses, the effect size was slightly larger (WMD -4.9; 95% CI, -7.1 to -2.7; 8 trials). It might be somewhat artificial to separate studies that describe a dual diagnosis population from those that do not because many people with alcohol dependence have additional psychiatric diagnoses and many studies don't report information on co-occurring diagnoses."
TEP Reviewer #2	General	6 month treatment time – does not take into consideration dropouts and missing data. This should not be in the scope of this review. This review should be reviewing evidence not making recommendations on trial design.	We did not make any recommendations about trial design. We were only describing that studies of short duration may yield misleading conclusions about treatment efficacy, due to fluctuations in drinking behavior that are typical of the course of alcoholism. We have revised the wording of the paragraph in the discussion to clarify our meaning based on this comment and comments from another peer reviewer. And we removed the mention of 6 month treatment time.
TEP Reviewer #2	General	Should emphasize the need for Federal Funding of large prospective studies to evaluate changes in harm and health consequences with various levels of drinking prospectively.	We have added this general idea (but without Mention of any particular funder) to the future research table (KQ 2 row).
TEP Reviewer #2	General	The report doesn't mention or highlight the cost-benefit analysis done for the COMBINE Study (Zarkin et. al.). Should be singled out as a model.	We have added a brief description of this analysis to the discussion in the ES and full report "…a cost study based on the COMBINE trial reported that several treatment combinations that include pharmacotherapy lead to reduced median social costs associated with health care, arrests, and motor vehicle accidents compared with medical management plus placebo."

Comment #	Section	Comment	Response
TEP Reviewer #2	General	Some concern with comparing studies done in diverse countries and cultures with differing pre-treatment experiences and non-pharmacological adjunctive therapy. This is almost like comparing apples and oranges. For instance, if one has congestive heart failure but some patients also are receiving treatment for diabetes while others have thyroid problems, lumping them together does not provide enough clarity. Also, suppose some CHF patients are getting a diuretic (equivalent to CBT) and others are not and all are given digitalis – there is a bias comparing these trials and should be evaluated separately. The co-interventions in the tables of studies are all over the map. This needs to be mentioned as a strong limitation of this work. In some ways this is more like effectiveness evaluation rather than efficacy evaluation. Personally, I would rather see the evaluation be done in trials with and without these additional interventions. A section should be written about this as a caveat and limitation of this sort of aggregate work. Also, perhaps a focus on one country or area might be more appropriate	It is an empiric question (whether the medications have varying efficacy or effectiveness for different cultures or in different countries). We have explored this potential heterogeneity in the report, with analyses stratified by country. Regarding the co-interventions, we added text throughout the report (abstract, results, and discussion) to clarify that effect sizes reflect the added benefits of medications. Also, we have added text to the limitations section of the Discussion explaining the variety of different psychosocial co-interventions used across trials, and that this heterogeneity limits our certainty about the effect of medications when used alone (with no psychosocial co-intervention) or when added to a particular psychosocial intervention. Reporting of previous and ongoing psychosocial interventions was variable across the included studies.
TEP Reviewer #2	General	It is not clear how risk of bias (or scientific weight) was taken into account in the meta-analysis and NNT estimates.	We explain this in the Methods section under Data Synthesis. “We did not include studies rated as high or unclear risk of bias in our main analyses, but did include them in sensitivity analyses.”
TEP Reviewer #2	General	It is not clear why in the nalmefene and some other compound reviews that the positive trials are highlighted in the writing section. Why are these results given more weight?	We disagree. We describe all included trials and did not exclude trials with negative results.
TEP Reviewer #2	General	Also, if exclusion criteria state that people with various psychiatric disorders are excluded why is the assumption that when co-morbidity is not reported it cannot be assumed that there was no significant co-morbidity. This is difficult to understand.	We excluded studies of individuals under the age of 18 years old. We did not have any other exclusion criteria related to study populations. So individuals with co-morbid conditions, including psychiatric conditions, were included in this systematic review. In addition, our fifth key question focused specifically on the differential effect of interventions on various subpopulations, including those with psychiatric disorders. We are not certain if this comment is referring to those issues or to some aspect related to our assessments of individual studies; if the latter, we did not make that assumption that studies excluding people with various psychiatric disorders.

Comment #	Section	Comment	Response
TEP Reviewer #2	General	There has been another quetiapine study reported by the NIAAA clinical trials group that is not included in the review. If nalmefene studies are added so should this one. The same should be done with the varenicline study also reported recently.	Thank you. We have updated our literature search and added any studies meeting our inclusion criteria. We did find a new quetiapine study that is likely the one mentioned here. It did not meet our inclusion criteria because it was <12 weeks. We did find 1 new study of varenicline that meets our inclusion criteria and has been added to the report.
TEP Reviewer #2	General	For sertraline, under Drinking Days. It is not clear why South Carolina is highlighted rather than being in the United States. This seems odd. Again not clear why only positive studies are highlighted for certain variables.	We deleted South Carolina, and put in U.S. to be consistent with the other entries. As mentioned in the responses to other comments, we did not only highlight positive studies. We describe the results of both positive and negative studies that reported the outcomes. We focus on the results of our meta-analyses (when we were able to do them), rather than results of individual studies.
TEP Reviewer #2	General	The PREDICT Study from Germany was not included in Naltrexone vs. Acamprosate comparisons.	Thank you. We have added it. Indexing of the Mann study was not complete at the time of our literature searches (thus our literature searches did not find it). The Mann study (Mann K, Lemenager T, Hoffmann S, et al. Addict Biol. 2012 Dec 12) is eligible and we have added it to the report.
TEP Reviewer #2	General	For the naltrexone OPRM1 studies, it should be made clear that the A118G SNP is infrequent in African Americans so several of the studies (Oslin et. al. and Anton et. al.) restricted their analyses to Caucasians. Also, since genetic epistasis might occur in different racial groups the race of the studies should be reported and this issue highlighted. Secondly, if the main trial results show no main effect of naltrexone it could be a failed clinical trial and not indicative of lack of pharmacogenetic effect. This should also be highlighted.	We have added information to address this first point to the results section of KQ 6. We added text to note the studies that restricted analyses to Caucasians and we report the racial information for each study in Table 35. We added the following in the results section of KQ 6: “The larger trial (N=627) that the study sample (N=220) was drawn from did not show a positive effect of naltrexone. The authors explained that the lack of an overall treatment effect of naltrexone in the larger trial might suggest that this was a sub-optimal sample in which to evaluate pharmacogenomic predictors of treatment.”

Comment #	Section	Comment	Response
TEP Reviewer #2	General	I am not sure I agree with not considering the Oslin et. al study in the main analysis of the OPRM1 gene. While it was retrospective analysis of several combined data sets it is really no worse then some of the other explorations published.	The study design for the Oslin et al study did not meet our inclusion criteria. It pooled data for a subset of subjects from 3 separate trials (and 1 of those was less than 12 weeks in treatment duration). However, we conducted sensitivity analyses that include this study (as explained in the report).
TEP Reviewer #2	General	I would encourage the authors to expand on the discussion of epidemiological findings and health outcomes and try to put some data into who drinking reduction during trials might translate into health benefits as gleaned from epidemiological studies.	We have expanded the discussion about health outcomes as suggested. For example, we added the following: "…A recent model estimated that increasing treatment coverage to 40% of all people with alcohol dependence in the European Union would reduce alcohol-attributable mortality by 13% for men and 9% for women. Further, a cost study based on the COMBINE trial reported that several treatment combinations that include pharmacotherapy lead to reduced median social costs associated with health care, arrests, and motor vehicle accidents compared with medical management plus placebo."
TEP Reviewer #2	General	Under the Policy Decision Section please consider adding that ancillary therapies such as CBT, 12 Step and CBI, have been shown effective in other studies and in some pharmacotherapy studies in their own right and therefor medications need to be judged but in an independent fashion from these ancillary therapies and together with them. This is crucial for decision makers.	We have added the following to that section: "Although we did not evaluate the effectiveness or comparative effectiveness of psychosocial interventions for alcohol use disorders (e.g., cognitive behavioral therapy, 12-step programs, combined behavioral intervention), such interventions have been evaluated within some of the included pharmacotherapy studies and in studies that were not included in our review. It may be important for decisionmakers to have information about the efficacy of psychosocial interventions from other sources. Further, decisionmakers may want information about the efficacy of medications when used independently of psychosocial interventions (which is limited) and when used together with them."
Peer Reviewer #2	General	This is an outstanding systematic review of medications for alcohol dependence. Thank you! The authors should be commended on the thoroughness of the review. The suggestions here are intended to make it even more useful to readers.	Thank you!

Comment #	Section	Comment	Response
TEP Reviewer #3	General	The authors did an excellent job. The results of this report are only as good as the design and population recruited in the designated studies, and of course, how well the studies were carried out (which is difficult to measure). For example, for disulfiram, we know this is a magic bullet to stop drinking if the patient takes the medication at the proper dose. The problem has been lack of compliance. We have not funded many disulfiram studies because of this as well as the difficulty of conducting an unbias trial. For instance, if a patient is taking disulfiram and drinks, the person will know if he/she is taking the medication. The compound also works as a psychological deterrant. However, if one doesn't drink doing the trial, both the medication and placebo groups will experience the pyschological deterrant effect, making it difficult to show an effect. Finally, in the US, disulfiram is being used as much as acamprosate and not that far behind for naltexone. Thus, it must be working for some individuals. Probably a sentence should be stated about disulfiram in the sturctured abstract.	Thank you! We appreciate all of these comments. We have added a sentence about disulfiram to the structured abstract as suggested.
TEP Reviewer #3	General	Finally, I believe that alcohol medications work for some but not everyone, like SSRIs for depression. Even though there is a small effect size for naltrexone and acamprosate, this does not necessarily mean that an individual gets a small effect. It appears that the medication has a major effect in some individuals but many may not experience any effect. This is not that unusual for complex diseases. For example, a recent meta analysis of antidepressions that includes positive and negative trials reported a small effect size of 0.3. Because of this, personalized medicine will be a major focus of alcohol research over the next decade.	We agree.
TEP Reviewer #3	General	Lastly, NIAAA started a new program of conducting multisite studies using a network of sites. Three studies have been completed and published (see attachment below). Two studies were published in 2012 and should have been included. They include quetiapine and levetiracetam. You might want to include levetracetam in your document although the results will be negative. Our recent study was with varenicline which we found a positive result.	We reviewed all of the suggested studies against our inclusion/exclusion criteria. Our update literature searches identified 2 of these studies, but did not find the varenicline study because indexing was not yet complete. The varenicline study meets our inclusion criteria, and has been added to the report. The newer quetiapine study was <12 weeks in duration. Levetiracetam was not an included medication because it is not commonly used for alcohol use disorders and searches during our topic refinement period (prior to conducting the review) suggested that there would be too few studies to synthesize. Also, discussions with key informant did not identify it as one of the clinically important medications to include.

Comment #	Section	Comment	Response
Peer Reviewer #3	General	This is an excellent review, that has a lot of extremely helpful clinical information. The key questions are appropriate and well defined. The population (studied) is well stated.	Thank you
Peer Reviewer #3	General	The only question is whether the target population is clinicians, researchers or others....	The audience is defined in the preface. The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well informed decisions and thereby improve the quality of health care services.
Peer Reviewer #4	General	the review is well done, well written. The criteria for recommendations requires a very high level of evidence. it would ahve been interesting to consider the rating of evidence in relation to medications for other psychiatric disorders such as depression or schizophrenia.	Thank you Yes, that might be interesting, but it was beyond the scope of this project.
Peer Reviewer #4	General	A casual reader may come away by thinking there is no reason to use pharmacotherapy.	The final report provides the evidence on the efficacy, comparative effectiveness, and harms of pharmacotherapy for alcohol use disorders and indicates that several medications have good evidence supporting their efficacy.
Peer Reviewer #5	General	The intention of this report is to provide important information for patients, clinicians, health system leaders, policy makers and other professionals regarding pharmacotherapy for adults with alcohol use disorders in outpatients setting. The report is timely, detailed, and valuable because it aggregates data from original sources and meta analysis to provide information regarding the use of medications in AUD. The report has clear clinical implications because it provides a detailed summary of the efficacy of FDA approved medications, and provides data on medications that show some promise in treatment of individuals with AUD.	Thank you

Comment #	Section	Comment	Response
Peer Reviewer #6	General	The stated purpose of this report is to inform patients, clinicians, health care system leaders, and policy makers to help make informed decisions to improve the quality of health care services. Also, the report should help inform clinical guidelines for reimbursement coverage policies. Given the lack of informed knowledge on the use of medications to treat alcohol use disorders there is an important need for a report that addresses these issues. Most of use with intimate knowledge of the research data has concluded that current clinical practice woefully undertreats patients with medications. As this reports states that only 10% of patients in treatment receive medications as a part of their recovery. Yet for medications to be successfully incorporated into treatment there needs to be an informed population that informed about the success of medical treatments, clinicians who are familiar with the use of medications, and a mechanism to adequately pay for treatment. A golden opportunity is wasted as this report fails to significantly add to our current knowledge base to inform prospective patients, clinicians, or entities responsible to reimbursement policies.	We agree with the peer reviewer that there are many opportunities to improve the quality of care for those with alcohol use disorders. A first step is to analyze available evidence to inform practice. We agree that this evidence can inform decision by clinicians and policymakers. We hope that the evidence in this systematic review can be disseminated to decisionmakers who can then incorporate these findings into practice and policies. We disagree that this is a wasted opportunity. We have identified areas where evidence can provide insight, and other areas where evidence is scant. In these areas of uncertainty and little evidence, we encourage providers to use clinical judgment, and engage with their patients in decisionmaking about the available treatment options.
Peer Reviewer #6	General	This report does an adequate job of identifying the relevant clinical population and explicating stating appropriate key questions. Where this report fails is in a thoughtful and clinically meaningful review of the relevant data. The methodology used here is a Cochrane type meta-analysis where studies are simply combined and an overall effect is reported. From an academic perspective, the meta-analysis presented here lacks creativity as it simply presents the data in a formulaic manor. This adds nothing to knowledge base and so is unlikely to advance of understanding of how to effectively and safely treat patients.	We disagree with the assertion that the review is not thoughtful or clinically meaningful. As we point out in the Introduction (and as this same reviewer points out in their initial general comment), medications for alcohol use are greatly underutilized, with a small percentage of eligible patients receiving medications. It is important to address the basic questions of whether or not the medications work, and which ones work best, and (if they do) to get that message to a large number of providers who might prescribe the medications. While some of the more nuanced questions suggested by this reviewer in subsequent comments are interesting, many of them are not among the questions deemed most important by our team, our Key Informants, and our Technical Expert Panel. Most of them are beyond the scope of our review.

Comment #	Section	Comment	Response
Peer Reviewer #6	General	In contrast, a recent meta-analysis by Maisel et. al, published in Addiction (2012) conducts a meta-analysis of naltrexone and acamprosate but asks the clinically interesting questions of the role of abstinence before initiating treatment and are there differences in treatment outcomes. Consistent with predictions based on the mechanism of action, the authors conclude acamprosate is more effective than naltrexone in fostering abstinence but naltrexone was more effective in subjects who were initially abstinent before starting naltrexone and more effective than acamprosate in reducing heavy drinking and alcohol craving. This is an example of the type of meta-analysis that leads to clinically meaningful guidance for health care providers and patients.	(this comment is a continuation of the previous comment). We agree that the review by Maisel addressed a clinically interesting question. We did not set out to conduct a detailed comparison of subgroups by whether they were initially abstinent. We provide some information in the Discussion (Applicability section) about the role of abstinence before initiating treatment: "Most studies required patients to abstain for at least a few days prior to initiating medication, and the medications are generally recommended for maintenance of abstinence. Acamprosate and injectable naltrexone are only approved for use in patients who have established abstinence, though the duration of required abstinence is not set. However, some studies enrolling patients who were not yet abstinent have reported reduction in heavy drinking with naltrexone or acamprosate."

Comment #	Section	Comment	Response
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	General	The rationale for the development of a once monthly extended release naltrexone injection was based on the premise that outcomes due to naltrexone therapy could be improved with a product that frees patients from the requirement to take a daily dose, and provides circulating naltrexone concentrations for one month or greater. Oral naltrexone was initially approved by the US FDA in 1994 for the treatment of alcohol dependence but its clinical use over the following years had been limited. It had been documented that this was at least in part due to the difficulty in adherence to a daily oral medication regimen [Pettinati 2000, Volpicelli 1997, Chick 2000 and Monti 2001 references 62, 106, 115, and 122 in the Review]. No direct head-head studies of 50 mg daily oral naltrexone and once-monthly 380 mg extended release injected naltrexone have been performed that would support the assumption that clinical results of treatment are similar and thus can logically be pooled. Consequently, we would like to comment on, and make note of the following: • Efficacy with respect to alcohol consumption outcomes: We would like to highlight the findings from the VIVITROL pivotal study in alcohol dependence. • Quality of life measures. We would like to highlight the findings from the VIVITROL pivotal study. • There are important differences in prescribing information for extended release injectable naltrexone (VIVITROL) and oral naltrexone, in particular their contraindications [www.vivitrol.com/Content/pdf/prescribing_info.pdf]	Thank you. In the new version of Table C, and in the Executive Summary sections on naltrexone, the naltrexone information is reorganized so that the information is separated for injectable naltrexone and various doses of oral naltrexone. We have already included the findings from the pivotal study (Pettinati 2009) in the report. We revised portions of the Discussion to specify the contraindications for oral naltrexone and those for injectable naltrexone separately. We now reference the updated prescribing information for injectable naltrexone.
Peer Reviewer #2	ES	ES-10 last paragraph: The authors conjecture that the impact of nalmefene was “not likely clinically significant”. I believe many alcohol experts would consider 1 drink per drinking day (7 drinks a week) a clinically significant decrease. Moreover, that sort of conjecture does not fit well in this beautiful systematic review.	We have deleted that as suggested.
Peer Reviewer #2	ES	ES-12 second to last paragraph: The paragraph on under-utilization seemed also to lack an evidence base despite research in this area. While some of the conjecture was possibly appropriate, it would be nice to have these parts of the review also referenced appropriately. An important barrier to primary care is that many (most?) general psychiatrists do not treatment AUD and are not comfortable managing AUD with medications; this will likely have to come first and I believe should be mentioned as a barrier. 15-17	Thank you for the suggested references. We have added references to the articles by Harris et al in to our paragraph about underutilization. We agree that this strengthens the paragraph.

Comment #	Section	Comment	Response
Peer Reviewer #2	ES	Focus on genetic factors in the Gaps seem potentially excessive (ES Table E). While it is exciting that we might be able to select medications based on pharmacogenetic studies, to me the discussion of genetically guided treatment should also include mention of the hefty placebo + medication monitoring response in COMBINE. The placebo response (perhaps in part a response of AUD to repeated BIs of MM)18,19 might markedly increase the benefit of medications and MM in real world primary care settings.13,14 For patients who tolerate the medications, I would wonder if we would want to only offer medications to those who might benefit from a genetically proven pathway—given the generally low risk profile of these medications—and the demonstrated benefit of MM and placebo.	Although the gaps related to KQ 6 take up more space in the table, we have not prioritized the gaps in the table. It takes more text to properly explain the gaps related to KQ 6. Also, with the addition of more gaps (from reviewer comments), they now take up a smaller proportion of the total space. We don't understand how the placebo + medical management (MM) response in COMBINE is related to the discussion of genetically-guided treatment or how it is within the scope of our discussion. This is an interesting theory (that repeated behavioral interventions of MM might markedly increase the benefit of medications and MM in real world primary care settings), but it is not supported by evidence currently. We have not added this theory to the report.
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	ES	Table C footnote: change per day to per month	We fixed this.
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	ES	ES-11; KQ 2: It is noted above that the results for naltrexone are insufficient and that the data were not poolable for oral naltrexone and injectable naltrexone. No direct head-head studies of 50 mg daily oral naltrexone and once-monthly 380 mg extended release injected naltrexone have been performed that would support the assumption that clinical results of treatment are similar and thus can logically be pooled. As the authors noted, in the pivotal study of VIVITROL in alcohol dependence, those receiving a 380 mg monthly dose of extended release injectable naltrexone (VIVITROL) had greater improvement on the SF-36 mental health summary score than those receiving placebo at 24 weeks (8.2 versus 6.2, p=0.044) [Pettinati 2009, referenced in the Review].	Thank you. In the new version of Table C, and in the Executive Summary sections on naltrexone, the naltrexone information is reorganized so that the information is separated for injectable naltrexone and various doses of oral naltrexone. We are still unable to pool any data for the quality of life or function outcomes (so that is still appropriate to indicate in the footnote where we have it). As noted by the comment, we have included the information from this study (Pettinati 2009) in the report.

Comment #	Section	Comment	Response
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	ES	ES-12: In the pivotal study of VIVITROL in alcohol dependence, those receiving a 380 mg monthly dose of extended release injectable naltrexone (VIVITROL) had greater improvement on the SF-36 mental health summary score than those receiving placebo at 24 weeks (8.2 versus 6.2, P <0.044) [Pettinati 2009, referenced in the Review].	We have already included this finding in the results. Thank you for the comment. We reviewed it and we affirm our decision to grade this as insufficient due to unknown consistency, imprecision (the between group difference of 2.0 does not reach the threshold typically considered clinically significant, but the confidence interval includes values that would support a variety of conclusions), and because this finding for SF-36 mental summary score was just one finding from one study for a single secondary outcome (and just the mental summary score of that secondary outcome).
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	ES	ES-14: In the pivotal study of VIVITROL in alcohol dependence where the majority of patients were not yet abstinent, there was a significant reduction in heavy drinking. Patients treated with 380 mg extended release naltrexone (VIVITROL) experienced approximately a 25% greater reduction in the rate of heavy drinking relative to the placebo-treated patients (*p* 0.03) [Garbutt, 2005 reference # 83 in the main body of the Review].	Yes, this is correct. We have already included these results in our report. We have added a reference to this study as suggested in the Discussion section on Applicability (in the sentence that says "However, some studies enrolling patients who were not yet abstinent have reported reduction in heavy drinking with naltrexone or acamprosate.").
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	ES	Table E and footnote a: There is a compelling analysis of extended release naltrexone injection in opioid dependent patients with the chronic liver condition of hepatitis C and HIV infection. The results indicate that extended release naltrexone injection can be used safely in patients with opioid dependence, even those with underlying chronic hepatitis C and/or HIV infections [Mitchell, 2012 reference below]. The authors may want to consider these data in this Review for alcohol treatments. Mitchell M, Memisoglu A, Silverman B. Hepatic safety of injectable extended-release naltrexone in patients with chronic hepatitis C and HIV infection. J Stud Alc Drug Dependence. 2012;73 (6):991-997	Thank you for the comment. For this review, we are only including studies of people with alcohol use disorders. We are not including studies conducted in other populations (if they don't also have an alcohol use disorder). For many of the included drugs, published studies have evaluated harms in other populations (e.g., healthy subjects, those with depression, etc.), but assessment of those studies is beyond the scope of our review. We identified this study and excluded it because the included population did not have alcohol use disorders.
TEP Reviewer #2	Abstract	Abstract does not use sufficient caveats. Ex. OPRM1 (needs more study). Limitations should be highlighted in abstract which comes across as too definitive.	We have revised the abstract to indicate that our meta-analyses for variation in naltrexone response related to OPRM1 polymorphisms found no statistically significant difference between AA homozygotes and those with at least one G allele , but confidence intervals were wide, and additional studies are needed.

Comment #	Section	Comment	Response
Peer Reviewer #2	TOC	The Table of Contents is inaccurate I believe. It says refs start on page 87; they start on page 96.	We fixed this.
TEP Reviewer #1	Introduction (and ES)	P ES-1 & p1: With the recent introduction of DSM-5, it seems important to make a reference to it while defining the spectrum of alcohol misuse.	We have added reference to DSM-5 (and brief explanation of how it differs) to the notes below Table A in the ES, the corresponding table for the full report, and the Discussion of both the ES and full report.
TEP Reviewer #1	Introduction (and ES)	Defining maximum recommended daily/weekly alcohol consumption would also be useful.	We have added a footnote to the same tables mentioned in the response to the prior comment explaining this, as suggested.
TEP Reviewer #1	Introduction (and ES)	P ES-1 & p 2: Lifetime risk for men is cited but not for women.	We have included lifetime prevalence rates in the following sentence for both men and women.
TEP Reviewer #1	Introduction (and ES)	P ES-1 & p2: Include cancers associated with AUD in listed medical disorders.	We have added the most common cancers associated with AUDs to this list in the full report. We did not add it to the ES because we want to keep it as concise as possible (we do have “cancer” on the list in the ES).
TEP Reviewer #1	Introduction (and ES)	P ES-1&p3: "Treatment may be delivered... other approaches" - suggest to change to "Treatment may be delivered in specialized or non specialized settings, within different levels of care including outpatient, intensive outpatient, residential, or inpatient settings.	We made revisions to the wording of this section in the full report to expand the explanation. The new version is: “Treatment may be delivered via individual outpatient counseling, intensive outpatient programs using group or individual methods, alcoholism treatment centers, or other approaches. Most treatment is currently delivered in specialty settings rather than in primary care settings. Primary care providers are typically trained to refer patients with alcohol dependence for specialized treatment, and primary care providers are generally unfamiliar with medications for treating alcohol dependence.”

Comment #	Section	Comment	Response
TEP Reviewer #1	Introduction (and ES)	P ES-3 Tb B & p3 Tb 2: Add oral for acamprosate/disulfuram dosing.	Done.
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Introduction (and ES)	pES-3 & p4 The authors may want to note that the VA does have criteria for use for extended release naltrexone injection in alcohol dependence. link: www.pbm.va.gov/clinicalguidance/criteriaforuse/naltrexoneinjcriteriafornonformularyuse.doc	We have added the point that injectable naltrexone is currently nonformulary at the VA to the Discussion and we reference the document in the suggested link.
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Introduction (and ES)	pES-3 & p4 The authors may want to note in that extended release naltrexone injection (VIVITROL) is not available in the UK.	We have added this in the Intro to the full report, but not in the Executive Summary (because we're trying to keep the latter more concise).
Peer Reviewer #1	Introduction	Good	Thank you
Peer Reviewer #2	Introduction	Two areas that would ideally be mentioned in the introduction. 1. One is DSM-5 alcohol use disorders and the abandonment of alcohol dependence as a label in DSM-5 (e.g. moderate to severe DSM-5 AUD (4+ symptoms) = alcohol dependence).1-3 DSM-5 should be added to Table	We have added this as suggested. It is now in the Table.

Comment #	Section	Comment	Response
		2. The second is brief approaches to identifying primary care patients at high risk for AUD who might benefit from assessment for moderate to severe AUD.4-7	We did not add a discussion of approaches to identifying patients with unhealthy alcohol use or those with alcohol use disorders because we feel that the screening/detection literature is quite large, and is beyond the scope of this report. It would require a great deal of space to describe that information appropriately. Additional information can be found in a systematic review conducted for the US Preventive Services Task Force (Jonas et al., *Ann Intern Med.* 2012;157(9):645-654; and related full technical report available at www.ncbi.nlm.nih.gov/books/NBK99199/pdf/TOC.pdf).
TEP Reviewer #3	Introduction	Page 1 Table 1. The amount of drinking should be included in definition of risky and harmful use. Actually for harmful use, the pattern and amount of drinking may vary depending on which organ is being damaged. Also, DSM V has replaced DSM !V. DSM V no longer uses the term dependence and abuse, but different degrees of alcohol use disorders (AUD). This should be mentioned in text.	Thank you. We have added information about DSM-5 to the text in multiple places, including Table 1, and a footnote to this Table with more explanation of the change in terminology from DSM-IV (and previous) to DSM-5. We also added a footnote to provide the limits for risky drinking.
TEP Reviewer #3	Introduction	Page 2, lines 54-56: I don't know if alcohol treatment including medications is being used more frequently. Only a small percentage of patients use alcohol medications.	We agree. It is a small percentage. We have included information about utilization of medications for alcohol dependence.
TEP Reviewer #3	Introduction	Page 4 Lines 15-16:NIAAA does not have a medications guidelines. Reference 35 is the Medical Management manual for COMBINE. We are currently working with SAMHSA to develop guidelines for alcohol medications.	Thank you for the clarification. We have revised the text to explain that this is the NIAAA Medical Management Treatment Manual, and have clarified what it covers.
TEP Reviewer #3	Introduction	Page 5: KQ4: Most of the subjects recruited in US trials probably go to primary care physicians, not for their drinking but for the problems that drinking is causing. Are you more interested in how well the medications are working in the real world? If so, naturalistic studies are needed, with high external validity and low internal validity. Nonetheless, we don't have studies in primary care settings or naturalistic studies.	Thank you. We agree that there is scant evidence in primary care (although we found a couple of eligible studies conducted in primary care settings). Our review was more focused on efficacy. In the Discussion (under Primary Care header), we provide some description of findings from more naturalistic studies, and have added description of two more of them for readers who are interested in where to turn for more studies with implications for implementation in primary care.

Comment #	Section	Comment	Response
TEP Reviewer #3	Introduction	Page 5, KQ 6: Personalized medicine is vital in medications development. I suspect that if we are going to predict who is going to respond to a specific medication, it may take a various factors including genetics, epigenetics, proteinomics (biomarkers), physiological characteristics (imaging), and human characteristics and history. Pharmacogenetics is interesting but we are only in early stages of research and it is not ready for treatment settings.	We agree.
Peer Reviewer #3	Introduction	the introduction is well written and organized. As above, the only aspect that seemed to be missing is the target audience, which does not seem well defined.	Thank you. The audience is defined in the preface. The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well informed decisions and thereby improve the quality of health care services.
Peer Reviewer #4	Introduction	No concerns	Thank you
Peer Reviewer #5	Introduction	The introduction clearly outlines why effective treatment for those with AUD is of paramount importance. It provides definitions for various conditions within the AUD spectrum, described how treatment for AUD has evolved and outlines the existing guidelines for treatment of AUD.	Thank you
Peer Reviewer #6	Introduction	The introduction in general nicely outlines the scope of the problem of alcohol use disorders and the methodology used in this report to answer the posed questions.	Thank you
Peer Reviewer #6	Introduction	My biggest concern is that the introduction does not present a good rationale for the specific questions that are posed in the report and why other questions have been omitted. For example I would have liked to see the following issues addressed in the report. From a clinical perspective, the type of questions that generate the most interest include issues of 1) how long do we keep patients on the medication, 2) what is the role of medication adherence in treatment outcomes, 3) is there a particularly good psychosocial approach to enhance treatment effectiveness, 4) do we need to detoxify patients before we initiate treatment, 5) what are the effects of these medications on craving and 6) are these medications safe to interact with other psychiatric and medical medications? This review does little to address these clinical issues. As mentioned above a thoughtful review of the literature could at least give our best guess at the moment as to answers of these questions. Clearly much work needs to be done but where data is lacking the report could point out the deficits in our knowledge.	We have revised the introduction and methods section, and have clarified the rationale for the key questions. As we point out in the Introduction, medications for alcohol use are underutilized, with a small percentage of eligible patients receiving medications. It is important to address the basic questions of whether or not the medications work, and which ones work best, and (if they do) to get that message to a large number of providers who might prescribe the medications.

Comment #	Section	Comment	Response
Peer Reviewer #6	Introduction	To be fair, the report does a credible job of pointing out gaps in empirical data such as the lack of good primary care studies, but the report barely exposes the tip of the ice burg when it comes to directly treating patients. At this point in our evolution, the question of does naltrexone and acamprosate work is far less interesting than how do we optimize treatment outcomes.	(this was a continuation of the previous comment). Thank you. Evidence about efficacy, comparative effectiveness, and harms is important information. Together with evidence about health care delivery, adherence, etc, it can lead to improved health care quality and health outcomes.
Peer Reviewer #6	Introduction	Similarly, as a policy maker or health insurer, the types of question I would be most interested are not addressed in this report. For example, I would want to know the health and economic consequences of offering and paying for treatment. This report correctly highlights that paucity of data looking at non-drinking health outcomes, but I believe there is a body of data that looks at social and economic consequences (see Baser, 2011, Am J. Managed Care, 17;S222-234 and Zarkin, 2010 Med Care 48:396-401).	We looked for evidence on health outcomes, and we have expanded the discussion section related to health outcomes. We have added a brief explanation of the Zarkin paper (cost-effectiveness from the COMBINE study) mentioned here, as it was directly relevant to the material in our discussion about health outcomes. The Baser paper did not report an eligible outcome.
TEP Reviewer #1	Methods	I/E criteria are justifiable. Search strategies are explicitly stated and logical.	Thank you
TEP Reviewer #1	Methods	Outcome measures are mentioned within the analytic framework (introduction section) but not clearly defined within the methods section. Some such as "number of heavy drinking days" require a more explicit definition. In addition, it is unclear how "accidents" differ from "injuries".	We have added the definition of heavy drinking days as a footnote to the I/E table (Table 3) of the full report. Heavy drinking days were defined as ≥4 drinks/day for women and ≥5 drinks/day for men. We also added a footnote to explain the issue raised here for accidents and injuries. Accidents typically refer to motor vehicle accidents. Injuries may be from a wide variety of alcohol-related problems (e.g., violence, falls).
TEP Reviewer #1	Methods	Statistical methods used are appropriate.	Thank you
Peer Reviewer #1	Methods	OK	Thank you
TEP Reviewer #3	Methods	Page 7, Methods: You might want to make the description of your meta-analysis more lay-friendly, depending on the readership.	We try to provide the information in multiple formats. This section of the full technical report is fairly detailed and technical, and is intended to be. The Executive Summary provides a less technical version, and the abstract an even less detailed, very high-level version.

Comment #	Section	Comment	Response
TEP Reviewer #3	Methods	Page 8, line 39: should also include levetiracetam (see attachment for article)	Levetiracetam was not an included medication because it is not commonly used for alcohol use disorders and searches during our topic refinement period (prior to conducting the review) suggested that there would be too few studies to synthesize. Also, discussions with key informant did not identify it as one of the clinically important medications to include.. We did not include every possible medication in this report. We focused on the medications deemed most important to decision makers.
TEP Reviewer #3	Methods	Page 8, Table 3: formatting, sentence runs into exclusion part of table.	We fixed this.
TEP Reviewer #3	Methods	Page 9, Line 5: Why is craving excluded? It is part of DSM V.	We focused on the outcomes of most importance to decision makers. The process for determining the scope is described in the Methods section and in the topic refinement and review protocol. From the judgment of our team, Key Informant input, TEP input, and public input, we did not feel that it was among the most important outcomes. Limitations in time and resources don't allow us to include every single outcome.
TEP Reviewer #3	Methods	Also, for the record, the FDA has designated % subjects with no heavy drinking as the primary endpoint for pivotal trials. However, you will not find that endpoint in most of the articles.	We included this outcome in our review and we found evidence on this outcome in many trials.

Comment #	Section	Comment	Response
TEP Reviewer #3	Methods	Page 9, Line 15: In every alcohol pharmacotherapy trial that I reviewed, if there is an effect of a medication, you see if by week 8. Although most alcohol clinical trials are at least 12weeks, there are some at 8 weeks. I am not sure why the exclusion is less than 12 weeks.	We set this cutoff based on epidemiologic literature, and after discussions with our Key Informants and TEP. We address this issue in the Limitations also: "We required that trials have at least 12 weeks of followup from the time of medication initiation, excluding trials of shorter duration. Some might consider this approach to omit potentially important information. However, longitudinal studies have found that shorter treatment periods may yield misleading conclusions about treatment efficacy, due to fluctuations in drinking behavior that are typical of the course of alcoholism[59,60]—suggesting that longer durations of followup might more accurately reflect the outcomes of greatest interest and importance."
TEP Reviewer #3	Methods	Page 10, lines 25-28: medication compliance is also very important. In fact, low dropouts and high medication compliance are essential in evaluating a medication.	Assessment of medication adherence was part of our assessment of the individual trials (we assessed fidelity and adherence as part of our risk of bias assessment).
TEP Reviewer #3	Methods	Page 10, Data Synthesis section: You might consider adding a sentence or two explaining what you did for those who may not be familiar with meta-analysis techniques.	We did not make any revisions to this part of the full technical report. Instead, we try to provide the information in multiple formats. This section of the full technical report is fairly detailed and technical, and is intended to be. The Executive Summary provides a less technical version, and the abstract an even less detailed, very high-level version.
Peer Reviewer #3	Methods	The authors report on the types of studies which are included in this review: the strategies are explicit and logical and the measures are appropriate with reasonable statistical methods.	Thank you
Peer Reviewer #4	Methods	the methods were good but very strict in terms of study inclusion.	Thank you
Peer Reviewer #5	Methods	The authors provide detailed description of the method they employed to select studies for this report. They also provide a list of studies they excluded and reasons for excluding them. The inclusion and exclusion criteria are well justified and the statistical approach is valid and scientifically sound.	Thank you

Comment #	Section	Comment	Response
Peer Reviewer #6	Methods	One methodological concern is that the authors based treatment effects on a random effects model to estimate pooled effects. While a random effects model is appropriate given the heterogeneity of the studies, this assumes that the continuous outcome measures (both within each study and between studies) are normally distributed. I suspect for outcomes such as percentage of drinking days and percentage of heavy drinking days, the data are highly skewed we may be underestimating the true effect size by failing to take this into account. For example, in a multi-centered trial conducted by the VA, the initial conclusion was that naltrexone offered no benefit over placebo. A subsequent analysis however, using typologies of outcomes, no drinking, low drinking, or high drinking (essentially converting the data into categorical non-parametric outcomes) did find a statistical difference in favor of naltrexone. What are the consequences to the meta-analysis if researchers have not appropriately presented their data?	We agree that a random effects model is appropriate given the heterogeneity of studies. "We acknowledge the reviewers point about understanding the assumptions for MA methods, and choice of the best methods. We followed the AHRQ methods guidance chapter on "Handling Continuous Outcomes in Quantitative Synthesis" by Rongwei Fu, PhD, et al. We noted that most of our included studies reported means and standard deviations (SDs), and noted that these suggested a normal distribution. We also assessed whether the mean is smaller than twice the SD in each intervention group, which would indicate that the data of an original study are likely to be skewed. Based on this information we felt that using the random effects models was appropriate. Further, many of the analyses and main conclusions are largely based on dichotomous outcomes (e.g., return to any drinking or return to heavy drinking), not continuous outcomes.
Peer Reviewer #6	Methods	In summary rather than describing the percentage of subjects in each study were smokers (how is this relevant?) the authors would do well to consider how each study and how the meta-analysis handled non-normally distributed data.	The percentage of smokers is relevant because it was considered a potential modifier of treatment effect. Regarding the normal (vs. not) distribution of continuous outcomes in the included studies, we followed the AHRQ methods guidance chapter on "Handling Continuous Outcomes in Quantitative Synthesis" by Rongwei Fu, PhD, et al. We noted that most of our included studies reported means and standard deviations (SDs), and noted that these suggested a normal distribution. We also assessed whether the mean is smaller than twice the SD in each intervention group, which would indicate that the data of an original study are likely to be skewed. Based on this information we felt that using the random effects models was appropriate.

Comment #	Section	Comment	Response
Peer Reviewer #2	Methods & Results	1. Implications for Primary Care (ES-13). I have several comments that are aimed at making the report more useful to primary care clinicians. I recognize that some of them would have been most useful at the stage that the KQs were refined, but I thought that they might still be addressed. a) Do patients need to abstain for medications to be effective? Although there is a paragraph on this topic in the Executive summary (ES), it was not quantitative, systematic or particularly helpful. Consider addressing this issue head on in the tables and tables: whether trials required abstinence prior to medications and if so for how long, and whether trials required patients say they had a goal of abstinence for eligibility? I suspect a number of users will want to have this easy to find. Given that acamprosate had 0 patients assessed for % heavy drinking days and NTX had 1423, it seems likely that the expectations for abstinence might have varied a lot between studies of the 2 medications (% heavy drinking days is mostly relevant if patients are drinking).	a) We have included text on this issue in the Applicability section. The vast majority of studies required patients to abstain for at least a few days prior to initiating medication, and the medications are generally recommended for maintenance of abstinence. Acamprosate and injectable naltrexone are only approved for use in patients who have established abstinence, though the duration of required abstinence is not set. However, we point out and reference studies enrolling patients who were not yet abstinent that have reported reduction in heavy drinking with naltrexone or acamprosate, with references to those studies. We have also added this issue (about whether patients need to abstain for medications to be effective) to the evidence gaps table (as suggested by this same reviewer in another comment), highlighting it as something that could be addressed by future studies.
Peer Reviewer #2	Methods & Results	b) How severe was the alcohol dependence of patients in each study? As alluded to patients with alcohol dependence in primary care settings likely have more mild alcohol dependence than those from treatment settings (compare the AUDIT in COMBINE to primary care studies).5,7,8 This is only mentioned on page ES-14. Just as recruitment methods are described in the Tables, a description of severity would be useful in the tables (# of AUD symptoms, # prior treatment episodes, etc.). This important issue needs more discussion and data. Since clinicians now need to map prior studies on to the new DSM-5 criteria (11 symptoms, 2-3 mild Alcohol dependence [AD], 4-5 moderate AD and 6+ severe AD), it will be important to report the severity of AUD among patients in prior studies.	We acknowledge that different studies may have recruited patients with different severity of alcohol dependence. We were not able to stratify our analysis based on this population characteristic due to limited reporting in individual studies. We added the following to the Applicability section of the Discussion: "The included literature used definitions from DSM-III or DSM–IV. DSM-5 (2013) describes a single AUD category measured on a continuum from mild to severe, and no longer has separate categories for alcohol abuse and dependence.[25] Using DSM-5 terminology, most participants in the included studies likely had moderate to severe AUDs. Thus, applicability of our findings to people with mild AUDs is uncertain."
Peer Reviewer #2	Methods & Results	Given that abuse has been shown to not be a useful construct and is being abandoned,1 I'd not focus attention there (e.g. ES-14 Table E # 1c).	We have deleted the research gap on alcohol abuse from the Table as suggested. For the text, the new portions on DSM-5 (described in responses to various other comments) have clarified that part of the text.

Comment #	Section	Comment	Response
Peer Reviewer #2	Methods & Results	c) It appears "encouragement trials" were excluded even for KQ#4? If not, why not, given the KQ primary care. Two articles that I would classify as encouragement trials were recently published and should—I believe—be included in the review. i. Care management for alcohol dependence. A recent VA trial of primary care "care management" for alcohol dependence focused on naltrexone. Based on the consent form (which the first author was kind enough to share with us) the patients were enrolled in a trial where they would be offered naltrexone. Thus while the study did not appear to meet your inclusion criteria for efficacy studies, this is a critically important effectiveness trial. 9 ii. The AHEAD trial (Saitz JAMA September 2013) of chronic care management (CCM) for alcohol and other substance use disorders (SUD) should also be mentioned. Key points are that this trial enrolled only patients who were NOT currently enrolled in primary care (74% from a public detox program) and only 10-12% had AUD without other SUD (60-67% heroin or cocaine, up to 56% homeless). The trial arms were routine primary care + access to a few sessions MET + info on treatment VERSUS CCM by an interdisciplinary team in a special primary care clinic (that encouraged medications for AD as well as opiate dependence). Both groups had better than expected outcomes but there was NSD between groups.10 This is important to mention because it does NOT undermine the likelihood that primary care management with medications can improve outcomes in primary care since this study recruited patients who were NOT engaged in primary care.	We required studies to compare one of the medications listed vs. placebo or another medication. Neither study meets that criteria. We agree that these studies deserve mention because they may have implications for providers who want to implement systems of care for people with alcohol use disorders. We have added them to the Discussion (both ES and full report) sections on primary care, and we highlight what they compared and their main findings in the Discussion. These studies do not isolate the comparison of interest, as there are multiple differences in the interventions between the groups. They compare care management programs with a control group, rather than use of a medication with placebo or with another mediation. New text in the report in the : "…Third, another U.S-based RCT (N=163) compared a primary-care based Alcohol Care Management (ACM) program with a specialty outpatient addiction treatment program (SC). A greater proportion of the ACM group received naltrexone than the SC group (65.9% vs 11.5%), the ACM group had a higher proportion of participants engaged in treatment over the 26 weeks, (OR 5.36, 95 % CI, 2.99 to 9.59), and the percentage of heavy drinking days was lower in the ACM group (OR 2.16, 95 % CI, 1.27 to 3.66). Overall abstinence did not differ between groups. Forth, the U.S.-based AHEAD trial (N=563) compared chronic care management (CCM) that included longitudinal care coordinated by a primary care clinician with no CCM for people with alcohol or drug dependence who were not currently engaged in primary care. Of those enrolled, 12 percent had alcohol dependence without also meeting criteria for other drug dependence. CCM included motivational enhancement therapy, relapse prevention counseling, on-site medical, addiction, and psychiatric treatment, social work assistance, and referrals.

Comment #	Section	Comment	Response
Peer Reviewer #2 (continued)	Methods & Results	c) It appears "encouragement trials" were excluded even for KQ#4? If not, why not, given the KQ primary care. Two articles that I would classify as encouragement trials were recently published and should—I believe—be included in the review. i. Care management for alcohol dependence. A recent VA trial of primary care "care management" for alcohol dependence focused on naltrexone. Based on the consent form (which the first author was kind enough to share with us) the patients were enrolled in a trial where they would be offered naltrexone. Thus while the study did not appear to meet your inclusion criteria for efficacy studies, this is a critically important effectiveness trial. 9 ii. The AHEAD trial (Saitz JAMA September 2013) of chronic care management (CCM) for alcohol and other substance use disorders (SUD) should also be mentioned. Key points are that this trial enrolled only patients who were NOT currently enrolled in primary care (74% from a public detox program) and only 10-12% had AUD without other SUD (60-67% heroin or cocaine, up to 56% homeless). The trial arms were routine primary care + access to a few sessions MET + info on treatment VERSUS CCM by an interdisciplinary team in a special primary care clinic (that encouraged medications for AD as well as opiate dependence). Both groups had better than expected outcomes but there was NSD between groups.10 This is important to mention because it does NOT undermine the likelihood that primary care management with medications can improve outcomes in primary care since this study recruited patients who were NOT engaged in primary care.	The no CCM group received a primary care appointment and a list of treatment resources including a telephone number to arrange counseling. The trial found no difference between groups for the primary outcome of abstinence over 12 months."
Peer Reviewer #2	Methods & Results	2. Disulfiram: a. Pre-op disulfiram: One review of disulfiram (Jorgensen et al) also included Tonnesen BMJ 1999 which randomly assigned supervised disulfiram 800mg to pre-op patients who drank heavily before elective colectomy (ref below). While it did not include the required outcomes, and was not blinded, given the limitations on the reviewed literature on disulfiram generally, it seemed worthy of mention.11	The Jorgensen review is included in our report. The Tonnesen trial does not meet our criteria because of the comparison group (comparison was routine care, and was not placebo or another medication) and because the treatment duration was only 1 month. We do not agree that this study should be mentioned, given the limitations mentioned by the reviewer along with the additional reasons that it didn't meet our criteria.

Comment #	Section	Comment	Response
Peer Reviewer #2	Methods & Results	b. Supervised disulfiram. Most clinicians who use disulfiram believe that monitored disulfiram is required for it to be effective. The review would ideally have addressed this both in the tables and more in the descriptive summary of results.	Thank you, very interesting. We have not had anyone else raise this issue as one of importance. We re-reviewed the 4 included disulfiram studies after getting this comment and they don't indicate using supervised disulfiram. One (Fuller '79) said that "…Intake of disulfiram was monitored throughout the year of the study." But, the surrounding description is more indicative of routine follow up similar to medical management rather than supervised administration.
Peer Reviewer #2	Methods & Results	c. A prior review that included disulfiram: I am not sure that this review was referenced.12	This is a single author review and analysis published in Addiction in 2005 that used a Swedish database (the database was published in English in 2003). This article does not have a methods section and does not describe a broad literature search, dual review, or other methodology that we would require to consider it for inclusion as a systematic review.
TEP Reviewer #1	Results	The amount of detail presented in the results section is mostly appropriate. CIs and/or p-values are not always included - such as on p32, some RD and WMD values are listed with no CI.	Thank you. We don't include the CI for every sensitivity analysis in the full report Results in an attempt to make the results less cumbersome, and to focus on the main results.
TEP Reviewer #1	Results	Characteristics of the studies are clearly defined. Key messages are explicit and applicable. This section would be more readable if there was more space between the major sections. In addition, fonts used for titles/subtitles do not always match. In many places, space will need to be deleted. Some text has been misplaced between pp34 & 35.	Thank you We fixed the fonts and formatting issues. We will format the report per the AHRQ publishing guidance.
TEP Reviewer #1	Results	Figures, tables and appendices are mostly adequate and descriptive.	Thank you
TEP Reviewer #1	Results	In appendix B, the term "wrong" could be replaced by "inadequate" or other similar term.	Thank for your suggestion. We changed it to "Ineligible" (e.g., ineligible population, ineligible intervention).

Comment #	Section	Comment	Response
TEP Reviewer #1	Results	In appendix C, tables C1/C2, C4/C5, C7/C8 have identical titles which is confusing.	They are slightly different, as the latter of each pair has "(continued)" at the end. Because these are large evidence tables, what used to be a single table in excel has been split into 2 tables to fit on the pages. For example, C1/C2 used to be a single table and both contain some of the risk of bias questions. We don't have a good idea for what to do about this other than to add "(continued)" to the table titles as we've done.
TEP Reviewer #1	Results	Investigators did include/exclude studies appropriately.	Thank you
Peer Reviewer #1	Results	Good	Thank you
TEP Reviewer #3	Results	Page 14, Characteristics of trials: I noticed in some of the descriptions of the various medications you state something about the bias, while in others you don't. For example, the risk of bias was not mentioned for acamprosate. Also several of the acamprosate studies were conducted for 6 to 12 months. I wonder about the dropouts and how they handled missing data in thier outcome analysis.	We have rechecked each section to ensure that information about risk of bias is included in the relevant tables and data synthesis portions of the report. The details for dropout rates from each study and of our assessments of adequacy of handling missing data are provided in the Risk of Bias appendix. We have chosen provide this information in the appendix rather than the body of the report to improve its readability.
TEP Reviewer #3	Results	Page 15, Table 5: It would be nice to see the dropout and compliance per arm (especially for disulfiram studies) so that the reader can judge what an unacceptable attrition and compliance rate might be. Also how was missing data handled.	These details are provided in the Risk of Bias appendix.
TEP Reviewer #3	Results	Page 19, outcomes: It is not clear what the bottom line is for the outcomes for acamprosate as well as the other medications. I guess in the discussion section you do give the bottom line. Interesting, with acamprosate and naltrexone you will get some outcomes that are positive while others are not. What does that mean?	We have also revised the abstract, executive summary, and the discussion for the final report to make the main messages more clear. Having some outcomes that are positive and others that are not means that they work for improving some outcomes, but not for others. We describe what they have been shown to improve (or to not improve, or when evidence was insufficient to determine) in the Results and Discussion.

Comment #	Section	Comment	Response
TEP Reviewer #3	Results	Page 19, line 22: What was the difference for the shorter trials?	As part of our attempt to focus on the most important information and to keep the length of the report somewhat manageable, we did not provide the data from every single stratified portion of every meta-analysis in the text. We focused on the most important findings. The details for this particular question are shown on the forest plot (Figure F-3) and not in the text.
TEP Reviewer #3	Results	Page 19, lines 5-11: I like this sentence which says of x studies, Y found a positive and significant effect. Please include this for each paragraph where an outcome's results are described	Thank you. We do not provide this information for every paragraph under every outcome. We feel that such an approach can be misleading and can distract from the main findings when used indiscriminately. It is appropriate and clear (and not misleading) when there are smaller numbers of studies that need to be described qualitatively. But, when there are larger numbers of studies, and a meta-analysis is conducted, the overall effect, confidence interval, and heterogeneity are the more appropriate way to describe the information.
TEP Reviewer #3	Results	Page 19 line 42: Again, what was the differences for the short duration trials?	The details for this particular question are shown on the forest plot (Figure F-7) and not in the text.
TEP Reviewer #3	Results	Page 20, Disulfiram: As mentioned above, there were not a lot of studies on disulfiram. In regards to the studies that were conducted, I suspect the compliance rate was low, particularly with Fuller et al 1986. Undoubtedly, there is a subpopulation that this medication works well, e.g. abstinent before taking the medication, socially stable, social support. As I said before, if they take the drug at the right dose, it is unlikely that drinking will occur.	We have added information about disulfiram in several places in the report, including in the main results that are described in the abstract, and in the ES and the discussion.
TEP Reviewer #3	Results	Page 22, Naltrexone: You might want to include two recent naltrexone articles (see attachment). Foa et al found an effect in alcohols with PTSD and Karl Mann conducted a multisite study of acamprosate and naltrexone. The design of the latter study was similar to COMBINE. Interestingly, the results were negative for both naltrexone and acamprosate.	We have reviewed these studies against our inclusion/exclusion criteria. Indexing of the Mann study was not complete at the time of our literature searches (thus our literature searches did not find it). The Mann study (Mann K, Lemenager T, Hoffmann S, et al. Addict Biol. 2012 Dec 12) is eligible and we have added it to the report. The Foa study did not meet our eligible study design criteria (Foa EB, Yusko DA, McLean CP, et al. JAMA. 2013 Aug 7;310(5):488-95). Our updated literature searched identified this study.

Comment #	Section	Comment	Response
TEP Reviewer #3	Results	Page 32, Results of naltrexone: Just an observation, most of the acamprosate studies conducted in Europe in the 90s had similar design (was orchestered by Lipha). The results were similar in terms of abstinence and number of days abstinent (drinking amounts were not measured). However, for the naltrexone studies, many were supported by NIAAA and were conducted on a variety of conditions and designs. Naltrexone was very consisted in exhibiting a small effect across these studies, while acamprosate studies had difficulty showing an effect outside of Europe.	Thank you. We include information in the Discussion about differences in findings (and where studies were conducted) for U.S. vs. non-U.S. studies: "Although the majority of included trials assessing the efficacy of acamprosate were conducted in Europe (16 of 22) and a minority were conducted in the United States (4 of 22), the opposite was true for naltrexone (27 of 44 in the United States and 8 of 44 in Europe). Further, the few studies of acamprosate conducted in the United States did not find it to be efficacious. It is unclear whether the different results were due to population differences or other factors. The European trials of acamprosate typically identified patients from inpatient settings or treatment programs, whereas the U.S.-based trials of acamprosate relied on advertisements and referrals. It is possible that this resulted in populations with differing alcoholism severity and differing potential for benefit. For example, studies of subjects recruited via advertisements may enroll people who have less severe disorders, and may be less applicable to patients with more severe forms of alcohol-use disorders."
TEP Reviewer #3	Results	Page 33, baclofen results: Why were the results so different between the 2 studies? I guess the answer is more studies need to be conducted.	Yes, it is hard to know. We agree that more studies are needed to determine whether it works (and we have included that in our evidence gaps table). They are two fairly small studies (with Ns around 80) and it is probably within the realm of plausible random error. Also, they were conducted in different countries and somewhat different populations as shown in the table of study characteristics.
TEP Reviewer #3	Results	Page 36, Citalopram: I wonder if it might be better to group all the SSRIs into one group. I doubt if there are any differences among them. Interestingly, we funded two tricyclic antidepressant studies in the 90s with more consistent results. We thought that perhaps the tricyclics may work better in depressed alcoholics although the side effects are worst.	This is an interesting question. If we found evidence that a single one of them worked, then we might have considered this to be more reasonable to pool data across the SSRIs. Our a priori plan was to analyze each medication separately. Pooling the SSRIs could be an interesting post-hoc analysis for some investigators to conduct.

Comment #	Section	Comment	Response
TEP Reviewer #3	Results	Page 39, Nalmefene: Yes, please include the two or three recent studies by Lundbeck. They also found that those who had higher drinking levels at baseline had the best outcomes with nalmefene.	We have added the new nalmefene studies to the report. We found these studies in our updated searches.
TEP Reviewer #3	Results	Page 41, quetiapine: Please include this multi-site study on quetipaine. It will not change your conclusions.	The study was <12 weeks duration, and therefore does not meet our eligibility criteria (Litten, Fertig, Falk, et al. Alcoholism: Clinical and Experimental Research, Vol. 36, No. 3 March 2012).
TEP Reviewer #3	Results	Page 47, Head to Head Trials: It is not surprising that you didn't find any differences since all had small effect sizes. The only chance of an effect might be if a study recruited more of a population that responded to particular medication than the other. Remember we believe that alcohol medications work for some but only for all (or most).	Thank you.
TEP Reviewer #3	Results	Page 49, Disuflram versus naltrexone: I thought in the meta analysis you were not going to give the results for just one trial.	The part of our report that this comment is referring to was specifically about placebo-controlled trials of medications used off-label—we explained that when we just found 1 study for a drug, that we moved all that information to an appendix and it is not in the main body of the report. This disulfiram vs. naltrexone section (about head-to-head studies of FDA-approved meds), therefore does describe the one study that we found. We describe what it found (we didn't meta-analyze it, since it is just 1 study).
TEP Reviewer #3	Results	Page 52, sertraline vs naltrexone: This study did find an effect for a combination of naltrexone and sertraline.	We did not review combinations of medications in this report.
TEP Reviewer #3	Results	Page 53, lines23-25: what was NNT for the Cochrane report?	We have added the NNT to the text.

Comment #	Section	Comment	Response
TEP Reviewer #3	Results	Page 53, Health Outcomes: Health outcomes can be divided into short-term (acute) and long-term outcomes. The short-term outcomes are usually acute, involving social, accident, injuries, arrest, job performance, etc. Long-term outcomes involves medical consequences, events that may take years of chronic drinking before seeing the results. Included in long-term are mortality and morbidity. The outcomes in clinical trials involved primarily short-term consequences. In fact, some of the trials may not go long enough to see these from occuring. Several epidemiologicql studies have correlated chronic drinking levels with various medical consequences. There has also been some question whether or not small changes in drinking really mean anything. However, recent data suggest that this may not be the case. For example, Rehm et al Addiction 106 Suppl 1:11-19, 2011 (see attachment) reported that any reduction in the dose of alcohol consumed, down to 10 g/day (one drink), will reduce the annual and lifetime risk of an alcohol-related death. A new publication will be coming out soon correlating reduction of heavy drinking versus consequences (again as few as one or two heavy drinking days have a long-term effect).	Thank you for this comment. We have expanded the Discussion portion of our report that addresses health outcomes and these issues, including references to some of the work by Dr. Rehm.
		Page 69, Adverse Effects: What are the methods used to reporting adverse effects? Methods probably varied among studies. Also, a general comment: all medications have a risk/benefit ratio. However, acamprosate and naltrexone are generally considered fairly safe medications. In all the trials, there have been either none or only a few serious adverse effects. The adverse event profile of disulfiram is probably more severe, although at doses of 250-500 mg/daily has not been problematic (see an article by J Chick, Drug Safety 20:427-435, 1999). Of course, topiramate probably has the worst adverse event profile, although it is not any different in patients taking topiramate for seizures and migraines.	Yes, that is correct (methods varied). We highlight this as one of the Key Points in the Adverse Effects key question: “Adverse events were often not collected using standardized measures, and methods for systematically capturing adverse events were often not reported.” And “While major harms were rarely reported in the studies, some minor harms (e.g., diarrhea) were reported more consistently.”
TEP Reviewer #3	Results	Page 81, lines 12-15: See the Foa paper (see attachment) for naltexone trial in PTSD alcoholics.	This study does not meet our eligibility criteria (Foa, Yusko, McLean, et al. JAMA. 2013; 310(5): 488-495). It is a single-blind (not double-blind) study that reports drinking outcomes and PTSD outcomes. We required studies to be double-blind RCTs to be included for drinking outcomes.

Comment #	Section	Comment	Response
TEP Reviewer #3	Results	Page 81: Genetic Polymophisms: The most impressive data so far on genetic polylmorphism is the work of Bankole Johnson with ondansetron (see attachments). Basically, what he found was that there are 5 genetic variants, two on the serotonin transport and three on the 5-HT3 gene, that are favorable to ondansetron. Approximately, one third of the alcoholic population has one of the five favorable variant. Also, he found that if the patients had more than one favorable variant, there was an enhancement of the outcome with ondansetron. Of course, this needs to be reproduced at other sites. Finally, Hank Kranzler (see attachment) also had an interesting findings with outcome for sertraline versus LL vs SS on the sertonon transporter, although it is preliminary.	Thank you for this comment. We have reviewed the Johnson studies (Am J Psychiatry 2011; 168:265–275 and Am J Psychiatry 2013; 170:1020–1031). Neither study meets our eligibility criteria. Both were less than 12 weeks in duration. These are in the appendix of excluded studies. The Kranzler study is included in the report.
Peer Reviewer #3	Results	There is an enormous amount of detail in this manuscript. While this was an impressive overview of the literature of pharmacologic treatments for alcohol dependence, at times the information seemed a bit redundant. Nevertheless, this is an important and clinically useful document.	Thank you. We hope that the Executive Summary and a journal article that stem from this work will provide the more concise version of the information for readers who prefer that.
Peer Reviewer #3	Results	The inclusion of a section on use of medications in primary care was a real plus, one of the most useful parts of the document.	Thank you
Peer Reviewer #3	Results	It may be difficult to include every study, but some medications that are currently being studied that were not mentioned include prazosin and varenecline.	We included both of these medications in our searches (and both are listed as eligible medications in our inclusion/exclusion table). We found no eligible studies for prazosin and we found 1 for varenicline. The one for varenicline was found from peer review comments (it had not yet been indexed at the time of our literature search); it was not in the draft report sent out for peer review. It has been added to the final report.
Peer Reviewer #4	Results	I had no issues here	Thank you
Peer Reviewer #5	Results	This is well written and well organized report. The results are clearly stated and tabulated. If anything the report is somewhat overwhelming in its length and detail.	Thank you. We agree that the full report is long and detailed. We hope that the Executive Summary and a journal article (or 2) will serve to provide a more concise, and less detailed, version of the material for readers (most) who are looking for those qualities.

Comment #	Section	Comment	Response
Peer Reviewer #5	Results	Although the authors report on some promising treatment options for the population under study, it may be useful to include data on three additional medications. Zonisamide, prazosin and varenicline have all been investigated for their use in individuals diagnosed with AD. The data for these medications seems promising and it is worth including these findings to raise awareness of regarding other treatment options.	Both prazosin and varenicline were eligible. We included both of these medications in our searches (and both are listed as eligible medications in our inclusion/exclusion table). We found no eligible studies for prazosin and we found 1 for varenicline. The 1 for varenicline was found from peer review comments (it had not yet been indexed at the time of our literature search). It has been added to the final report. Zonisamide was not eligible. We developed the list of eligible medications as described in the Methods section. It could be included in future updates if literature on it emerges.
Peer Reviewer #6	Results	The authors do a fairly good job of presenting their results, however, it is not so clear how the researchers came to their conclusions to include or exclude a study. This is one area where the authors made an intellectual contribution to the data analysis than simply rely on brute force. The authors use “risk of bias” as a reason to exclude studies from their meta-analysis. From a practical perspective we are trying to weed out studies that are biased and thus unlikely to be replicated by consequent unbiased studies. It is ironic that the methodology used here managed to discard to the trash bin the original study that first showed naltrexone to be an effective treatment for alcohol dependent subjects. The study found that naltrexone treatment reduced relapse rated to heavy drinking, alcohol craving, and did not have a substantial effect to reducing risk of a slip to any drinking. Twenty one years later with over 40 randomized trials conducted in four continents by dozens of researchers, the conclusions from the 1992 manuscript remain intact.	Thank you. It is correct that our main analyses did not include studies rated as high or unclear risk of bias. But, such studies were included in sensitivity analyses to assess whether their inclusion would lead to different findings. For overall decisions about inclusion or exclusion, our criteria are detailed in the Methods section. The comments here are about the following study: Volpicelli JR, Alterman AI, Hayashida M, et al. Naltrexone in the treatment of alcohol dependence. Arch Gen Psychiatry. 1992 Nov;49(11):876-80. We understand the problem here. We erroneously listed this study in the Appendix of Excluded studies. It is, in fact, an included study. However, fixing that listing does not change our findings because the study reports data from a subset of subjects from Volpicelli 1995, which was already included in the report and in our sensitivity analyses (because of the risk of bias rating).

Comment #	Section	Comment	Response
Peer Reviewer #6	Results	Arguably the most reliable treatment effect in alcohol treatment research (both medication and psychosocial treatments) was first reported in a manuscript that was not even included in the data analysis because it had the "wrong outcome measure". The study reported on a survival analysis for return to heavy drinking and somehow this was not an appropriate outcome for the data analysis. If survival analysis was too complicated an outcome to be used in the data analysis it is a simple matter to calculate the percentage of subjects who simply returned to heavy drinking, a drinking that was used in the other studies presented in the meta-analysis. The omission of this study and other such omissions raises questions as to the methodology used to weed out studies.	This is a continuation of the previous comment. As noted in our response above, we have included the study in question here. In our tables of included studies, it now shows up in the same rows with Volpicelli 1995. We have added a footnote stating: "Data entered is from Volpicelli 1995, which reported pooled results of 99 subjects. Data from a smaller subset (N=70) of this sample was reported in Volpicelli 1992. For our data analyses, we used data from Volpicelli 1995 to use the larger, more complete sample and did not use data from Volpicelli 1992 to avoid double counting." We also inserted similar footnotes under the forest plots that contain data from Volpicelli 1995 to clarify why Volpicelli 1992 is not also included in them. We have double-checked our inclusion/exclusion and did not find any other such omissions, and no others were identified by reviewers.
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Results	p. 22: VIVITROL was evaluated at doses 190 mg and 380 mg per month, reference 83: Garbutt et al.	Thank you. We fixed this (changed per day to "per month").
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Results	p. 61: 190 mg per month (change twice)	Thank you. We fixed this.

Comment #	Section	Comment	Response
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Results	Return to Any Drinking While the authors state that the effect on return to heavy drinking did not reach statistical significance for injectable naltrexone, it should be noted that VIVITROL, (380 mg per month extended release naltrexone injection), was approved by FDA based on sustained abstinence in a sub group of patients. Among the subset of patients (n=53, 8% of the total study population) who abstained completely from drinking during the week prior to the first dose of medication, compared with placebo-treated patients, those treated with VIVITROL 380 mg were more likely than placebo-treated patients to maintain complete abstinence throughout treatment [1]. 1. VIVITROL prescribing information Alkermes, Inc. Revised July 2013 accessed October 22, 2013 www.vivitrol.comlContent/pdflprescribing info.pdf	Consistent with our planned synthesis approach, we did not use subgroup analyses to make conclusions about whether medications have evidence of efficacy. We have not routinely reported data from all such subgroup analyses for other medications, and thus do not think it would be consistent for us to highlight this subgroup analysis from one Vivitrol study.
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Results	Return to Heavy Drinking (p 32) The authors state that the effect on return to heavy drinking did not reach statistical significance for injectable naltrexone. However, we would like to highlight that the event rate of heavy drinking was the primary outcome measure for the pivotal study in alcohol dependence. Patients treated with 380 mg extended release naltrexone (VIVITROL) experienced approximately a 25% greater reduction in the rate of heavy drinking relative to the placebo-treated patients *(p* ::: 0.03) [Garbutt, 2005 referenced in the review].	This study is included in the report, and it is included in our meta-analysis for heavy drinking days. It is not in the meta-analysis for return to heavy drinking, because the study enrolled patients who were still drinking (rather than those with required abstinence who could then possibly return to heavy drinking).
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Results	Heavy Drinking Days (p 32) A pre-specified analysis of percent heavy drinking days revealed a significant difference between extended release naltrexone 380 mg and placebo *(P* ::: 0.0017) in the pivotal study. The median per subject event rate (percent heavy drinking days) was 10.2% for the extended release naltrexone 380 mg extended release naltrexone (VIVITROL) treatment group and 19.8% for the placebo group following the last dose in the study (includes all subject drinking data utilizing LOCF) [4]. 4. Data on file from Clinical Study ALK2I-003, Alkermes Inc, Waltham MA	This information is from the Garbutt 2005 study that we have already included in the review. We added the data sent by the commenter to the relevant analyses.

Comment #	Section	Comment	Response
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Results	Drinks per Drinking Day (p 32) A post-hoc analysis of the pivotal study was performed on drinks per drinking day. For the 30 days pre-baseline, the average number of drinks per drinking day in all subjects were: mean (SD) of 8.5 (4.6) and 8.5 (4.0) for the 380 mg extended release naltrexone (VIVITROL) and placebo groups respectively. The post baseline (up to the last dose + 30 days) number of drinks per day were: mean (SD) of 4.8 (3.1) and 5.8 (3.3) for the 380 mg extended release naltrexone (VIVITROL) and placebo group, respectively. The pre-baseline median was: 7.7 and 7.6 for the 380 mg extended release naltrexone (VIVITROL) and placebo group, respectively; and the post-baseline median was 3.8 and 5.1 for the 380 mg extended release naltrexone (VIVITROL) and placebo group, respectively [4] . 4. Data on file from Clinical Study ALK2I-003, Alkermes Inc, Waltham MA	Thank you. This information is from the Garbutt 2005 study that we have already included in the review. We added the data sent by the commenter to the relevant analyses.

Comment #	Section	Comment	Response
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Results	Review Key Question 3 and Safety of extended release naltrexone injection in patients with chronic liver conditions The authors state a current limitation/evidence gap on whether naltrexone can be used for people with various liver conditions. The authors also stated that "the FDA removed the black box warning for hepatotoxicity for injectable naltrexone, but it is unclear whether naltrexone should be used in people with various chronic liver conditions." We would like to highlight that there is there is a published analysis of extended release naltrexone injection in opioid dependent patients with the chronic liver condition of hepatitis C and HIV infection. The results indicate that extended release naltrexone injection can be used safely in patients with opioid dependence, even those with underlying chronic hepatitis C and/or HIV infections [2]. The authors may want to consider these data in this review for alcohol treatments. It should also be noted that the monthly naltrexone dose with VIVITROL is one fourth that of daily oral immediate-release naltrexone (380 mg vs. 1,500 mglmonth). Since VIVITROL is an 1M injection, naltrexone released from the intramuscular space avoids first pass hepatic metabolism and is associated with a greater area-under the curve for naltrexone and lower exposure to the principle active metabolite 6-~-naltrexone [5,6] than for 50 mg daily oral naltrexone. 2. Mitchell M, Memisoglu A, Silverman B. Hepatic safety of injectable extended-release naltrexone in patients with chronic hepatitis C and HIV infection. J Stud Alc Drug Dependence. 2012;73 (6):991-997. 5. Dunbar JL, Tumcliff RZ, Dong Q, et al. Single- and multiple-dose pharmacokinetics oflongacting injectable naltrexone. *Alcohol Clin Exp Res. 2006;30:480-490.* 6. TurncliffRZ, Dunbar JL, Dong Q, et a!. Pharmacokinetics oflong-acting naltrexone in subjects with mild to moderate hepatic impairment. *J Clin Pharmacal. 2005 ;45:1259-1267.*	Thank you for the comment. For this review, we are only including studies of people with alcohol use disorders. We are not including studies conducted in other populations (if they don't also have an alcohol use disorder). For many of the included drugs, published studies have evaluated harms in other populations (e.g., healthy subjects, those with depression, etc.), but assessment of those studies is beyond the scope of our review.
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Discussion	p. 87: See comments above on p. 86; edit footnote for dose per month	We fixed this.

Comment #	Section	Comment	Response
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Discussion	p. 89: It should be noted that there are differences in the prescribing information for oral naltrexone and VIVITROL (380 mg per month extended release naltrexone injection), particularly with respect to contraindications. Extended release naltrexone injection is not contraindicated in patients with acute hepatitis or liver failure, or for those with anticipated need for opioids. Please also note that the VIVITROL prescribing information was recently updated in 2013. link: http://www.vivitrol.com/Content/pdf/prescribing_info.pdf	Thank you. We have revised this section to specify the contraindications for oral naltrexone and those for injectable naltrexone separately. We now reference the updated prescribing information.
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Discussion	p. 91: In the VIVITROL pivotal trial for alcohol dependence, significant effects on alcohol consumption were demonstrated vs. placebo (Garbutt 2005, reference 83 in the Review). The primary outcome measure was event rate of heavy drinking. Patients treated with 380 mg extended release naltrexone (VIVITROL) experienced approximately a 25% greater reduction in the rate of heavy drinking relative to the placebo-treated patients (P value 0.03).	Thank you. We realize that this outcome in this study was statistically significant. Our statements here in the discussion are about the overall findings, focused on our assessment of all the relevant data; not just on one study and one outcome.
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Discussion	p. 92: It should be noted that there are differences in the prescribing information for oral naltrexone and VIVITROL (380 mg per month extended release naltrexone injection), particularly with respect to contraindications. Extended release naltrexone injection is not contraindicated in patients with acute hepatitis or liver failure, or for those with anticipated need for opioids. Please also note that the VIVITROL prescribing information was recently updated in 2013. link: http://www.vivitrol.com/Content/pdf/prescribing_info.pdf	We have noted this here and we refer back to the Harms section for details about the contraindications for injectable naltrexone. We now specify the contraindications for oral naltrexone and those for injectable naltrexone separately. We now reference the updated prescribing information.
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Discussion	p. 92: In the pivotal study of VIVITROL in alcohol dependence, those receiving a 380 mg monthly dose of extended release injectable naltrexone (VIVITROL) had greater improvement on the SF-36 mental health summary score than those receiving placebo at 24 weeks (8.2 versus 6.2, P <0.044) [Pettinati 2009, referenced in the Review].	We have already included this finding in the results. This finding for SF-36 mental summary score was just one finding from one study for a single secondary outcome (and just the mental summary score of that secondary outcome).
TEP Reviewer #1	Discussion/ Conclusion	Implications of the major findings are clearly stated.	Thank you

Comment #	Section	Comment	Response
TEP Reviewer #1	Discussion/ Conclusion	Investigators might consider including results for po vs. IM naltrexone in Table C (ES-9)/Table 34 (p87).	We have added these as suggested
TEP Reviewer #1	Discussion/ Conclusion	Limitations are described adequately.	Thank you
TEP Reviewer #1	Discussion/ Conclusion	On p91, investigators comment on effect of IM naltrexone on return to any drinking and return to heavy drinking but not effect on number of drinking days.	This information has been added to Table 34 (and Table C of the ES).
TEP Reviewer #1	Discussion/ Conclusion	The future research section is clear and will be useful and easily translated into new research.	Thank you
Peer Reviewer #1	Discussion/ Conclusion	OK	Thank you
Peer Reviewer #2	Discussion/ Conclusion	1. COMINE Medical Management. I was slightly surprised that the medical management (MM) arms of the COMBINE trial were not further described in the section on primary care. MM was specifically designed to simulate the sort of management a primary care team could provide—monitoring medication adherence and side effects, and encouragement of sober support. Although MM was at more frequent intervals than most primary care clinicians see patients, chronic care management by nurses for diabetes or CHF can occur at such intervals. My own sense is that that COMBINE trial opened the way to consider management of alcohol dependence in primary care, and that the efficacy of the MM-placebo arm in particular deserves mention.13,14 Before COMBINE was published in 2006, NIH and VA grant review sections were hesitant to fund primary care trials for AUD. (i.e. I doubt the Oslin and Saitz studies above would not have been funded.)	We have added description and discussion of COMBINE MM as suggested to both the ES and the full report Discussion: "…Some included studies conducted in non-primary care settings used interventions that may be adaptable for delivery in primary care. For example, in the COMBINE study (Anton et al., JAMA. 2006; 295(17): 2003-2017) providers delivered a medical management intervention comprised of up to nine manual-guided counseling visits. The first visit was approximately 45 minutes and follow-up visits were about 20 minutes each. Medical management included advice for reducing drinking, inquiries about medication side-effects, and emphasis on the importance of taking medications as prescribed."

Comment #	Section	Comment	Response
Peer Reviewer #2	Discussion/ Conclusion	2. The discussion would ideally address DSM-5. This will discussion will benefit from the assessment of severity of AUD for patients in each study (Ideally symptom count if available in many studies), recommended above.	We added the following to the Applicability section of the Discussion: "The included literature used definitions from DSM-III or DSM–IV. DSM-5 (2013) describes a single AUD category measured on a continuum from mild to severe, and no longer has separate categories for alcohol abuse and dependence. Using DSM-5 terminology, most participants in the included studies likely had moderate to severe AUDs. Thus, applicability of our findings to people with mild AUDs is uncertain."
Peer Reviewer #2	Discussion/ Conclusion	3. Gaps in research. For future research (Table E ES) a major gap is our lack of knowledge on whether patients need to stop drinking before starting medications in order for them to benefit. Given that many patients who want help with their drinking are not ready to abstain, this is an important unanswered question.	We agree. We included this point in the Applicability section. We agree that it is an important evidence gap and have added it to the Table as suggested.
TEP Reviewer #3	Discussion/ Conclusion	Page 86, lines 27 and 28: Epidemiological studies suggest that small reductions in drinking can make a difference in medical consequences.	We have expanded the section about health outcomes and evidence from epidemiologic studies as suggested by several reviewers.
TEP Reviewer #3	Discussion/ Conclusion	Page 86, last paragraph: As mentioned above in General Comments, the message is that it works for some but not all, similar to SSRIs for depression.	Yes, that is correct.
TEP Reviewer #3	Discussion/ Conclusion	Page 87, table 34: You might consider including all medications reviewed in this report, especially topiramate since it is a positive finding.	We have separated FDA-approved medications from the off-label medications to keep with the organization throughout the report. We also think that one of the main messages is that naltrexone and acamprosate have the best evidence of efficacy currently, so we first highlight that information, and then mention the findings for the other medications with some evidence of efficacy (topiramate, valproic acid, and nalmefene)
TEP Reviewer #3	Discussion/ Conclusion	Page 88, lines 35-36: Again, Rehm article (see attachment) suggest that small reduction in drinking can made a difference.	We have expanded the section about health outcomes and evidence from epidemiologic studies, including a reference to this article and its main findings.
TEP Reviewer #3	Discussion/ Conclusion	Pages 88-89, Harms section: As mentioned above, acamprosate and naltrexone are relatively safe medications.	Thank you. Rather than describing medications as safe or unsafe, we try to focus on what the evidence is for various types of adverse effects.

Comment #	Section	Comment	Response
TEP Reviewer #3	Discussion/ Conclusion	Page 90, lines 9-11: In O'Malley's study, even Medical Mangement (developed in COMBINE) may be too intense intervention for most primary care settings.	We added a description of COMBINE's medical management here and have included a description of what is involved so that readers can understand the intensity involved.
TEP Reviewer #3	Discussion/ Conclusion	Page 90, Genetic Polymorphisms: As discussed above, personalized medicine will most likely be an integration of genetics, epigenetics, proteomics,physiological characteristics (imaging), patients characteristics and history. Also, see discussion above on genetic variants and ondansetron reponse. Genetics polymorphisms and other factors for personalized response is still not ready for clinicians.	Thank you.
TEP Reviewer #3	Discussion/ Conclusion	Page 91, lines 3-4: As mentioned above, reference 35 is not an NIAAA guideline. We are working with SAMHSA to create a medications development guideline.	Thank you for the clarification. We have revised the text to explain that this is the NIAAA Medical Management Treatment Manual, and have clarified what it covers.
TEP Reviewer #3	Discussion/ Conclusion	Page 91, lines 16-17: Hank Kranzer has two studies although they might have been 2 month treatment trials.	Thank you. We found these studies in our updated literature searches and have reviewed all of Kranzler's potentially relevant studies, and have included the one that is 12 weeks or longer.

Comment #	Section	Comment	Response
TEP Reviewer #3	Discussion/ Conclusion	Page 91, lines 27-39: We really don't know why the European acamprosate trials differ from the US acamprosate studies. Population might have been a consideration. However, when we conducted an analysis of the COMBINE data set, we did not find an effect for acamprosate in the most severe drinking subjects (unpublished data). Most of the European trials were conducted in the 1990's using a similar design. Interestingly, Mann et al (see attachment) (conducted in Germany) used a similar design as COMBINE and found no effect for acamprosate. Perhaps, experimental design had an influence. The early European trials exhibited a somewhat higher external validity and a lower internal validity than the COMBINE and Karl Mann's study. But at the moment, we just don't know for sure.	Thank you. We agree that we don't know why they differ. We include information about this in the Discussion: "Although the majority of included trials assessing the efficacy of acamprosate were conducted in Europe (16 of 22) and a minority were conducted in the United States (4 of 22), the opposite was true for naltrexone (27 of 44 in the United States and 8 of 44 in Europe). Further, the few studies of acamprosate conducted in the United States did not find it to be efficacious. It is unclear whether the different results were due to population differences or other factors. The European trials of acamprosate typically identified patients from inpatient settings or treatment programs, whereas the U.S.-based trials of acamprosate relied on advertisements and referrals. It is possible that this resulted in populations with differing alcoholism severity and differing potential for benefit. For example, studies of subjects recruited via advertisements may enroll people who have less severe disorders, and may be less applicable to patients with more severe forms of alcohol-use disorders."
TEP Reviewer #3	Discussion/ Conclusion	Page 92, lines 36-41: I question the statement that treatment periods of less than 6 months may yield misleading conclusions. First, references 216 and 217 are old references. In all the alcohol pharmacotherapy trials that I have observed, if there is an effect, you see in 6 weeks. Only in one trial did I see the effect disappear and that was in the multi-site Vivitrol trial. Using the intent to treat the effect disappeared after three months. However, in the subpopulation that the FDA used for approval, the effect did not disappeared. We also conducted analysis (not published) on two long pharmacotherapy studies: the Vivitrol study and the John Krystal 12 month VA naltrexone study. Basically, the averages of the outcome (we used % subjects with no heavy drinking outcome, same as FDA requirement) were fairly consisted across the total period of treatment in both data sets. Finally, conducting longer follow-up without treatment doesn't necessarily tell you about the efficacy of the medication. Yes, some might get better but others may need the medication. For example, patients who stop taking SSRI for depression, may relapse to depression without the medication.	In the Limitations of the review section in the Discussion, revised the wording to clarify our meaning based on this comment and comments from another peer reviewer. And we removed the mention of 6 month treatment time. We describe that studies of short duration may yield misleading conclusions about treatment efficacy, due to fluctuations in drinking behavior that are typical of the course of alcoholism. Longer studies are more likely to provide reliable data for the outcomes we were evaluating.

Comment #	Section	Comment	Response
TEP Reviewer #3	Discussion/ Conclusion	Page 93, lines 3-8: One of the challenges in the next decade is to compare levels and frequency of drinking with levels of consequences, both short- and long-term.	Thank you. We have added something similar to this statement to our Evidence Gaps table, in the row for KQ 2: "…These could include large prospective studies to evaluate harm and health consequences with various levels of drinking."
TEP Reviewer #3	Discussion/ Conclusion	Page 94, Conclusions: Interesting the NNT for acamprosate and naltrexone is higher than previous meta-analyses.	Not by much. And there are a lot of new studies in this field. So, it is not really very surprising that it would change some.
TEP Reviewer #3	Discussion/ Conclusion	Page 95, Table 36: #1 Since the effect size is small for alcohol medications acamprosate and naltrexone, it is not surprising that there were no differences. Only if a trial recruited more patients that responded to drug A than drug B would you get a difference.	Thank you. This comment does not seem to suggest any change to the report, but is just reflecting on the reviewer's views.
TEP Reviewer #3	Discussion/ Conclusion	lines 11-15: You might mention that DSM V no longer has dependence and abuse but Alcohol Use Disorder.	We have added this to the report in several places. We deleted the future research need about studies for people with alcohol abuse (since that is no longer an entity with the new DSM-5 definitions).
TEP Reviewer #3	Discussion/ Conclusion	#4: naturalistic studies (what happens in real world) might be more useful. Unfortunately, there is no data for naturalistic or primary care studies.	Those might be useful also, but we stand by the evidence gap and the need listed here, with a focus on primary care (these could also be naturalistic studies).
TEP Reviewer #3	Discussion/ Conclusion	#5 and 6 are really personalized medicine and can be combined.	They are both aspects of personalized medicine, but typically would be different studies, with a different focus.
Peer Reviewer #3	Discussion/ Conclusion	Implications are clearly stated.	Thank you
Peer Reviewer #3	Discussion/ Conclusion	Perhaps because of the length of the document, the final discussion alludes to previous sections-might have been nice to have it there too.	Yes, that is correct, we didn't want to repeat too much information. We're trying to cut down on the redundancy and length. And it increases chances to make errors to have the same information in too many places.
Peer Reviewer #3	Discussion/ Conclusion	Should be noted that placing the of areas for research in a table was really effective.	Thank you
Peer Reviewer #4	Discussion/ Conclusion	This was good	Thank you

Comment #	Section	Comment	Response
Peer Reviewer #5	Discussion/ Conclusion	The discussion, as the rest of the report, is well written underlying the problems physicians phase when prescribing medications for individuals diagnosed with AD. They are also right in pointing out that medications are underutilized and point out to the need for better dissemination of information regarding pharmacological treatment for AD. The authors do a very good job of outlining gaps in the literature and provide suggestions for future research in this area.	Thank you
Peer Reviewer #6	Discussion/ Conclusion	In general the conclusion of this report presents a very conservative estimate of the effectiveness of medications (especially naltrexone) to treat alcohol use disorders. There are few if any novel implications from this report.	We disagree with the assertion that the conclusion is conservative and there are no novel implications from this report. We stand by our assessment on the effectiveness of medications. As we point out in the Introduction, medications for alcohol use are greatly underutilized, with a small percentage of eligible patients receiving medications. It is important to address the basic questions of whether or not the medications work, and which ones work best, and (if they do) to get that message to a large number of providers who might prescribe the medications. There are many opportunities to improve the quality of care for those with alcohol use disorders. A first step is to analyze available evidence to inform practice. We hope that the evidence in this systematic review can be disseminated to decisionmakers who can then incorporate these findings into practice and policies.

Comment #	Section	Comment	Response
Peer Reviewer #6	Discussion/ Conclusion	Of the non-FDA approved medications for the treatment of alcohol use disorders, the authors present a fairly positive assessment for the effectiveness of topiramate. This is based on only 4 studies, only two of which were included in the "meta-analysis". One of the studies that showed a positive effect for topiramate included 54 placebo subjects of which over 50% of the subjects (34) dropped out before completing the 12 week trial compared to 19 subjects in the topiramate group. Since relapse was assumed for dropouts, it is not surprising that topiramate had lower relapse rates. Two other studies were conducted by the same first author. In these studies subjects were not detoxified prior to receiving treatment and most subjects continued to drink heavily throughout the trial. Since topiramate can reduce alcohol withdrawal symptoms it is not clear if topiramate was effective because it reduced withdrawal symptoms rather than acted as a relapse deterrent. Given that subjects in the study continued to drink heavily on about half of the study days, the overall results from the study could be confounded by a reduction in alcohol withdrawal symptoms. While not necessarily a problem it is worth noting that the larger study used in the meta-analysis was sponsored by a pharmaceutical company and so the "low risk of bias" assessment may be a little generous.	We stand by our assessment of topiramate. The 4 studies of topiramate present consistent findings that favor topiramate over placebo. We do understand that it is debatable for whether the strength of evidence should be low or moderate (and we did debate this with our team). We agree that the study mentioned here that had over 50% of the 54 placebo-treated subjects drop out has a high risk of bias, and we had already rated that study as high risk of bias. Having multiple studies from the same first author is not a criteria that we use to critically appraise literature or to alter conclusions. Interesting points about why topiramate was effective, but the fact remains that the outcomes showed efficacy for several alcohol consumption outcomes. Many of the included studies in this review were sponsored by pharmaceutical companies. This was not one of our criteria to assess risk of bias. We assessed potential for publication bias and selective outcome reporting as described in the Methods and Limitations of the Discussion, and we did not find evidence of either type of bias.
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Conclusions	p. 94: In the pivotal study of VIVITROL in alcohol dependence, those receiving a 380 mg monthly dose of extended release injectable naltrexone (VIVITROL) had greater improvement on the SF-36 mental health summary score than those receiving placebo at 24 weeks (8.2 versus 6.2, P <0.044) [Pettinati 2009, referenced in the Review].	We have already included this finding in the results. Thank you for the comment. We reviewed it and we affirm our decision to grade this as insufficient due to unknown consistency, imprecision, and because this finding for SF-36 mental summary score was just one finding from one study for a single secondary outcome (and just the mental summary score of that secondary outcome).
Peer Reviewer #2	Table of Excluded Studies	1. Table of excluded studies. In evaluating studies excluded it was challenging with all the sections separated. A single alphabetical list of studies excluded (First author, Journal and year) and a table indicating why they were excluded would be much more useful.	We revised the organization of the Appendix to provide a single alphabetized list of studies (rather than separate lists for various exclusion reasons) excluded at the full-text level. We provide a reason for exclusion at the end of each citation.

Comment #	Section	Comment	Response
Peer Reviewer #2	Appendixes	The many reference sections would be much more user friendly if they had a HEADER ON EACH PAGE saying what section the references were for. This reviewer spent a considerable amount of time trying to find the references for different sections when I was referring back and forth to the tables and references.	There is 1 set references for the Executive Summary and 1 for the full report. We have retitled the set for the ES to make it more clear what it pertains to ("References for the Executive Summary")
Public Reviewer #1; Alkermes, Inc., Bernard Silverman, maker of Vivitrol	Appendix D	Table D-3; QoL: see comments above in discussion of Key Question 2. Health Outcomes in the Executive summary; also, change footnote to show dose per month not day	We have already included this finding in the results. Thank you for the comment. We reviewed it and we affirm our decision to grade this as insufficient due to unknown consistency, imprecision, and because this finding for SF-36 mental summary score was just one finding from one study for a single secondary outcome (and just the mental summary score of that secondary outcome).
TEP Reviewer #1	Clarity & Usability	The report is mostly well structured and organized but for some points mentioned earlier. The main points are clearly presented. The conclusions can be used to inform policy and/or practice decisions. Implications for policy decisions could be further developed.	Thank you. We have addressed each of the specific comments from the reviewer that this refers back to.
Peer Reviewer #1	Clarity & Usability	Yes	Thank you
TEP Reviewer #3	Clarity & Usability	Authors did an excellent job in conducting this meta analysis (and a lot of work). I made numerous comments to help strengthen the report. I also attached other articles that the authors should consider. Finally, I also attached two articles that NIAAA wrote on the future directions of medications development for alcohol treatment. I agree with the report that there is evidence for improved outcome for acamprosate, naltrexone, and topiramate. I believe that under the right condition, disulfiram can also be useful. Also, results of a recent multi-site trial suggest that varenicline may also be effective in treating alcoholism. We know that these medications work for some, but not all. The challenge over the next decade to determine which patients respond best to a specific medication.	Thank you! The comments were helpful!
Peer Reviewer #3	Clarity & Usability	This document is well structured, with the main points clearly presented and the conclusions could be used by clinicians, educators and for policy decisions.	Thank you
Peer Reviewer #4	Clarity & Usability	Fairly technical in nature	Yes, the full report is fairly technical. We hope that the Executive Summary and a journal article (or 2) will provide more concise, and less technical, versions.

Comment #	Section	Comment	Response
Peer Reviewer #5	Clarity & Usability	As outlined earlier the report is clear and very detailed. Its findings can be used to inform policy and guide practice decisions. However, its length may be a minor deterrent.	Thank you. We hope that the Executive Summary and a journal article will provide more concise, and less technical, versions that will not deter readers because of their length.
Peer Reviewer #6	Clarity & Usability	The sheer volume of tables and statistical data make it difficult to find useful information that can be easily digested by the target audience. As presently offered, this report is not likely to generate significant interest. While the authors have generated a considerable volume of extracted data and gone through their meta-analysis systematically, there is little intellectual creativity in analyzing the data into meaningful clinical observations or guidance for reimbursement policy.	We appreciate that the technical report is long and dense. We hope that the executive summary and a journal article will provide more concise versions of the most useful information.
TEP Reviewer #1	Typos	- pES-14, line 19: Remove period after United States - pES-17: Remove period before citation 19	We fixed these