Comparative Effectiveness Research Review Disposition of Comments Report

Research Review Title: Comparative Effectiveness of Second Generation Antidepressants in the Pharmacologic Treatment of Adult Depression—An Update to a 2007 Report

Draft review available for public comment from December 23, 2010 to January 20, 2011.

Research Review Citation: Gartlehner G, Hansen RA, Morgan LC, Thaler K, Lux LJ, Van Noord M, Mager U, Gaynes BN, Thieda P, Strobelberger M, Lloyd S, Reichenpfader U, Lohr KN. Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression: An Update of the 2007 Comparative Effectiveness Review. Comparative Effectiveness Review No. 48. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center, Contract No. 290-2007-10056-I.) AHRQ Publication No. 12-EHC012-EF. Rockville, MD: Agency for Healthcare Research and Quality. December 2011. Available at: 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 E-mail. 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: http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=727
Published Online: July 2011

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Clarity and Usability	Yes, this is a well-structured and well-organized report. The clarity of the report might be improved at times with greater brevity. The conclusions can be used to inform policy and practice decisions although as noted above the quantity of information provided may limit the dissemination.	We have revised the report and eliminated redundancies.
Peer Reviewer #3	Clarity and Usability	The report is very well structured and organized for the massive amount of data summarized. The main points are presented well. The conclusions are very informative for policy and practice. No overall suggestions.	Thank you.
Peer Reviewer #4	Clarity and Usability	Yes, the authors have done an extremely good job at summarizing vast quantities of data into comprehensible bits.	Thank you.
Peer Reviewer #5	Clarity and Usability	Clarity is good in some places, but obscured by too many words in other places. Usability is greatly hampered by the decisions to include multiple places to summarize the results, and the decision to include so many questions in one report.	We have revised the report and eliminated redundancies. To some, extent, however, we are bound to the structure of MMA reports.
Peer Reviewer #2	Executive Summary	You should go to http://www.nappp.org and see the White Paper: Failure to Serve, that reviews the literature and concludes that "medication only approaches to depression are scientifically unsubstantiated, and are tantamount to a hoax or incompetent treatment". Your work further exposes the severe limitations of anti-depressants in the treatment or understanding of the etiology of depression. Since Seligman's work on conditioning learned helplessness we have had more sophisticated models and have ignored them and made primary care (where most treatment for depression occurs) a "drug deliver system" rather than a healthcare system. Hopefully, the new ACA and Integrated care will stopp this and elevate the quality of care. Still, state departments of mental health are moving quickly to undermine integrated care by passing state Medicaid Plans that lock the mentally ill, like chattel, into mandatory referral to antiquated and drug and case management only systems. Thanks for your amazing work and amazing agency	Thank you for your comments.
Peer Reviewer #4	Executive summary	This detailed review was a pleasure to read, easy to follow despite a large amount of information. The summary tables were understandable and helpful. Pg 23 line 13. It seems unusual to present a reduced incidence of an adverse effect with an NNT. A higher NNH might be more accurate as the way it is presented suggests that bupropion treats sexual dysfunction, which is obviously different than suggesting that it is associated with lower rates than other medications.	We have removed this NNT.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	General	There has been a great deal of interest in whether young adults have an increased rate of suicidal ideation in response to treatment with second generation antidepressants, enough to compel the FDA to require warnings on these antidepressants. The data on which these decisions were based suggests that there may be differences between young adults (late adolescents) and older adults in the way that they respond to second generation antidepressants and in vulnerability to side effects. The reasons for comparing age only at the older end of the age spectrum might be worth articulating as depression is common in young people and fraught with this controversy.	We mention this FDA work in the section on suicidality. Children and adolescents, however, were not included in this report.
Peer Reviewer #4	General	The summary table states that there are no substantial differences between second generation antidepressant medications for either efficacy or effectiveness and yet it appears that select comparisons reveal differences between medications such as escitalopram and citalopram (NNT=13).	There are some statistically significant differences. The magnitudes of the differences, however, are small and likely not clinically relevant.
Peer Reviewer #4	General	It may be difficult for readers to reconcile some of these findings and to extract what is clinically meaningful without some guidance from the authors.	We have tried to provide a concise summaries in the key points and also in the executive summary. The scope of the report, however, is large and it has been a challenge to condense this huge amount of information.
Peer Reviewer #4	General	Many comparisons of interest in this report involve head to head trials of two active compounds. One issue that was rarely discussed was whether the trials were designed to be superiority studies, equivalence trials or non-inferiority trials. This distinction is relevant for understanding whether a trial has achieved its primary goal in comparing two treatments (see for example, Sackett, ACP J Club. 2004 Mar-Apr;140(2):A1 and elsewhere).	Most of the studies used a two-sided p-value but were clearly underpowered to detect a realistic difference between two interventions. Not a single study defined an equivalence or non-inferiority delta and tested accordingly. In meta-analyses, however, the lack of power can be overcome by pooling studies.
Peer Reviewer #4	General	Acknowledgement of this design issue and commentary on whether the studies in fact were testing what they asserted to be testing might be helpful	See above.
Peer Reviewer #1	General Comments	This report represents a tremendous amount of work and provides a clinically meaningful synthesis of a large body of literature. Although the information is quite clinically meaningful and the target population, audience and key questions are explicit, the relevance to clinicians is somewhat limited by the volume of material. Simply put, few clinicians or decision-makers will review document of this length in any depth. To this end, the summary tables are quite useful.	The Eisenberg Center will further condense the material to make it more readable for clinicians.
Peer Reviewer #3	General Comments	The manuscript is of high quality and is carefully researched. It is too long for most readers, but the summaries at the beginning of the MS are very good. Quite comprehensive.	Thank you.
Peer Reviewer #4	General Comments	Yes. It is possible that some questions do not deserve the same 'weight' as others, but all are clinically meaningful.	Thank you.

Source: http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=727
Published Online: July 2011

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	General Comments	This is the second report I have seen with a Table 1 and Table 2 that seem to be covering at least some of the same ground, and are for some reason separated by just a page or 2. I am not sure if this is an AHRQ style issue, or an issue for this EPC, but it does not make sense to me. I found the report to be too broad to digest. It would be better if it could be divided into 2 reports - one way to do this would be drug vs drug and class vs class, and a separate one for comparing long and short-acting drugs. I also found the non-depression symptoms section to be way to large - who decided what symptoms belong on that list?? Hopefully not influenced by drug company marketing efforts to differentiate their drugs!	The scope of the report is large with almost 300 included studies. We divided large tables into smaller ones to make it easier for readers to follow the structure. Accompanying symptoms were selected on clinical grounds and based on results from the STAR-D study
Peer Reviewer #1	Methods	The methods are clearly stated and appropriate. There may be arguments for other endpoints that may be more clinically meaningful but the quantity of literature on such endpoints would be even more limited.	Many thanks.
Peer Reviewer #3	Methods	The methods are very sound and present no significant problems. One might quibble with what constitutes a "low - medium - high" dosage range in the tables, but that is debatable.	Because no research on dosing equivalence was available, we developed this roster to be able to highlight studies with issues in dosing equivalence in a standardized manner.
Peer Reviewer #4	Methods	Yes, although it would have been interesting to see more discussion regarding whether the studies were in fact testing what they claimed to be testing (superiority, non-inferiority or equivalence). It is likely that in many cases the authors of the randomized trials did not set pre-determined criteria regarding what would constitute a clinically significant difference between treatment outcomes and discussion regarding the potential problems with interpretation of such studies would have enriched the discussion.	Many studies appeared to be underpowered. To some extent we could overcome this issue by conducting meta-analyses.
Peer Reviewer #5	Methods	The first page of methods, describing the changes, seems too lengthy and goes into unnecessary detail. I do not understand the change in methods for indirect comparisons. It is stated that previously a network meta-analysis was done but this time a mixed treatment comparison analysis using Bayesian methods. It is my understanding that the term Network meta-analysis is used in place of the longer term mixed treatment comparisons - they are the same thing. A search of the literature seems to support this. Did you mean that last time you conducted adjusted indirect comparisons? I have several other comments regarding wording in the text that is confusing - please see the pdf.	We believe that it is necessary to briefly outline methodological changes that have occurred during the update. We agree that there could be misunderstandings regarding the term “network meta-analysis”. In the original report we used a method termed “network meta-analysis” by Chumley. We have changed the wording to make this clearer.
Peer Reviewer #5	Methods	Pg 32, Line 32 CDER - How is this different to the FDA documents identified by the SRC? This seems like a vague and unclear statement. At the least it is repeated below.	There is no difference. We have deleted this sentence.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Methods	Pg 32, Line 36 "SRC" –Comment: This term has not been introduced before - probably will confuse the reader	"SRC" is introduced in the section on "Topic Development".
Peer Reviewer #5	Methods	Pg 32, Line 37 "We received dossiers from TBD…" – Comment: I think these are known.	Due to delays, they were not known at the time of the submission of the draft report.
Peer Reviewer #5	Methods	Pg 32, Line 39 "The SRC also searched the following sources for potentially relevant unpublished and ongoing Literature…" – Comment: and unpublished data relating to published studies.	We have changed the wording.
Peer Reviewer #5	Methods	Pg 32, Line 50 "…medications outside our…" Comment: replace "outside" with "within"	We have changed the wording.
Peer Reviewer #5	Methods	Pg. 33 line 5 "RCTs of at least 6 weeks duration and an adult study population were eligible for inclusion…" Comment: replace "and" with "in"	We have changed the wording.
Peer Reviewer #5	Methods	Pg 33 line 8 "…we included placebo controlled trials…" Comment: It seems odd to describe it as though you did not include PCTs here, and below that you do include them for the mixed comparison analysis So, you really did include all PCTs regardless of the direct evidence.	We have changed the wording.
Peer Reviewer #5	Methods	Pg 33, line 16 -17 "…outcome of interest." Comment: Requiring a control group - a 2 arm study? Excludes any single-arm observational study.	For harms we also included large uncontrolled studies.
Peer Reviewer #5	Methods	Pg 33, line 19-20 "Outcomes for efficacy or effectiveness,…" Comment: remove "efficacy or"	We have changed the wording.
Peer Reviewer #5	Methods	Pg 33, line 20-21 "If no study measuring health outcomes was available for a particular indication or population subgroup, we included intermediate outcomes (e.g., changes in depression scores)." Comment: This seems to be contradicted by the sentences that come after. If we follow this sentence, you would not include any MADRS or HAMD data if a study reported some SF-36 outcomes. Clearly that is not the case.	We have changed the wording.
Peer Reviewer #5	Methods	Pg. 33, line 25 "(e.g., 50 percent improvement of depression scores for response)." Comment: It is not clear w hat definition you have chosen for response - is it 50%? Or are you accepting whatever rate the study set?	All studies defined response as a 50% improvement of scores.
Peer Reviewer #5	Methods	Pg. 33, line 28 "…and drug interactions." Comment: remove this text; How did you look for these? 'Drug Interactions' is not really an outcome measure.	We have deleted this part of the sentence.
Peer Reviewer #5	Methods	Pg 33, line 30 "…included meta-analyses in this CER…" Comment: delete meta-analyses and insert systematic reviews	We did not include systematic reviews if they did not conduct quantitative analyses.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Methods	Pg 34, line 13-16 Quality Assessment "Items assessed included selection of cases or cohorts and controls, adjustment for confounders, methods of outcomes assessment, length of followup, and statistical analysis." Comment: This paragraph seem to restate a lot of what is stated nicely in the paragraph above. Even if you want to keep parts of this, it should not be separated from the similar statements by the comments on dual review.	The paragraph above refers to RCTs. This section refers to observational studies.
Peer Reviewer #5	Methods	Pg 34, lines 30-38 "Studies that met all criteria were rated good quality. The majority of studies received a quality rating of fair. This category includes studies that presumably fulfilled all quality criteria but did not report their methods to an extent that answered all our questions. Time constraints precluded our contacting study authors for clarification of methodological questions. Thus, the fair quality category includes studies with quite different strengths and weaknesses. We rated studies that had a fatal flaw in one or more categories as poor quality and, generally, excluded them from our analyses. If no other evidence on an outcome of interest was available, however, we may comment on findings from poor studies." Comment: This paragraph seem to restate a lot of what is stated nicely in the paragraph above. Even if you want to keep parts of this, it should not be separated from the similar statements by the comments on dual review.	We have combined the 2 paragraphs.
Peer Reviewer #5	Methods	Pg 34, line 39 "In addition to internal and external validity…" Comment: remove "and external"	We have changed the wording.
Peer Reviewer #5	Methods	Pg 34, line 40 "To evaluate comparative evidence,…" Comment: I am not clear that this is not an issue of applicability - variations in dose as a characteristic. It certainly does not really belong under quality assessment. Also, you have not stated here what you actually did with this information. Did you use it in analyses? Qualitative synthesis - e.g. stratification?	Only few studies had dosages that were not comparable. We highlight such studies in the text.
Peer Reviewer #5	Methods	Pg 34, line 43 "…previously created and use in this CER…" Comment: add a d to use	We fixed the typo.
Peer Reviewer #5	Methods	Pg 35, lines 39-48 Applicability – first paragraph. Comment: I do not think that this paragraph really matches the guidance on this in the EPC guide. Simply saying how you identified effectiveness trials does not describe how you assessed applicability.	We agree, the assessment of applicability does not entirely follow the new guidance for CERs. Because this report was an update, AHRQ agreed that we do not have to go back and create applicability tables for the entire report.
Peer Reviewer #5	Methods	Pg 36 line 34 "…we examined these additional factors…" Comment: replace examined with reported	We have changed the wording.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Results	As noted above, the quantity of detail provided will be reviewed by a very limited audience. The tables help considerably in this area. At some places, the language distinguishing between the effect size and the strength of the evidence was difficult to parse. (For example, when stating that there is strong evidence for some minimal superiority of one medication over another.)	We have revised the report and tried to make such statements clearer.
Peer Reviewer #3	Results	The results are well founded. Basically, there are few differences between SGA drugs in the treatment of depression.	We agree with this summary.
Peer Reviewer #4	Results	It is possible that there are some sections where the reporting of the results was excessively detailed, but this is not a major flaw	In "key points" we have focused on the most important findings.
Peer Reviewer #5	Results	I included all of my comments on results in the pdf. The review team has done a large amount of work here. Some sections are synthesized well (KQ1 and also harms), while others are not quite there yet (KQ 2 and 3)	Many thanks.
Peer Reviewer #5	Results KQ1	Pg 53 or page 32 line 12-22 Finally, as explained in Chapter 2, we graded the strength of evidence for all major comparisons and outcomes in the key points. Table 8 summarizes the key questions that we are addressing in this chapter. We focus in this chapter chiefly on head-to-head studies. If no head-to-head evidence was available, we report on placebo controlled studies. We include information only on studies for which our quality ratings were good or fair; most studies were rated fair, so we specifically call out quality ratings only for good trials or studies. Poor-quality studies are listed in Appendix E; in a very few cases in which a poor-quality study may have had the only relevant information on a major comparison or outcome, we will cite this information in the detailed analysis text. Summary tables in the detailed analyses subsections have only good or fair quality studies. Comment: This seems like a completely unnecessary paragraph - it restates what I just read in methods. Maybe cut this down to 2 sentences.	We agree, this is redundant. We have deleted most of this text.
Peer Reviewer #5	Results KQ1	Pg 55 or page 40 line 5-6 86 articles on meta-analyses or systematic reviews Comment: remove "articles on"	We have removed "article on".
Peer Reviewer #5	Results KQ1	Pg 55 or page 40, line 6-7 We incorporated data from additional placebo-controlled studies for indirect comparisons only. Comment: How many	We have added the respective number.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Results KQ1	Pg 55 or page 40 lines 41-54 The following second-generation antidepressants are currently approved by the US Food and Drug Administration (FDA) for the treatment of depressive disorders in adults: bupropion, citalopram, desvenlafaxine, duloxetine, escitalopram, fluoxetine, mirtazapine, nefazodone, paroxetine, sertraline, trazodone, and venlafaxine. Of these, the following are selective serotonin reuptake inhibitors (SSRI): citalopram, escitalopram, fluoxetine, paroxetine, and sertraline. Selective serotonin and norepinephrine reuptake inhibitor (SSNRI) and serotonin norepinephrine reuptake inhibitor (SNRI) include the following: desvenlafaxine, duloxetine, mirtazapine, venlafaxine. All other second-generation agents include bupropion, nefazodone, and trazodone. The FDA has not approved the SSRI fluvoxamine for treatment of major depressive disorder (MDD) but the Agency for Healthcare Research and Quality (AHRQ) included it on the list of medications of interest for this review. Comment: This paragraph seems unnecessary here - it should be an overview of the results, not a general overview.	We have removed this paragraph.
Peer Reviewer #5	Results KQ1	Page 56 or page 42, line 3-9 Tables 10 through 14 provide selected information on all these studies. They are grouped according to the main drug classes compared—SSRIs vs. SSRIs (Table 10); SSRIs vs. SSNRI and SNRI (Table 11); and SSRI vs. other second-generation antidepressants (Table 12); SNRIs vs. SSNRIs and SNRIs (Table 13); SNRIs vs. other second-generation antidepressants (Table 14) and other second-generation antidepressants vs. other second-generation antidepressants (Table 15)—and then listed alphabetically by the specific drugs compared. Comment: I personally don't need this level of detailed description - I can read it for myself and would not get confused without this. I would rather get right into the results without so much preamble.	The reasoning for this level of detail is that we assume that most readers will not read the report cover to cover but rather pick individual chapters. With providing more detail at the outset of each chapter, we hope to make it easier for such readers to be able to follow the report.
Peer Reviewer #5	Results KQ1	Page 56 or page 42, line 30-31 Investigators rarely assessed quality of life and functional capacity; if they did, they typically considered these as only secondary outcomes. Comments: Study investigators?	We have changed the wording to “study investigators”.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Results KQ1	Page 56 or page 42, line 31-34 Most studies employed both physician-rated scales (e.g., HAM-D, MADRS, Clinical Global Impressions Scale [CGI]) and patient-rated scales (e.g., Hospital Anxiety and Depression Rating Scale [HAD-A], Battelle Quality of Life Scale [BQOLS]). All studies used physician-rated scales to assess the main outcome measures. Comments: These 2 sentences do not mesh. Please be clear! I am not clear on the point of both sentences. Maybe what is important is that there are more physician rated svales reported??	We have reworded this section.
Peer Reviewer #5	Results KQ1	Page 63 or page 49, line 37-57 Direct evidence was considered sufficient to conduct meta-analyses for six drug-drug comparisons: • Citalopram vs. escitalopram [and all other bullets] Comment: For all of these bullets, the order of results is reversed from what you stated your preference was in the methods section. You made a case for preferring response rates - those should be presented first. In fact, for key points I don't see why you would present the rest.	We have streamlined the key points and focus now on response rates.
Peer Reviewer #5	Results KQ1	Page 63 or page 49, line 47-50 Escitalopram is still patent protected, whereas citalopram is also available as a generic drug. Above mentioned results are based on meta-analyses of head to- head trials. Comment: remove this sentence	We have removed this sentence.
Peer Reviewer #5	Results KQ1	Page 64 or page 50, line 4-5 Table 16. Number of head-to-head trials of selective serotonin reuptake inhibitors for treating major depressive disorders Comment: I do not think this table is useful at all. It repeats what we saw in the multiple tables above. It is also strange to have it inserted in the middle of the bullets.	These tables were in the wrong place. We have moved them to the overview section where the text refers to them.
Peer Reviewer #5	Results KQ1	Page 66 or page 52, line 3-9 Very few comparative effectiveness trials were available; their findings were generally consistent with those from efficacy trials. Eighteen studies (N = 4,050) comparing one second-generation antidepressant with another indicated no differences in health-related quality of life. Quality of life, however, was rarely assessed as a primary outcome measure. Comment: It is strange to present this information here - it should have come first as you clearly stated these were your primary outcome measures and study designs. Also, its confusing hen you switch back and forth between bullets and paragraphs without some sort of signal to the reader about what has changed.	We have deleted this sentence from the key points. We use bullet points just for the listing of meta-analyses.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Results KQ1	Page 66 or page 52, line 11-12 Seven studies, all funded by the maker of mirtazapine, reported that mirtazapine has a significantly faster onset of action than citalopram, fluoxetine, paroxetine, and sertraline (Table 18). Comment: I don't remember that this was an outcome you discussed in the methods section. How do you decide what is a meaningful difference? And insert this text before the work significantly "statistically (not necessarily clinically)"	Onset of action is an outcome of interest for the report. We added "statistically" to the sentence to be clear that we do not mean clinically significant.
Peer Reviewer #5	Results KQ1	Page 87 or page 72, line 33 We identified no head-to-head trials in a population with dysthymia. Comment: remove this sentence	In this case we disagree with the Peer Reviewer. We think that it is important to point out that we did not find any head-to-head evidence.
Peer Reviewer #5	Results KQ1	Page 87 or page 72, line 33 The significant differences in population characteristics in placebo-controlled trials Comment: remove significant insert meaningful	We have changed the sentence.
Peer Reviewer #5	Results KQ1	Page 87 or page 72, line 42-43 Two studies provide mixed evidence about the general efficacy of fluoxetine for the treatment of dysthymia. Comment: I think mixed evidence is a confusing term	We have changed the wording to "conflicting".
Peer Reviewer #5	Results KQ1	Page 87 or page 72, line 45-49 A subgroup of patients older than 60 years showed a significantly greater improvement than those on placebo; a subgroup of patients younger than 60 years did not show any difference in effectiveness between paroxetine and placebo. Comment: Shouldn't this go in the section on subgroups??	We talk about this study in more detail in KQ5. We have removed this statement from the key points.
Peer Reviewer #5	Results KQ1	Page 90 or page 76, line 24 Two of these trials compared fluoxetine daily with fluoxetine weekly Comment: Not clear if these are the same formulation (immediate release) or if the weekly one is an extended release formulation.	It is the same formulation. This is explained in the introductory paragraph of this chapter.
Peer Reviewer #5	Results KQ1	Page 90 or page 76, line 26-29 We could not find any studies on other medications, such as bupropion or fluvoxamine, that are available as both immediate and extended release formulations. Comment: delete sentence	We would like to retain this statement because we think it provides important information for readers.
Peer Reviewer #5	Results KQ1	Page 91 or page 77, line 23-25 Efficacy of Immediate - Versus Extended-Release Formulations: Key Points Comment: I think these key points needs qualifiers in relation to study quality and the overall SOE.	We have added the SOE rating to the key points.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Results KQ1	Page 91 or page 77, line 29-31 Two RCTs reported similar rates of maintenance of response and relapse for patients treated with fluoxetine daily or fluoxetine weekly during the continuation phase. Comment: Did you explain why this is even possible with fluoxetine somewhere in the earlier discussion of the development of this question? I don't remember it. Might want to say that this is probably only possible with this specific drug. Also, in this specific case the suggestion is to treat initially with daily dosing and then switch to weekly. Is there a mechanistic reason for this?	This is explained in the introductory paragraph of this chapter. The reason is the long half-life of fluoxetine.
Peer Reviewer #5	Results KQ1	Page 91 or page 77, line 29-31 After 12 weeks of treatment, significantly more patients on venlafaxine XR experienced a response to treatment than patients treated with venlafaxine IR (data not reported; Comment: Yikes! I would not think that this deserves such a strong statement in the key points then. You don;t know how much of a difference there was in absolute terms. I assume you rated this fair quality, but the SOE has to be low or even insufficient?	The data was reported as a graph only. We could not deduct exact differences between the two formulations. We agree with downgrading the SOE to low.
Peer Reviewer #5	Results KQ4	Page 123 or page 108, line 44 Adverse Events and Discontinuation Rates: Key Points Comments: This section on key points is very well written.	Thank you
Peer Reviewer #5	Results KQ4	Page 133 or page 118, line 5 Except for sexual dysfunction Comment: may have missed the discussion on this in the methods, it's hard for me to see how sexual side effects are at the same level of seriousness as suicidality. It is certainly not life threatening.	We adopted the definition of the FDA which states that "a serious adverse event is any medical occurrence that results in death, is life threatening, requires hospitalization, results in persistent or significant disability or incapacity, or is a congenital birth defect". We view sexual dysfunction as a significant incapacity.
Peer Reviewer #5	Results KQ4	Page 133 or page 188, line Except for sexual dysfunction, trials and observational studies were too small and study durations too short to assess the comparative risks of rare but serious adverse events such as suicidality, seizures, cardiovascular adverse events, serotonin syndrome, hyponatremia, or hepatotoxicity. Comment: This sentence is confusing.	We have reworded this sentence.
Peer Reviewer #5	Results KQ4	Page 142 or page 127, line 13 Adherence: Key Points Comment: I am surprised to not also see persistence here, as it may be the more relevant outcome - using observational evidence. Adherence in trials is somewhat to quite artificial - we would expect that if a specific drug results in lower adherence rates then they would have worse outcomes. So, its an intermediate outcome with unclear meaning	We appreciate this comment. We have used adherence as an overall term for both, adherence and persistence. Any study assessing persistence was eligible for inclusion. We have revised the report accordingly and distinguish now between adherence and persistence.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Discussion	The discussion is clear and coherent. I have specific suggestions for consideration: Section citalopram versus escitalopram (page 17, manuscript page 3): Although a difference between drugs is noted as significant, a more important question is whether it is clinically meaningful; a mean difference on the MADRS of 1.52 points is not meaningfully different. Note that no NNT is given.	Thank you. We try to emphasize in the discussion that statistically significant differences are likely not clinically relevant.
Peer Reviewer #3	Discussion	Paroxetine vs. duloxetine (page 18): It is unclear why the studies were too heterogenous to pool changes in HAM-D but not response rates. Same for the sertraline versus venlafaxine data.	Statistical heterogeneity also depends on the outcome measure of choice. Relative outcome measures (e.g. ORs) are usually more stable than changes on scores.
Peer Reviewer #3	Discussion	Speed of response (page 18): It should be noted that the alleged speed of response with mirtazapine may be related to differential effects on sleep alone. This is reflected in the differential adverse effect profile of somnolence with mirtazapine.	We have added a sentence to the discussion.
Peer Reviewer #3	Discussion	Anxiety (page 20): Drug-placebo differences on anxious symptoms with bupropion are minimal or, perhaps, non-existent (q.v., Tomarken et al. J Affect Disord 2004; 78:235-241). In the Trivedi et al. study of anxiety symptoms in MDD (J Clin Psychiatry 2001; 62:776-781), neither bupropion nor sertraline meaningfully separated from placebo on either HAM-D or HAM-A (although this study is commonly touted as indicating that bupropion is equally effective to sertraline on anxious symptoms). Another paper commonly cited is Rush AJ, et al. Neuropsychopharmacology 2001; 25:131-138. However, that was not a placebo controlled trial.	The Trivedi study is a pooled analysis, and was therefore ineligible for the efficacy section of this report. The Rush trial is included in this section, it was a head-to-head trial of bupropion and sertraline. Tomarken et al. did not meet our eligibility criteria because it is a placebo controlled trial with only 19 patients. Overall, we agree with these comments and the conclusion in this section state that bupropion and sertraline are equally efficacious for anxiety and that no difference was seen between bupropion and placebo.
Peer Reviewer #1	Discussion/ Conclusion	The conclusions are clearly articulated and in many ways this review hits the nail on the head when identifying limitations of the available body of literature. These conclusions suggest a clear agenda for future research. However, given the large number of key questions, it may be helpful to further prioritize among the unanswered questions.	At the end of the discussion we tried to prioritize future research needs
Peer Reviewer #3	Discussion/ Conclusion	Pain (page 21): It should be noted that all four referenced drug-placebo studies in the table used duloxetine 60 mg./day, which was marginally better in pain and fibromyalgia trials in contrast to 120 mg./day. Also, the following reference is not in the paper: reference Perahia DG et al. J Psychiatry Res 2009; 43:512-518.	Yes, the doses of the placebo-controlled trials are listed in the table as 60mg/day. No trials of 120mg/day are included. The publication Perahia DG et al. J Psychiatry Res 2009; 43:512-518 was excluded because it does not include a control group, rather compares two different methods of switching to duloxetine in non- or inadequate responders.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Discussion/ Conclusion	Sexual functioning: You should at least comment on the pooled analysis of duloxetine trials by Delgado PL et al. J Clin Psychiatry 2005; 66:686-692.	We included this study in KQ4 on adverse events
Peer Reviewer #3	Introduction	Table 5. Note extra space in brand name Remeron.	We have removed the extra space.
Peer Reviewer #3	Introduction	Dosing ranges (page 49): Recommend following changes: Desvenlafaxine Pristiq® 50-100 mg. Duloxetine Cymbalta® 30-60 mg	The dosing classification method used in this report follows a structured process that begins with doses included in the FDA-approved product labeling. We then carefully reviewed these dosing ranges for gross inconsistencies with clinical practice. This method previously demonstrated a dose-response relationship for the high-medium-low classification (Hansen et al., Medical Decision Making, 2009). We acknowledge that our dosing classification method might not perfectly reflect all dosing practices, but we believe they are able to identify gross inequities across studies and the method is appropriate for using in systematic review.
Peer Reviewer #3	Introduction	For low-medium-high doses for venlafaxine, recommend using the same ranges for both immediate and extended release products (since kinetics don't differ much other than Tmax).	The dosing classification method used in this report follows a structured process that begins with doses included in the FDA-approved product labeling. We then carefully reviewed these dosing ranges for gross inconsistencies with clinical practice. This method previously demonstrated a dose-response relationship for the high-medium-low classification (Hansen et al., Medical Decision Making, 2009). We acknowledge that our dosing classification method might not perfectly reflect all dosing practices, but we believe they are able to identify gross inequities across studies and the method is appropriate for using in systematic review. The immediate and extended release dosing ranges follow this approach.
Peer Reviewer #3	Discussion/ Conclusion	Suggest arranging forest plots by date and not alphabetically.	We believe that both ways are equally fine. Since we did not conduct any cumulative metaanalyses, rearranging forest plots is probably not necessary.
Peer Reviewer #3	Discussion/ Conclusion	Fluoxamine misspelled in Figure 14.	We have fixed the typo.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Discussion/ Conclusions	I found this section to be too long - repeating much of what was stated multiple times already in overviews and then in key points, and now here in both text and table. So, that is 4 times, not counting the executive summary! I suggest cutting it down to a discussion of what the findings mean. So, it would be just a few paragraphs.	We have streamlined the discussion but we are bound to the format of AHRQ reports.
Peer Reviewer #5	Discussion/ Conclusions	Page 163 or page 148, line 30 No evidence Comment: I am under the impression that no evidence results in an insufficient rating.	Many thanks for pointing this out. We have changed the ratings accordingly.
Peer Reviewer #5	Discussion/ Conclusions	Page 169 or page 154, line 28 Results for Efficacy and Effectiveness in Major Depressive Disorders Comment: Results [This comment refers to the Discussion section]	We have streamlined the discussion but we are bound to the format of AHRQ reports.
Peer Reviewer #5	Discussion/ Conclusions	Page 175 or page 160, line 27 Applicability of Results Comment: I think that the EPC guidance on how to approach this section suggests more details using a PICO format.	Because it is an update of an older report, this section does not entirely follow the current guidance.
Peer Reviewer #5	Evidence Table KQ5	Page 502, line 12 Comment: Is this overall Œ not by D1 vs D2?	The comment in the Evidence Table has been removed.
Public Reviewer #1	Executive Summary	You should go to http://www.nappp.org and see the White Paper: Failure to Serve, that reviews the literature and concludes that "medication only approaches to depression are scientifically unsubstantiated, and are tantamount to a hoax or incompetent treatment". Your work further exposes the severe limitations of anti-depressants in the treatment or understanding of the etiology of depression. Since Seligman's work on conditioning learned helplessness we have had more sophistocated models and have ignored them and made primary care (where most treatment for depression occurs) a "drug deliver system" rather than a healthcare system. Hopefully, the new ACA and Integrated care will stop this and elevate the quality of care. Still, state departments of mental health are moving quickly to undermine integrated care by passing state Medicaid Plans that lock the mentally ill, like chattel, into manditory referral to antiquated and drug and case management only systems. Thanks for your amazing work and amazing agency! Jerry Morris, PsyD, MS(Pharm), MBA, ABMP, ABPP, CCM, NCSP; Board Certified Medical Psychologist	Many thanks for this interesting link.