Comparative Effectiveness Review Disposition of Comments Report

Research Review Title: Adverse Effects of Pharmacologic Treatments of Major Depression in Older Adults

Draft review available for public comment from May 13, 2018 to June 13, 2018.

Research review citation: Sobieraj DM, Baker WL, Martinez BK, Hernandez AV, Coleman CI, Ross JS, Berg KM, Steffens DC. Adverse Effects of Pharmacologic Treatments of Major Depression in Older Adults. Comparative Effectiveness Review No. 215. (Prepared by the University of Connecticut Evidence-based Practice Center under Contract No. 290-2015-00012-I.) AHRQ Publication No. 19-EHC011-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2019. Posted final reports are located on the Effective Health Care Program search page. DOI: https://doi.org/10.23970/AHRQEPCCER215.

Comments to Research Review

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

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the 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.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #1	Executive Summary	Pages 14-16. Table B. The format of Table B makes the findings difficult to interpret. Perhaps, it could benefit from instructions as to how to interpret the contents. For example, on page 14, lines 4 and 20, does "less" mean fewer adverse events or does it mean less severe? If it is describing a number, fewer would be correct.	We have changed the word "less" to "fewer" since this is more accurate. We have added more text prior to the table explaining interpretation of table information. We also included a legend for SOE interpretation in the text preceding the table rather than just as a footnote in Table B.
Peer reviewer #1	Executive Summary	Page 17, line 7. Consider adding "not-placebo-controlled" in parentheses.	This was added.
Peer reviewer #1	Executive Summary	Page 17, lines 46-54. Discussion. Were any of the included studies specifically inclusive of nursing facility residents and if so, likely a Patient Health Questionnaire (PHQ-9) was performed rather than a HAM-D or MADRS. If all nursing facility studies were excluded, please indicate why. Were all of the residents deemed too unstable or subject to polypharmacy? Given that subjects with dementia and nursing facility residents were populations of interest, it would be helpful to know if any were included and what were the specific findings in these populations.	The nursing facility setting was not a reason for exclusion in this review. None of the included studies were specifically a nursing facility population. This statement was added to the ES and report discussion. Furthermore, the full report PICOTS "S" states "We were interested in non-acute care settings for KQs and CQ such as specialist or generalist outpatient setting, rehabilitation facility and nursing facilities". No studies used PHQ-9 for depression screening/diagnosis and no exclusions were made based on this criterion.
Peer reviewer #1	Executive Summary	Page 18, lines 8,9. The sentence indicates that only outpatients were included (i.e., not inpatient or urgent care settings). Are the authors thus excluding nursing facility residents? The terminology is confusing given that nursing facility residents are "institutionalized". Please clarify.	No, nursing facility residents were not excluded. We added the word "hospitalized" to inpatient as that is what is intended. Furthermore, the full report PICOTS "S" states "We were interested in non-acute care settings for KQs and CQ such as specialist or generalist outpatient setting, rehabilitation facility and nursing facilities".

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #1	Executive Summary	Page 18, lines 53, 54. Is this sentence suggesting that you advocate for more observational studies? Why would observational studies be better than controlled studies for assessing adverse effects in older adults?	” Although observational studies are considered more “real-world” in terms of applicability and may better capture harms in particular than the “protocolized” nature of RCTs, we have edited the statement to state “…well controlled studies powered to assess adverse events..”.
Peer reviewer #3	Executive Summary	Executive summary: Page ES-1, line 22: The description of the Beers criteria focus is a little off – more accurate might be to say “…identifies potentially inappropriate medications that are best avoided for most adults with specific conditions, or…”	We have made this edit.
Peer reviewer #3	Executive Summary	Page ES-1, line 27-28, and Page 4, line 21: This suggestion for alternative meds is not a formal AGS recommendation; rather, it was a workgroup whose work was reviewed by AGS and the Beers panel.	We have edited the language to reflect the work group rather than AGS makes this recommendation.
Peer reviewer #3	Executive Summary	Page ES-7, table C, footnote a: This footnote says “This cohort study” – but I thought there were 2 observational studies.	There are two observational studies that met inclusion criteria. Given the new table format, both observational studies are referenced appropriately.
Peer reviewer #3	Executive Summary	Page ES-9, and full report, page 40: The references that begins with “Steinmen” are misspelled (it is Steinman)	We have corrected this.
Peer reviewer #5	Executive Summary	The Structured Abstract summarizes the main aspects of the project aptly, and the Executive Summary is a useful overview providing a greater amount of detail. Although I recognize that Table B contains nearly the full array of information resulting from the analyses, I find it rather difficult to understand and not very useful.	We have made several revisions to the tables based on all comments received to improve the utility of the table.
Peer reviewer #5	Executive Summary	On p. ES-8, in the last sentence, the clause "particularly those of observational design and studies powered to assess adverse events" does not read right and is confusing. What does "those" refer back to?	“those” refers to “studies”, which has been corrected.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Executive Summary	Table A (p. ES-1) may be somewhat misleading in terms of several of the included medications. Although part of the SSRI group of medications, fluvoxamine is almost always used in individuals with obsessive compulsive disorder and would not be used as a first-line antidepressant medication. Also, trazodone is often prescribed in low doses to assist with insomnia but is almost never used as an antidepressant at doses that are effective in treating depression. Data on harms of trazodone may still be relevant to clinicians but it is not likely to be used as a first-line treatment for MDD.	There was considerable discussion during the protocol development as to which therapies to include as interventions and what term should be used to label them (i.e. first-line). The included drug therapies were determined in consideration of practice guidelines but the major guidelines were published prior to several newer antidepressants coming to the market. Thus, the input of the KI and TEP panel was solicited to determine which therapies would be of most interest to clinicians reading this report. This clarification has been added to the "Intervention" of the PICOTS section of the introduction. We have removed the word "first-line" as it can be confusing and isn't important to further classify the antidepressants beyond the list we provide and how they were selected for inclusion.
Public Reviewer #1, American Psychiatric Association	Executive Summary	On p. ES-1, the second paragraph notes that "Approximately 15-20 percent of adults older than age 65 in the United States have experienced depression." This statement references the CDC website (CDC Promotes Public Health Approach to Address Depression among Older Adults. https://www.cdc.gov/aging/pdf/cib_mental_health.pdf. Accessed Nov 2, 2017) which in turn references a Geriatric Mental Health Foundation site that is no longer available (www.gmhfonline.org/gmhf/consumer/factsheets/depression_latelife.html). Thus, the actual study that provided this statistic is unclear. The text as written does not allow one to determine whether the 15-20% rate is a lifetime prevalence rate or whether it refers to depressive episodes that have occurred after age 65. If more specific information is available, it would be useful to cite in lieu of the current citation.	According to the source, this statistic reflects the estimated prevalence. We have changed the sentence in the ES and main report to clarify this.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Executive Summary	On p. ES-2, in the analytic framework (Figure A), and throughout the rest of the document, it would be preferable to use "5 or more concurrent prescription medications" rather than the word "polypharmacy". The latter concept is used in multiple contexts with differing definitions and sometimes is used to indicate prescribing of multiple medications in the same class or of a general type (e.g., psychotropics) or inappropriate use of multiple medications.	We recognize the term "polypharmacy" may have more than one interpretation and for this reason we provide a definition of the term "5 or more concurrent prescription medications" as it is used in this report.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Executive Summary	On p. ES-2, in the paragraph on results, the text notes that the majority of studies relied on spontaneous reporting of adverse events and the following page notes that "none of the RCTs were powered or designed to capture adverse events." It is not clear why that aspect of the study design wouldn't have had a greater (negative) effect on the risk of bias estimates for the individual studies yet 13 were listed as low risk of bias. In addition, in the tables in the appendix, a number of these studies are rated as having a high risk of bias on selective outcome reporting or other factors. Still others were rated as "unclear" on sequence generation, allocation concealment and blinding of outcome assessors. These issues lead to questions about the ratings of low risk of bias for the individual studies and, in turn, lead to questions about the appropriateness of the strength of evidence ratings.	The Cochrane Risk of Bias tool was used to assess the risk of bias for each included RCT. Each domain was rated as "high", "low" or "unclear" risk of bias. According to AHRQ EPC Methods, a "summative" risk of bias for each study is recommended. We classified the overall risk of bias as "low", "medium", "high" according to the collective domain assessment for a given trial. If the majority of domains were "unclear" the summative risk of bias of "unclear" was assigned. If the two investigators assessing summative ROB felt that the domains rated "high" for a given study collectively were sufficient to consider the full study results biased, the trial was given an overall rating of "high". When evaluating the domain of "study ROB" in the SOE grading, the totality of studies for that given intervention/comparator/outcome was considered. If several studies contributed data but had varying risk of bias ratings, we considered where the majority of data came from using the sample sizes and assigned the SOE rating accordingly. Thus if the majority of sample came from a low risk of bias trial, the ROB domain in the SOE rating would have been rated as low. Being underpowered would contribute to imprecision and would be considered when rating the domain of "precision" in SOE. Suspicion of selective outcome reporting was more clearly documented in SOE grading and has contributed to the lowering of SOE for several outcomes, as it was a common finding.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Executive Summary	On p. ES-4, the layout of Table B parallels the layout of the systematic review. However, this makes it difficult to use the information in determining whether a given medication is more or less likely to be associated with adverse effects than another medication. The review states that the group-wise comparisons (e.g., SSRI vs. SNRI) were originally intended to identify medication class effects, but that this was difficult due to the paucity of studies. Given this, it may be more helpful to readers to replace or supplement this table with one organized by medication. Also, for the comparisons relative to placebo, please add "vs. placebo" to the relevant comparisons in the acute, continuation, and maintenance columns to improve the ability to read these columns from top to bottom without needing to look at the far-left column. Similarly, add the appropriate comparisons in these columns for the mirtazapine vs. paroxetine comparisons and the vortioxetine vs. duloxetine comparisons. The use of + to +++ for the strength of evidence ratings in the table is initially confusing before reading the legend, as a reader may erroneously infer that this represents the magnitude of the difference rather than the strength of evidence rating. Because the frequency and severity of an adverse effect are crucial pieces of information for clinicians, it would be important to incorporate this information into the report throughout, particularly in summary tables.	The a priori determined unit of analysis was drug class. Data are not analyzed on the individual drug basis in this report. In some instances, where only a single drug was represented in a class, we provide that drug name so the reader can understand the limitation brought forth with just one drug representing a class. In addition to the footnote explanation for SOE, we have added text preceding the table as to the interpretation of the symbols for SOE. We have added (SOE) in the column headers as well. Frequency of individual adverse events on the study level are reported in Appendix C Table 3. Severity information, particularly for the outcomes of "any adverse event" and "withdrawal due to adverse events" are provided in the text when the studies provided this detail, which was infrequent.
Public Reviewer #1, American Psychiatric Association	Executive Summary	On p. ES-8, a crucial statement is buried in the middle of the discussion. It may be worth highlighting the statement that "Readers should not assume a failure to find a difference means the given interventions are similar in adverse event profiles." Or, add this as a footnote to the summary tables.	We have moved that limitations paragraph to first in that section to better highlight this statement. We have clarified that this applies to conclusions with low SOE or the findings for which SOE were not graded.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Executive Summary	On p. ES-9, the main conclusion of the review is not very helpful to clinicians. It is impossible to know how to apply the information that "acute treatment with SNRIs (duloxetine and venlafaxine), but not SSRIs (escitalopram and fluoxetine), led to a greater number of adverse events compared with placebo" unless data is provided on the actual type and relative frequencies of those adverse events. It would also be important to know if those differences are clinically significant or just statistically significant.	. The NNH for graded outcomes with statistical significance has been added in the key messages, abstract, and key points to improve translation of results. When the data concerning specific harms contributing to these more general outcomes were reported in the studies, we report this detail in the report text. The number of events as reported per study are provided in Appendix C Table 3.
Public Reviewer #1, American Psychiatric Association	Executive Summary	p. ES-1 Heading "Rational" should be "Rationale"	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Executive Summary	p. ES-1 3rd paragraph "American Geriatric Society" should be "American Geriatrics Society"	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Executive Summary	p. ES-2 1st paragraph "PsychInfo" should be "PsycInfo", here and throughout the rest of the document	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Executive Summary	p. ES-3 Methods "SIADH "should be written out	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Executive Summary	p. ES-4 Table B "Less" should be "Fewer" (adverse events and study withdrawals are countable items)	This has been corrected.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Executive Summary	p. ES-8 1st paragraph, 2nd to last line comma after "criteria"	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Executive Summary	last paragraph decision-makers is spelled wrong (and in several other places in document)	According to the AHRQ publication guide, descionmakers is left without a hyphen.
Peer reviewer #3	Abstract	Abstract: Page viii, line 14, and full report page 35, line 7: This makes it sound like vilazodone, trazodone, and vortioxetine are considered first-line treatments. The authors clarify later in the report that these are drugs that clinicians in daily practice often choose first, rather than being guideline-recommended as first-line, but here and in a few other places it sounds like these are guideline-recommended as first-line choices.	Clarifying language as to how these drugs were selected for inclusion is presented in the background section of the ES preceding Table A. We recognized that the same language was not used in the PICOTS section of the main report and have now added the same language in that section. We have also added similar language to the abstract. We have removed the word "first-line" from the title and throughout as it may be confusing. We have listed the included drugs and also how they were decided upon.
Peer reviewer #3	Abstract	Page viii, line 47: Here and elsewhere in the report, I found the phrasing "we identified single RCTs" to be confusing – would consider wordsmithing or otherwise clarifying.	We changed the language to "..for a given comparator and outcome, conclusions are often based on a signle study".
Peer reviewer #6	Abstract	Including results focused on withdrawal due to adverse events in the abstract and conclusions would provide more balance.	We present results for withdrawal due to adverse events in the abstract and conclusions.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #6	Abstract	Given that the large observational study was considered low SOE, it is not clear these results should be prominent in the abstract. Although the sample for the observational study was limited to those with a depression diagnosis (not necessarily MDD), it is likely that those who did not receive treatment had less severe depression --which can not be completely adjusted for based on the severity categories that were used in this study. This could explain some of the findings (which are related to depression severity). It is not clear that the authors of the observational study could fully control for channeling that can occur (preferential selection of antidepressants based on patient characteristics) given information from administrative data. Furthermore, it is not clear how this observational study defined "adverse events" from the administrative data. This observational study provides information on important outcomes that were not included in RCTs (albeit low), however, it is not clear whether these results should be positioned as prominently in the report.	We recognize that the observational study itself has limitations associated with its internal and external validity which have been considered in the grading of strength of evidence for data used from this study. However, observational data are particularly important for harms outcomes and can give insight into effects associated with treatments particularly where RCT evidence is scare, as in this report. We have added more detail to the limitations of this study as per this suggestion and the suggestions of other reviewers.
Peer reviewer #1	Introduction	The Introduction is appropriate	Thank you.
Peer reviewer #1	Introduction	Page 20, line 49. Given that several trials are the subject of the sentence, I think “analysis” should be “analyses”. Consider changing “analyzed” to “evaluated”.`	We have made these corrections.
Peer reviewer #1	Introduction	Page 22, line 40. The American Geriatrics Society and the American Medical Directors Association specifically discourage use of the term, “nursing home” and prefer “nursing facility”. Please consider replacing throughout the report	We have replaced “nursing home” with “nursing facility” throughout the report.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #1	Introduction	Page 22, lines 31 – 38. There is a mention of three trials within which there was no association between baseline cognitive function and depression outcomes, but there are no citations. In a metaanalysis, the group out of Sunnybrook Health Sciences Centre in Toronto, found that antidepressant treatment was efficacious and that cognitive status did not change. Thompson S, Herrmann N, Rapoport MJ, Lanctot K. Efficacy ad safety of antidepressants for treatment of depression in Alzheimer's disease: a metaanalysis. Can J Psych 2007;52:248-55. Similarly, Weintraub reported that older adults who responded to sertraline at 12-weeks post initiation did not sustain that improvement at week 24, although cognition did not decline. Weintraub, D.; Rosenberg, P.B.; Drye, L.T.*; et al.* Sertraline for the treatment of depression in Alzheimer disease: week-24 outcomes. *Am. J. Geriatr. Psychiatry* 2010;18:332-40. Neither of these papers were included or excluded in the AHRQ report. Please evaluate.	This paragraph discusses the findings of a systematic review and meta-analysis by Benraad and colleagues, which is referenced. The sentence in question refers to 3 of the trials included in that analysis that did analyze relationship of cognitive function and depression outcomes. Of note, the contextual question methods, which are reported in the "method" chapter, state that "This question is not based on a systematic review and strict screening and inclusion criteria were not applied as the aim of the CQ is to provide a qualitative overview of the state of the evidence without formal systematic review or analytic plans." The Thompson et al citation provided in this peer review comment did not indicate any information regarding age of the subjects and thus was not included in the CQ narrative. Weintraub et al was a RCT that did not meet inclusion for KQs1 or 2 because of age and major depression diagnosis criteria. These patients were not required to be 65y of age and older and were diagnosed with depression of Alzheimer's disease which is a different diagnosis than major depressive disorder.
Peer reviewer #3	Introduction	No comments	NA
Peer review #4	Introduction	Good overview and well written.	Thank you.
Peer reviewer #5	Introduction	The Introduction reads OK, in general.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #5	Introduction	P. 1, first sentence under "Antidepressants versus Placebo:" I'd suggest this should start "Several placebo-controlled analyses have examined...."	We have edited the sentence similar to what was suggested by this and another comment.
Peer reviewer #5	Introduction	P. 2, 3rd full sentence: A review of 11 trials is described. What is the reference for this?	We have added a citation to this sentence.
Peer reviewer #5	Introduction	P. 2, next paragraph, 1st sentence: "Evaluating antidepressant efficacy in relapse and remission...." seems like a misstatement. Should "relapse" be "response," instead?	This correction has been made.
Peer reviewer #5	Introduction	P. 5, 2nd last paragraph, 1st sentence: "We excluded studies that evaluated non-pharmacologic interventions...."The statement here should make it clear that the exclusion was only for studies in which the primary focus was on a non-pharmacologic intervention, but that studies involving non-pharmacological interventions as a comparator to an antidepressant medication were acceptable under the study criteria (even though it appears that none were actually found to be eligible for inclusion).	We have made this clarification as suggested since this is in fact what was intended.
Peer reviewer #6	Introduction	The introduction is very clear and concise.	Thank you.
Public Reviewer #1, American Psychiatric Association	Introduction	On p. 1, the distinctions between a contextual question and a key question are not defined until later in the document (p.11) and may be worth mentioning here. Also, the abbreviation CQ should be defined here as it is used later in the document (p. 4) without being defined.	We keep the methods related to KQ vs. CQ in the methods chapter so as not to introduce methods in the introduction of the report. We have defined contextual question as CQ where is first appears.
Public Reviewer #1, American Psychiatric Association	Introduction	On p. 2, top paragraph, it would be helpful to know how response was defined even though this is described later in the document.	The definitions of response and remission for the meta-analysis in reference were added to this paragraph.
Public Reviewer #1, American Psychiatric Association	Introduction	On p. 3, in the 3rd paragraph under Efficacy in Subgroups, it is important to note that many of the individuals in the systematic review of Boyce and colleagues had some degree of cognitive impairment, with co-occurring dementia in many of the subjects. Also, this review included individuals who had depression but did not meet MDD criteria. Each of these factors may have influenced their findings	We have Re-written the CQ and this no longer applies.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Introduction	On p. 4, in the 2nd paragraph, the Beers Criteria are discussed, including the 2015 modifications that added SSRIs and TCAs to the list of drugs to be avoided in older adults with a history of falls or fractures. It would be helpful to note the specific studies that led to this modification to the Beers Criteria so that the citations and conclusions of this review can be compared to the information used at the time the Beers Criteria were updated.	We have added citations to the text of the discussion where we address the 2015 Beers criteria and relation to our findings, which were provide by the partner American Geriatrics Society.
Public Reviewer #1, American Psychiatric Association	Introduction	p. 1 last paragraph "analysis" should be "analyses"	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Introduction	p. 2 2nd paragraph, 2nd sentence not a sentence as written, several options to change it, e.g.: "…with MDD that included only three trials,…" or " (…trials(n=1063) and found no…"	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Introduction	p. 2, 2nd paragraph, 4th sentence missing "were" – should be "rates were observed"	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Introduction	p. 2, 3rd paragraph under Antidepressants vs Each Other "A network meta-analysis suggests…" problem with subject/verb agreement throughout	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Introduction	p. 3, 1st paragraph, 3rd sentence should be "antidepressants" and last sentence should be "outcome"	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Introduction	p. 3, middle of 2nd paragraph "with wither" should be "as having"	This has been corrected.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Introduction	p. 3, 1st paragraph under Efficacy in Subgroups "regardless the comparator" should be "regardless of the comparator"	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Introduction	p. 4, 1st paragraph, "analysis" should be "analyses"	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Introduction	p.5, 1st full paragraph needs a colon before numbering and a semicolon before "2)"	This has been corrected.
Public Reviewer #1, American Psychiatric Association	Introduction	p. 6, 1st line suggest deleting "non-invasive" as there are no "invasive" psychotherapies, also a spacing issue between lines	The term "non-invasive" was removed.
Peer reviewer #1	Methods	The inclusion and exclusion criteria are appropriate for the population of interest. Please see my comments below related to flaws in the search strategy.	This comment is addressed below in "General Comments".
Peer reviewer #1	Methods	Definitions and diagnostic criteria are appropriate.	NA
Peer reviewer #1	Methods	I am uncertain about the assessment of mirtazapine (see comments below).	NA
Peer reviewer #1	Methods	Page 29, lines 33-37. This explanation may reflect sound methodology, but to a non-statistician, this sentence is off-putting. Could the test for statistical heterogeneity be explained in lay language?	The p-value for the Cochrane Q statistic is interpreted as presence of heterogeneity or not with a p-value of 0.10 used as a threshold. The I^2 quantifies the amount of heterogeneity present. The description in the methods has not been revised, as it is written in accordance with similar systematic reviews.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #3	Methods	Page 6, lines 52-56: Functional outcomes could be benefits or harms, depending on the direction of effect	We recognize that certain outcomes in this report could be either a benefit or harm and report data regardless if the result was suggesting a benefit or harm.
Peer reviewer #3	Methods	Well-done overall	Thank you.
Peer reviewer #4	Methods	It has been explicitly stated and due to the relationship of this review with the Beer's criteria feel that the age cut off was warranted	NA
Peer reviewer #5	Methods	All these aspects of the methods appear very acceptable to me. On p. 9, you could provide a bit more detail (or perhaps do some rewording) to clarify how a "third reviewer resolved disagreements through consensus," as it is not intuitively obvious how a third review achieves full consensus.	This statement was confusing and has been reworded.
Peer reviewer #6	Methods	Included medications: Since this report is focusing on first line therapy, it is not clear why studies that focus on comparisons with TCAs are included. It is well known that TCAs are less tolerable than second generation antidepressants and these medications are typically not used for depression. Also, trazodone is not a first line antidepressant, and it is rarely used to treat depression. It is not clear why this medication was included in the review.	TCAs were not included as an intervention of interest. Only studies that evaluated one of the drugs in Table 1 were eligible. These drugs could have been in comparison to a TCA. The drugs selected for inclusion focused on therapies that were considered most likely to be used in this population, according to the expert opinion of the partner, key informants, technical expert panel and public comments at the protocol development stage.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #6	Methods	The report focuses on older adults with major depressive disorder, although the definition of MDD in the inclusion criteria is broad and includes self-report in order to capture observational studies. While this strategy has merit, it would be helpful to present in the results a summary description for how MDD is determined in the included studies as we know outcomes may be related to depression severity. It is not clear that the observational study that was included met all of the exclusion criteria (comorbid seizure disorder or psych conditions)	In the ES, page ES-7, discussion, applicability, it states "Major depression was mostly diagnosed using DSM criteria. Based on scores from the Hamilton Depression Rating Scale (HAM-D) or the Montgomery-Asberg Depression Scale (MADRS) for study eligibility, the population represents those with moderate severity depression." This information is also in the full report, page 57 of 63, Chapter 4-Discussion, Applicability. The exclusion criterion in reference specifies "studies that focused enrollment solely on one of the given populations". For example a study that required MDD and comorbid seizures or other psychological conditions. Thus, the observational study meets the criteria for this review.
Peer reviewer #6	Methods	page 8, line 2: I believe this should read "September 2017"	This correction was made.
Peer reviewer #6	Methods	page 8, figure: recommend "sex" in place of "gender" throughout. Also, "withdrawal due to adverse events" seems like it needs to be deleted from the subgroup box.	We have made both changes throughout the report.
Public Reviewer #1, American Psychiatric Association	Methods	p. 8, 2nd paragraph clinicaltrials.gov is spelled wrong	This correction was made.
Public Reviewer #1, American Psychiatric Association	Methods	p. 10, 4th paragraph missing a conjunction and a close parenthesis	This correction was made.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Methods	p. 11, 1st paragraph after bullets should be “influenced”	This correction was made.
Public Reviewer #2, Evan Mayo-Wilson	Methods	The data sources used in this report are inadequate to answer the research question.	The protocol for this review was established in consultant with an extensive expert panel (TEP), public comment, and the partner. This included input on the data sources. The evidence and conclusions made in this report are a direct reflection of the evidence found following the a priori methods.
Peer reviewer #1	Results	The results section is adequately detailed, with the exception of addressing the doses actually taken by study subjects. While these data are in Appendix C, I would like to see them included in some format in the Results section, then discussed.	We added information regarding drug dose as a limitation to the applicability of this review. Most studies did not allow dosing of the full range of dose considered to be “usual dose” for older adults according to guidelines or regulatory documents. Therefore, the data in this review are not reflective of the higher but typical doses that may be used in clinical practice for efficacy. This has been added to the ES and full report. We have also added sentences in the results section summarizing typical doses used.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #1	Results	In my attached file, I suggest a few studies for investigator consideration. In particular, there are a couple in older adults with dementia. I did not search for each of the other characteristics specified in the subgroups, but feel that the report could benefit from broadening the search strategy to include terms relevant to the subgroups.	Thank you for these suggestions. These citations do not meet the inclusion criteria requiring the study to include patients 65y of age and older. We did not search for specific subgroups because the search strategy was designed to broadly to capture literature regarding older adults and major depression and one of the drugs of interest. If a study was pertinent to the overall KQ or to a subgroup, and met all inclusion criteria, the citation was included. Attempts to search for only a subgroup would lead to a narrower search and potentially missing papers of interest to this review.
Peer reviewer #1	Results	Page 33, line 17. To my way of understanding the strength of evidence based on methodology, “placebo” and “no treatment” studies should not be combined. There is a mix of evidence, some high and some low SOE. Consider separate sections for placebo-controlled and no treatment trials.	As per our methods, page 10, “We evaluated SOE separately for RCT and observational studies”. The observational study data was rated separately and the SOE rating is always accompanied with the given outcome.
Peer reviewer #1	Results	Pages 34-35, Table 4. Comparative Adverse Effects of SSRI vs. Placebo. Please add a definition of ROB to the table footnote.	We have added this abbreviation.
Peer reviewer #1	Results	Page 39, line 53. “..,data “do” not suggest…” Data is plural.	This has been corrected.
Peer reviewer #1	Results	Page 44, line 8. Rephrase the sentence beginning, “the risk of serious adverse events were less…” Perhaps, “Serious adverse events were less…”	This sentence was rephrased.
Peer reviewer #1	Results	Page 44, line 42. “Difference” instead of “different”.	This has been corrected.
Peer reviewer #1	Results	Page 46, line 43. Delete the “s” on “makes”.	This has been corrected.
Peer reviewer #1	Results	Page 48, line 16 and Page 49, line 42. “The risk of any adverse event…”was not different” instead of “were no different”.	This has been corrected.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #1	Results	Page 49, line 38. “Trazodone” is mis-spelled.	This has been corrected.
Peer reviewer #1	Results	Page 50, Table 13 Footnotes. Please add DSST and RAVLT to the footnote abbreviations.	These abbreviations were added.
Peer reviewer #1	Results	Page 52, line 6-7. Make “experience” past tense. Consider breaking into two sentences. Hip fracture and suicide do not seem to go together.	We have made these suggested revisions.
Peer reviewer #2	Results	The amount of detail is excellent. The forest plots are well-done and easy to interpret.	Thank you.
Peer reviewer #3	Results	See general comments above. Otherwise, clear and seems complete	Thank you.
Peer reviewer #4	Results	Felt that the representation of the data was clear and presented well.	Thank you.
Peer reviewer #4	Results	Here are some other studies that were discussed among Beers Criteria members on this topic. Torvinen-Kiiskinen S, Tolppanen AM, Koponen M, et al. Antidepressan t use and risk of hip fractures among community dwelling persons with and without Alzheimer’s disease. Int J Geriatr Psychiatry. 2017;32(12):e107-e115. Macri JC, Iaboni A, Kirkham JG, et al. Association between antidepressants and fall-related injuries among long-term care residents. Am J Geriatr Psychiatry. 2017;25(12):1326-1336. Naples et al. Am J Geriatr Psychiatry. 2016 December ; 24(12): 1221–1227. doi:10.1016/j.jagp.2016.08.008.	Thank you for sharing these citations. They do not meet inclusion criteria for this review thus were not added.
Peer reviewer #5	Results	The “layered” approach seems quite effective, in which the Key Points are presented first, then progressively greater detail about the various analyses and their outcomes. The figures and tables are adequate and quite helpful. I have no comments with respect to the particular studies that were included versus excluded. P. 14, 2nd indented bullet under Key Points: “evidence was limited in number” Number of what? “limited in the number of studies?”	The evidence was limited in number of trials and reported outcomes. This has been added to complete the sentence.
Peer reviewer #5	Results	P. 14, 2nd sentence under “Study Characteristics;” perhaps better to say “Fragus et al. investigated exclusively patients with heart failure....”	This change has been made.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #5	Results	P. 31, 1st sentence under "Results:" change "as evidence of" to "according to"	This change has been made.
Peer reviewer #5	Results	P. 33, KQ2: insert "by" between "e.g.," and "age"	This change has been made.
Peer reviewer #5	Results	P. 33, 2nd paragraph under "Age:" the abbreviations here for "weeks" and "years" are inconsistent with the style of the rest of the report.	These abbreviations were spelled out, consistent with the rest of the report.
Peer reviewer #6	Results	As a general comment for the Key Points for each medication/class, I found it helpful when the number of studies was mentioned (which is inconsistent throughout sections). It would be helpful if the estimates from the metaanalyses results was provided in the bullet points.	We have added the number of trials throughout the key points as suggested.
Peer reviewer #6	Results	SSRIs key points (page 33). 3 -missing key point regarding SSRI-SSRI	We did not feel the data warranted a key point thus none was presented.
Peer reviewer #6	Results	Page 34, Table 6: wrong citation is given for the Coupland study.	This has been corrected.
Peer reviewer #6	Results	Page 36, Figure 3: Would be helpful to add the results of the meta-analysis of any adverse event (as done with SNRI)	We have modified Figure 3 as suggested.
Peer reviewer #6	Results	Page 40, line 10-11: Confidence intervals available from original study?	No confidence interval was reported in the original study.
Peer reviewer #6	Results	Page 40, SRNI key points: Is the first bullet referring to the results from meta-analysis? If so, it would be helpful to add "meta-analysis or x RCTs.... (as was done with SSRIs.	We have made this revision as suggested.
Peer reviewer #6	Results	Page 43, figure 5, should make some comment regarding the results for serious adverse events for acute phase.	Comments regarding serious adverse events and withdrawal due to adverse events follow the figure.
Peer reviewer #6	Results	Summarize SNRI-SSRI under key points? Could fit under SSRI or SNRI, but is probably worth summarizing at the key point level.	We did not feel the data warranted a key point thus none was presented.
Peer reviewer #6	Results	Add Bupropion to key points.	We did not feel the data warranted a key point thus none was presented.
Peer reviewer #6	Results	page 53, lines 12 and paragraph "Risk Factors for Falling..." May consider adding confidence intervals for estimates.	Confidence intervals were not reported in the original study.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Results	On p. 15, Table 4 (and subsequent tables), it would again be helpful to have supplementary tables that summarize comparisons and outcomes by drug, not just by drug class vs. broad categories of comparators. In addition, it would be helpful to have tables that note average or median drug doses and the frequency and, if available, the relative severity of each type of adverse effect with each medication. Simply knowing that an increased risk exists is not clinically useful to patients or clinicians without knowing the likelihood that an adverse effect will occur for a particular patient or with a typical medication dose. When tables (and text) report medication-related differences in continuous variables (e.g., blood pressure, serum sodium), it would be useful to know if these statistically significant differences are clinically significant.	The pre-specified methods determined the comparisons would be made per class for SSRI and SNRIs. Given the limited amount of drugs within each class that were represented, we present the names of the SSRIs and SNRIs with pooled results. When data for a single drug was available, the drug level comparison is presented (ex. duloxetine vs. placebo is the primary data for SNRIs vs. placebo and is reported in such way). Table 1 and 2 in Appendix C report dose information that was reported in each study in the "intervention/comparisons" column. Not all studies reported mean dose or percentage of patients taking each possible dose. When the data was reported we have recorded it there. We have also added summary data regarding dose within the "study characteristics" paragraph that begins each results section for the unique drug comparisons. We recognize the importance of interpreting continuous outcomes within the context of minimally important difference. Standard minimally important differences are not readily defined for the continuous outcomes found to be significant in this report (such as serum sodium) and thus none were provided.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Results	On p. 18, the 1st full paragraph describes the study of Coupland et al. (2011). Given the unique features of this study and the importance of its findings to the systematic review, it would be helpful to discuss the study design and its strengths and limitations in more detail. In particular, in interpreting the study results, it is essential to know that comparison subjects not taking an antidepressant had had depression diagnosed at some point, but there are still likely to have been differences between individuals taking an antidepressant and those not taking an antidepressant that could influence the findings (e.g., depression severity, concomitant physical illness, prior medication benefits or intolerance). For example, individuals with severe depression may have been more likely to receive an antidepressant but also more likely to die by suicide without any direct causal link between antidepressant use and suicide risk. Similarly, other factors may have influenced the choice of a specific antidepressant and affected the relative effects of specific medications on outcomes. Thus, the authors conclude "As this is an observational study, it is susceptible to confounding by indication, channeling bias, and residual confounding, so differences in characteristics between patients prescribed different antidepressant drugs that could account for some of the associations between the drugs and the adverse outcomes may remain." The study also found that SSRIs and several other antidepressants were associated with an increased risk of several adverse outcomes as compared with TCAs. As this is contrary to expectations, it may suggest differences in relative dose or other factors are present that would influence other study outcomes. These implications of the Coupland et al. study should also be incorporated into the later sections of the document (e.g., pp. 35 ff).	We agree with the comments made and have incorporated these suggestions into the "Discussion", "Limitations", second-to last paragraph where we had already introduced limitations regarding this cohort study.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Results	On p. 30, the findings with trazodone are described. It would be particularly important to note the mean or median dose of trazodone that was prescribed because doses for antidepressant effects are much higher than those for sedative properties (e.g., as sometimes used to treat insomnia).	Table 1 and 2 in Appendix C report dose information that was reported in each study in the "intervention/comparisons" column. Not all studies reported mean dose or percentage of patients taking each possible dose. When the data was reported we have recorded it there. We have also added summary data regarding dose within the "study characteristics" paragraph that begins each results section for the unique drug comparisons.
Public Reviewer #1, American Psychiatric Association	Results	On p. 33, the bottom of the 1st paragraph under Age, notes that "commonly reported adverse events that led to withdrawal in both groups included nausea, anxiety and depression;" however, this is confusing as depression was the reason that individuals were being treated and seems difficult to conceptualize as an adverse event.	We can see how this may be counter intuitive but we are reporting what was reported in the study, and in fact depression was a reported adverse event.
Public Reviewer #1, American Psychiatric Association	Results	p.12, 3rd line, the sentence that begins "Although we…" should be double-checked for wording as it seems confusing.	We have reworded this sentence as it was confusing.
Public Reviewer #1, American Psychiatric Association	Results	p. 14, last bullet under SSRIs v placebo there is an "SOE" after hyponatremia but it doesn't say what level and in the table it says it wasn't done – probably should be deleted	This correction has been made.
Public Reviewer #1, American Psychiatric Association	Results	p. 17, 4th line "singe" should be "single"	This correction has been made.
Public Reviewer #1, American Psychiatric Association	Results	p. 17, 2nd paragraph the sentence that begins "Fragus et al.…" should be double-checked for wording as it seems confusing.	We have edited this sentence based on this comment and one from another reviewer.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Results	p. 17, 3rd paragraph “events” should be “event”	This correction has been made.
Public Reviewer #1, American Psychiatric Association	Results	p. 18, 1st paragraph “effects” should be “events”	This correction has been made.
Public Reviewer #1, American Psychiatric Association	Results	2nd to last sentence needs a comma after “suicide” (it is there in last sentence). This issue with dependent clauses at the start of a sentence missing the comma is throughout the document. Paragraph under study characteristics should be “TCAs”	This correction has been made.
Public Reviewer #1, American Psychiatric Association	Results	p. 19, first paragraph, sentence starting “Two studies…” – this is really awkward and incomplete as written	This sentence has been corrected.
Public Reviewer #1, American Psychiatric Association	Results	p. 20, 2nd sentence should start “A single study…” and it should be SNRI (see page 21 as well) last paragraph should be “data did”, not “data does”	These corrections have been made.
Public Reviewer #1, American Psychiatric Association	Results	p. 25, 1st paragraph last sentence needs to be reworked. 3rd paragraph the start of the sentence up to the colon needs to be reworked. 6th paragraph “different” should be “difference”	These corrections have been made.
Public Reviewer #1, American Psychiatric Association	Results	p. 26, last line, “low both trials” should be “low in both trials”.	This correction has been made.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Results	p. 27, Results “makes” should be “make”	This correction has been made.
Public Reviewer #1, American Psychiatric Association	Results	p. 28, under Trazodone no “and” between “mortality and suicide”	This correction has been made.
Public Reviewer #1, American Psychiatric Association	Results	p. 29, top of page delete “while” if using semicolon	This correction has been made.
Public Reviewer #1, American Psychiatric Association	Results	p. 30, last paragraph “Razodone” and “mirtazapine” should both be “trazodone”	This correction has been made.
Public Reviewer #1, American Psychiatric Association	Results	p. 31, 1st sentence under Results should be double-checked for wording as it seems confusing	This correction has been made.
Public Reviewer #1, American Psychiatric Association	Results	p. 32, 1st sentence under Results, “Vortioxetine decreased risk” should be “Vortioxetine had a lower risk…”	This sentence is accurate as written and this format is consistent throughout the report.
Public Reviewer #1, American Psychiatric Association	Results	p. 33, 1st paragraph should be “experienced “ 1st paragraph under Age, 4th sentence “Withdrawal was similar…” Last paragraph all the years, weeks, etc. need to be written out	These corrections have been made.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Results	p. 34, 2nd line of 1st full paragraph "was two randomized phases" should be "was conducted in two randomized phases"	This change has been made.
Public Reviewer #1, American Psychiatric Association	Results	p. 34, last line of same paragraph should be "odds ratio was" or "odds were" (probably the first) and see last paragraph as well	This correction has been made.
Peer reviewer #1	Discussion	The implications of the major findings are clearly stated and the review of studies adequately described. See my attached file for missing literature. In the Research Gap section on page 57, the investigators conclude that there were no data to evaluate subgroups of interest except age. I believe they missed studies that could find additional literature to evaluate subgroups further.	The provided citations did not meet the criterion of studying a sample of patients 65y of age and older. We did not search for specific subgroups because the search strategy was designed to broadly to capture literature regarding older adults and major depression and one of the drugs of interest. If a study was pertinent to the overall KQ or to a subgroup, and met all inclusion criteria, the citation was included. Attempts to search for only a subgroup would lead to a narrower search and potentially missing papers of interest to this review.
Peer reviewer #1	Discussion	The only section which addressed potential future research was the section called Research Gaps, unless I missed something. This section could be expanded to include recommendations to study the subgroups	The need for research to fill the gaps in subgroup data is presented in the "evidence gaps and future research needs" in the "discussion".
Peer reviewer #1	Discussion	Page 54, line 39. Delete "both".	This correction has been made.
Peer reviewer #1	Discussion	Page 55, line 4 and 19. Remember "data" is plural. Modify accordingly in those three sentences: "data do not suggest"; "data… were scarce"; and "data suggest"	We reviewed the full report for the singular vs. plural use of the word data and made any necessary corrections.
Peer reviewer #1	Discussion	Page 56, line 8. Consider changing to, "This cohort study was not included in the prior review."	This correction has been made.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #1	Discussion	Page 56, 57. Limitations. A large limitation of this report is that the dose of antidepressants was not assessed, or if assessed, not included in the report. In older persons, it is quite typical that antidepressants cannot be titrated to "therapeutic ranges" due to limiting side effects. One of the most frequent issue, I identified while in nursing facility practice, was addition of subtherapeutic doses of antidepressants that were never titrated to the therapeutic range. Therefore, in some studies the questionable efficacy may be attributable to subtherapeutic doses and the report of adverse effects may be understated. In fact, per Appendix C. Evidence tables, several studies report a proportion of study cohorts on subtherapeutic doses. Orrell M; Collins E; Shergill S; Katona C. Management of depression in the elderly by general practitioners: Antidepressants. Family Practice [Fam Pract], ISSN: 0263-2136, 1995;12 (1): 5-11. Wells KB; Norquist G; Benjamin B; Rogers W; Kahn K; Brook R. Quality of antidepressant medications prescribed8. General Hospital Psychiatry [Gen Hosp Psychiatry], ISSN: 0163-8343, 1994;16 (1):4-15. PMID: 8039682 Rojas-Fernandez C; Thomas VS; Carver D; Tonks R. Suboptimal use of antidepressants in the elderly8 Clinical Therapeutics [Clin Ther], ISSN: 0149-2918, 1999;21 (11):1937-50.	We collected data regarding doses of antidepressants used in each included trial, which is reported in Appendix C Table 1-2. We have added information to the ES and report "discussion", "applicability" section regarding the limitation in applicability of this review related to drug dose. Almost all antidepressants were dosed in the lower half of the "usual dose in older adults" suggested by guidelines. Therefore data are not reflective of higher, but possibly typical doses that clinicians may prescribe for better drug efficacy.
Peer reviewer #1	Discussion	Page 57, line 9. This sentence has an unnecessary, double-use of "adjustment" and "adjusted for".	This error has been corrected.
Peer reviewer #1	Discussion	Page 57, line 34. Descisionmakers is misspelled and should be hyphenated.	Per the AHRQ publication guide, decsionmakers is without a hyphen.
Peer reviewer #1	Discussion	Page 57, line 36. Data is plural, thus, "… there were no.."	This correction has been made.
Peer reviewer #1	Discussion	Page 57, line 41. This sentence doesn't make sense. Are you attempting a to compare antidepressant-antidepressant to antidepressant-placebo trials? If so consider, "Limited data were more available for direct comparisons of antidepressants than for comparisons to placebo".	Yes, this correction has been made.
Peer reviewer #1	Discussion	Page 57, line 42. Should "are" be "is" as the sentence subject is singular?	Yes, this correction has been made.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #2	Discussion	No, the implications are not balanced by a further discussion of confounding by indication and how clinicians may utilize these data to IMPROVE outcomes for older adults with depression who require treatment with antidepressant pharmacotherapy. I fear the results of reviews such as this contribute to fear of prescribing among primary care physicians, furthering the epidemic of un-treated depression in late-life	We have added a paragraph to the discussion, findings relative to what is known section, regarding the concept of confounding by indication and how clinicians can apply the findings of this report with consideration of this concept.
Peer reviewer #3	Discussion	See general comments above. Otherwise, clear and cogent discussion.	Thank you.
Peer reviewer #4	Discussion	Appreciated the explicit recommendation on studies being powered to detect differences	Thank you.
Peer reviewer #5	Discussion	The conclusions are quite comprehensively outlined and explained, and appear consistent with what the data support. Limitations are adequately discussed, and needs for future research are clearly stated. P. 35, bottom paragraph: It is not clear here what therapy/therapies were compared to placebo when evaluating cognitive impairment.	We have revised this paragraph as the intervention and comparator in reference was not always clear.
Peer reviewer #5	Discussion	Pp. 37-38: A number of statements under "Limitations" and "Research Gaps" are rather garbled and unclear in meaning. The following are some suggestions about possible rewordings: "Applicability" paragraph, line 6: change "with the severity" to "and the severity"	This change has been made.
Peer reviewer #5	Discussion	Same paragraph, line 9: It is unclear to say "all the results should not be extrapolated…" Do you mean to say simply that "thus none of the results should be extrapolated….?"	We removed the word "all" as that was confusing.
Peer reviewer #5	Discussion	"Limitations," 1st paragraph, lines 1-2: "No evidence was found for a number of the interventions of interest in this review, nor for many of…." Same paragraph, lines 3-4: "Most of the available data featured comparisons to placebo and few direct data were found to inform comparative harms…." Same paragraph, lines 5-6: "…small number of trials and limited sample sizes…."	These changes have been made.
Peer reviewer #5	Discussion	Next paragraph, line 2: "Many outcomes suffered from the rareness of events where, for example, only one or two events occurred…" "Limitations," 3rd paragraph, lines 4-6: "although adjustments were made for dementia, antihypertensives, sedatives and hypnotics, and prior falls, other factors such as hypotension were not adjusted for."	These changes have been made.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #5	Discussion	"Research Gaps," paragraph 1, lines 5-7: "There were many outcomes….eligible studies, yet these are important to clinicians and decisionmakers according to the key informants, technical expert panelists and partners on this project who helped shape the list of outcomes of interest."	These changes have been made.
Peer reviewer #5	Discussion	Same paragraph, line 9: change "shaped" to "identified" Next line: Change clause beginning "they are globally important" to "information about their influence is highly important for the care of older depressed patients."	These changes have been made.
Peer reviewer #5	Discussion	Next paragraph, lines 1-3: "Aside from comparisons to placebo, limited data were available for direct comparisons among antidepressants. While a decision must first be made as to whether or not to treat MDD with antidepressants, with more severe depression the more telling decision is likely to be which antidepressant to prescribe…"	These changes have been made.
Peer reviewer #5	Discussion	P. 38, "Conclusions," 2nd sentence: "Further characterization of the comparative safety of first-line antidepressants is difficult because…."	These changes have been made.
Peer reviewer #6	Discussion	Limitations: Please include issues discussed in the general remarks regarding the limitations of the Coupland study. Although several factors were adjusted for in the models, covariates were derived from administrative data which has limitations.	We have added additional limitation of this study in the "discussion", "limitations", second to the last paragraph, based on the comments received from peer and public review.
Peer reviewer #6	Discussion	Could make more explicit that people that tolerate treatment were randomized, thus adverse events were less likely to occur with continued treatment (rather than less likely to be "recorded")	We have made this revision in the discussion.
Peer reviewer #6	Discussion	Research gaps: Can the authors provide any direction for future research in this area? Since observational studies are likely to be the source to evaluate many of these outcomes, can the authors offer any advice to move the field forward ? This would be a great place to add a sentence that RCTs need to include outcomes that are important for the care of older adults and that studies should include patients with MCC. The evidence gaps section (page 18) pointed out some specific research needs with regard to important outcomes and subgroup analyses that that could be mentioned in this section.	We have added the suggested sentence information into the future research needs of the discussion.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Discussion	On p. 36, in the discussion of findings in relationship to what is already known under KQ2, it would be helpful to place these results in context as they relate to the American Geriatrics Society Beers Criteria. The inclusion of many of these medications in the Beers Criteria is intended to prevent harms of treatment. Nevertheless, the Beers Criteria are also used in clinical decision support systems and by pharmacy benefit managers, which can have the unintended consequence of reducing or restricting use of these medications even when they are clinically needed to treat depressive symptoms. If these findings are intended to inform application of the Beers criteria, it is crucial that the limitations of these data be emphasized. Given the limited confidence in many of these findings in this group of patients, modification of the Beers Criteria or a footnote to the Beers Criteria that emphasizes the poor quality of this data on antidepressant harms may be warranted.	We have added a paragraph in the "discussion" section "findings in relationship to what is already known" regarding the 2015 Beers Criteria. However, since these may (or may not) change in the near future with a planned update, it may no longer be clinically relevant to compare our findings to the 2015 criteria. How the partner decides to utilize the findings presented in this report for future Beers Criteria is outside of the scope of our EPC and the statements in this report.
Public Reviewer #1, American Psychiatric Association	Discussion	p. 36, 4th full paragraph, notes that "sertraline and fluoxetine caused less dizziness than venlafaxine"; however, it is not clear whether these medications caused dizziness less often or whether they were associated with a lesser severity of dizziness. (This type of wording should be assessed and addressed throughout the document.)	We have corrected this statement and also checked the use of the word "less" throughout the document. We have replaced "less" with "fewer" where appropriate.
Public Reviewer #1, American Psychiatric Association	Discussion	p. 36 heading should "what is" be capitalized?	This correction has been made.
Public Reviewer #1, American Psychiatric Association	Discussion	p. 36, 4th paragraph under Findings "This review" should probably be changed to "That review". This review implied this review from AHRQ I was reading	This correction has been made.
Public Reviewer #1, American Psychiatric Association	Discussion	p. 37, top 2 paragraphs there are 2 places that say "this findings" – needs to agree "this finding" or "these findings"	These corrections have been made.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Discussion	p. 37, under Applicability The sentence on DSM criteria is awkward – too many "was"s	We removed one "was".
Public Reviewer #1, American Psychiatric Association	Discussion	p.38, 1st full paragraph There are several sentences that are awkward or could benefit from reworking with conjunctions, commas, etc.	We have made revisions to this paragraph.
Public Reviewer #1, American Psychiatric Association	Discussion	p. 38, 1st paragraph under Research Gaps, the sentence that reads "We found no evidence for several therapies of interest other than SSRI or SNRI," should be double-checked for wording as it seems confusing. 4th sentence delete "but" or "yet"	We have made these changes.
Public Reviewer #1, American Psychiatric Association	Discussion	p. 38, 2st paragraph under Research Gaps the first two sentences should be double-checked for wording as they seem confusing.	We have made revisions based on other comments and suggestions.
Public Reviewer #1, American Psychiatric Association	Discussion	p. 38, last line delete "help"	This change has been made.
Public Reviewer #1, American Psychiatric Association	Conclusion	On p. 36, in the discussion of findings in relationship to what is already known under KQ2, it would be useful to know whether these findings on adverse events are similar to findings seen in younger individuals. Although the review is appropriately limited in scope to individuals over age 65 years of age, similar findings in younger and older individuals would add credibility to the conclusions. Differences in the findings would be less compelling because older individuals might be expected to experience different or greater risks of medications.	A paragraph has been added to the discussion, findings relative to what is known, regarding antidepressant harms in younger patients relative to our findings.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	Conclusion	On p. 38, in the conclusions, it is essential to emphasize that the majority of these studies were done in individuals with few concomitant physical conditions and that factors such as drug-drug interactions and pharmacokinetic properties must be taken into consideration in the context of the individual patient before relying on this comparative harms data to make prescribing decisions. This is especially true for fluoxetine, which appears preferable from the conclusions of this review. Nevertheless, it can be quite problematic and often is not recommended for use in older individuals because of its high levels of plasma protein binding, metabolism via via CYP2C19 and CYP2D6, its long half-life (4-6 days with chronic use), and its even longer halflife active metabolite, norfluoxetine (4-16 days).	In the "Discussion" chapter under "applicability" we have added "Resulting drug-drug interactions and pharmacokinetic changes must be taken into consideration when prescribing antidepressants."
Peer Reviewer #1	Appendix	Page A-1. Appendix A, Search Strategy. While the analytic framework includes subgroups with risk of falls/fractures, dementia/cognitive decline, nursing home setting, polypharmacy, and frailty, none of these terms were included in your search strategy. How were older adults with subgroup characteristics identified? I am concerned that you missed a significant body of relevant literature by not including these search terms. Given that these are characteristics of interest, I had hoped to see each addressed in Table summaries and the results section	We did not search for specific subgroups because the search strategy was designed broadly to capture literature regarding older adults and major depression and one of the drugs of interest. If a study was pertinent to the overall KQ or to a subgroup, and met all inclusion criteria, the citation was included.
Peer reviewer #1	General	The report is clinically meaningful with information helpful to clinicians interested in initiating antidepressant therapy in older adults. However, I do not feel the target population, while appropriately defined, was sought in the search strategy given that search terms pertaining to the subgroups were not included, i.e., frailty, dementia, Alzheimer's disease, risk of fall/fracture, etc. Thus, I am concerned that a significant body of literature may not have been found.	We did not search for specific subgroups because the search strategy was designed to broadly to capture literature regarding older adults and major depression and one of the drugs of interest. If a study was pertinent to the overall KQ or to a subgroup, and met all inclusion criteria, the citation was included. Attempts to search for only a subgroup would lead to a narrower search and potentially missing papers of interest to this review.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #1	General	KQ 1 is explicitly stated, although, I am concerned that because of some under dosing, the report may fail to achieve its goal of answering KQ 1.	We have added dosing information as requested in another comment from this reviewer.
Peer reviewer #1	General	KQ 2. Is not explicitly stated and because of that, I feel that we have no more information on the treatment of depression in older adults with dementia, frailty, risk of falls/fracture, etc., than we did prior to the development of this review.	NA
Peer reviewer #2	General	The report is timely, scholarly, and clinically meaningful. The key questions are explicitly stated, and the population (outpatient with moderate depression) is reasonably well-described. Including a description of cognitive status of the participants in the trials would be useful. The majority of SOE is moderate or low. My biggest concern is that there is little to no mention of confounding by indication. For example, depression in an of itself in older adults a risk for falls. While the authors state that the studies were not designed or powered to detect harm, further statements that balances the risks of untreated depression on falls and cognition would help to place these analyses in a more clinical context.	We have added a paragraph to the discussion, findings relative to what is known section, regarding the concept of confounding by indication and how clinicians can apply the findings of this report with consideration of this concept.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #3	General	This is a well-written, comprehensive, and fair review of the literature on this topic. While overall the review is very good, several areas may benefit from further exploration, emphasis, or clarification, as follows: Throughout the report, many of the key points bullets and summary statements focus on the statistically significant findings. This has the potential to detract from the key message that for most of the drugs of interest, many of the outcomes were not studied at all, or were evaluated in a very limited way (due to paucity of studies, limited duration of followup, non-representative populations, etc.). This key message is acknowledged but seems relatively underemphasized relative to its importance in the summary statements, although it is explored more thoroughly in the discussion…. This may inadvertently lead casual readers to conclude, for example, that SSRIs and SNRIs are particularly harmful because they are the main ones noted to have significant findings of harm. Perhaps the more accurate message to emphasize is that certain drugs are demonstrated to increase risk of adverse events; little else is known about many of the drugs of interest, and thus we lack evidence to meaningfully compare drugs in this class and assess their relative harms. Ultimately this is the most important clinical question, because clinicians are likely most interested in knowing if one drug is more effective and/or more harmful than the others, and thus should be preferentially prescribed or avoided.	We agree with the interpretation of this reviewer. We have strengthened the last key message bullet, to emphasize that little else can be stated about other antidepressants and harms because we lack comparative harms evidence.
Peer reviewer #3	General	Throughout the report, and particularly in the summary statements, he contrast of Aes vs. withdrawals due to Aes is a little confusing, since they sound so similar. It might help to wordsmith a little to ensure that readers appreciate the difference between these 2 outcomes	Throughout the report we are careful and use the terms “adverse events” or “withdrawal due to adverse events”. The report methods define these outcomes explicitly.
Peer reviewer #3	General	It is unclear why the outcomes of “any adverse events” and “withdrawals due to adverse events” are emphasized more than other outcomes throughout the report – it would help to explicitly clarify why.	In this report little data were available for specific harms. The most reliable data were available for these two more generic harms of “adverse events” and “withdrawal due to adverse events”. Thus, this was the emphasis of the key messages.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #3	General	In the discussion it would help to add something brief about the extent to which studies were industry-funded and the potential of funding sources to induce bias.	11 of the 19 RCTs reported industry sponsorship. We have listed the suggested information in the "limitations" section of the "discussion"
Peer reviewer #3	General	The key questions are appropriately stated, and the target population and primary audience made clear enough. The report and its findings have clinical significance, but would be more meaningful to clinicians and decisionmakers if: a) you would consistently state the findings in terms of, or relative to, the medications about the use of which clinicians must make decisions (e.g., it is more readily meaningful to present findings in the form that "medication A was associated with more of harm X than placebo was," rather than that "the comparison of medication A with placebo favored placebo;" and b) you could present effect size information about the statistically significant findings, as well as the strength of the evidence. In the case of these findings, that might mean estimating the Number Needed to Harm (NNH), analogously to how the report indicates the NNT for efficacy in the first paragraph under "Antidepressants versus Placebo" in the "Contextual Question" section on p. 2.	We have adjusted the wording, particularly in the key points, to improve clarity. We have add NNH to the outcomes that were graded for strength of evidence where there were statistically significant differences found.
Peer reviewer #4	General	This report is clinically meaningful and in light that the audience has been explicitly defined.. it has limited the number of trials that can be included.	NA
Peer reviewer #4	General	Also, feel that the authors have adequately represented and researched key questions based on discussions of the TEP	Thank you.
Peer reviewer #6	General	Given similar efficacy among first line antidepressants, having information about differences in adverse events among agents is important for selection of therapy. Thus, this report is clinically relevant and meaningful. The target population and audience are explicitly defined and the key questions are relevant.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #1, American Psychiatric Association	General	This AHRQ systematic review addresses an important clinical issue given the prevalence of depression in older adults and the associated need for antidepressant medication treatment in many individuals. Also, as pointed out in the review objectives/rationale, treatment is often selected based on factors such as relative likelihood of adverse events. Thus, we concur with the value of this topic for clinicians, patients, and other stakeholders.	Thank you.
Public Reviewer #2, Evan Mayo-Wilson	General	Most adverse events in clinical trials are not reported in public sources such as journal articles. Thus, it is incorrect to claim that drugs does not cause adverse events based on public data. The claim that "but not selective serotonin reuptake inhibitors (SSRIs)" suggests that SSRIs don't cause AEs during acute treatment, which is certainly false.	The claims in this report are based on the evidence base included and the data reported in those included studies.
Peer reviewer #2	Clarity and Usability	The conclusions could be presented in a more balanced fashion to improve clinical relevance.	We have made modifications to the conclusion based on the totality of comments and believe it is improved.
Peer reviewer #4	Clarity and Usability	Believe they do provide additional information. Unfortunately, the findings really illustrate the importance of reframing how we look at data for medication surveillance and older adults.	NA
Peer reviewer #5	Clarity and Usability	The report seems well organized according to a format that appears to have been fairly well established in advance, as is perhaps required. The findings and conclusions are somewhat relevant to policy and practice decisions, but as indicated in the report, the scope of the new knowledge produced is disappointingly limited because of the scarcity of research evidence on which findings/conclusions could be based. The implications of the project thus seem to be primarily that it has highlighted the needs for more research that is specifically oriented toward the questions that this review had hoped to address. As noted previously, providing effect size information (such as NNH) in addition to the strength of evidence (SOE) information would help clinicians formulate the risk-benefit ratios applicable to the use of antidepressant medications in the older adult population.	We have add NNH to the outcomes that were graded for strength of evidence where there were statistically significant difference found.
Peer reviewer #6	Clarity and Usability	The report is clear and well structured. The main points are clearly presented.	Thank you.